universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

„Identification of potential Ribonuclease A inhibitors by in silico screening“

verfasst von

Katsiaryna Bulyha

angestrebter akademischer Grad

Magistra der Pharmazie (Mag.pharm.)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 449
Studienrichtung lt. Studienblatt:	Diplomstudium Pharmazie
Betreut von:	Univ.-Prof. Mag. Dr. Gerhard Ecker

Danksagung

An dieser Stelle möchte ich mich ganz besonders bei meinem Betreuer, Herrn Univ. Prof. Dr. Gerhard Ecker, bedanken. Ohne sein offenes Ohr für alle Fragen, seine Hilfestellung und Feedback wäre das Verfassen dieser Arbeit nicht möglich gewesen.

Für das ausgezeichnete Arbeitsklima möchte ich mich außerdem bei der gesamten Pharmacoinformatics Research Group bedanken. Danke Andi, Doris, Michael, Marta, Lars, Amir, Floriane, Daria, Eva und Daniela für die Unterstützung und den Rat bei allen Problemen und Herausforderungen.

Mein spezieller Dank gilt meiner Familie, meinen Eltern Elena und Sergei und meinem Bruder Alex, für ihren Beistand in allen Situationen.

Ebenfalls möchte ich meinen Freunden und Kollegen danken, die mich durch das gesamte Studium begleitet haben.

Danke

Table of Contents

I. BIOLOGICAL BACKGROUND 9

A. RIBONUCLEASE A 9
B. OBESITY 11
Introduction 11
Facts according to WHO 12
Causes and mechanisms 12
Effects on health 15
Treatment 17
Ribonuclease A as a new target for treating obesity 20

II. AIM OF STUDY 21

III. COMPUTATIONAL BACKGROUND 22

A. COMPUTER-AIDED DRUG DESIGN 22
Structure-based design 22
B. SEQUENCE ALIGNMENT 23
Clustal W 23
C. PHARMACOPHORE MODELING 25
Introduction 25
Structure-based Pharmacophore Generation 26
Virtual Screening 27
D. CHEMGPS-NP 29
Work-flow 29
E. DOCKING 32

IV. MATERIAL AND METHODS 34

A. SEQUENCE ALIGNMENT 34
B. DATABASE SEARCH 35
C. PHARMACOPHORE MODEL AND SCREENING 36
Generation of Pharmacophore Model 36
Screening a Commercial Database 36
D. CHEMGPS-NP 37
E. DOCKING 37
Ligand and Protein Preparation 37
Docking workflow 38
Redocking - Validation of the method 39

Scoring *39*
Outcome Processing *39*

V. RESULTS AND DISCUSSION 40

A. SEQUENCE ALIGNMENT 40
B. DATABASE SEARCH 42
C. PHARMACOPHORE MODEL AND SCREENING 44
Generation of a structure based Pharmacophore Model *44*
Screening a Commercial Database *47*
D. CHEMGPS-NP 48
E. DOCKING 53
Redocking *53*
Docking results *56*
Compound Selection for Experimental Testing *56*

VI. CONCLUSION AND OUTLOOK 57

VII. REFERENCES 58

VIII. ANNEX 63

A. HIT COMPOUNDS (ID NUMBERS FROM LIFECHEMICALS) 63
B. ABSTRACT 80
C. ZUSAMMENFASSUNG 81
D. LIST OF ABBREVIATIONS 82
E. LIST OF FIGURES 83
F. CURRICULUM VITAE 85

I. Biological Background

A. Ribonuclease A

Synonyms: Ribonuclease pancreatic, Ribonuclease 1

Abbreviations: RNase 1, RNase A

Ribonuclease A is an endonuclease that catalyzes the cleavage of RNA on the 3' side of pyrimidine nucleotides. It acts on single-stranded as well as on double-stranded RNA [Uni-Prot © 2002–2013]. RNase A cleaves the RNA by transphosphorylation and hydrolysis reaction. Several subsites exist within the catalytic center of RNase A, where RNA binds. The subsitese are labeled as P, R and B that are binding sites for phosphate, ribose and nucleobase of RNA. The main subsite is considered P_1, while subsites B_1 and B_2 on each side of P_1 are partially conserved. However, B_1 binds pyrimidines, while B_2 prefers purines [Samanta et al. 2011]. RNase A comprises four disulfide bonds as well as two *cis*-amide bonds, see Figure 2 [Marashall et al. 2007]

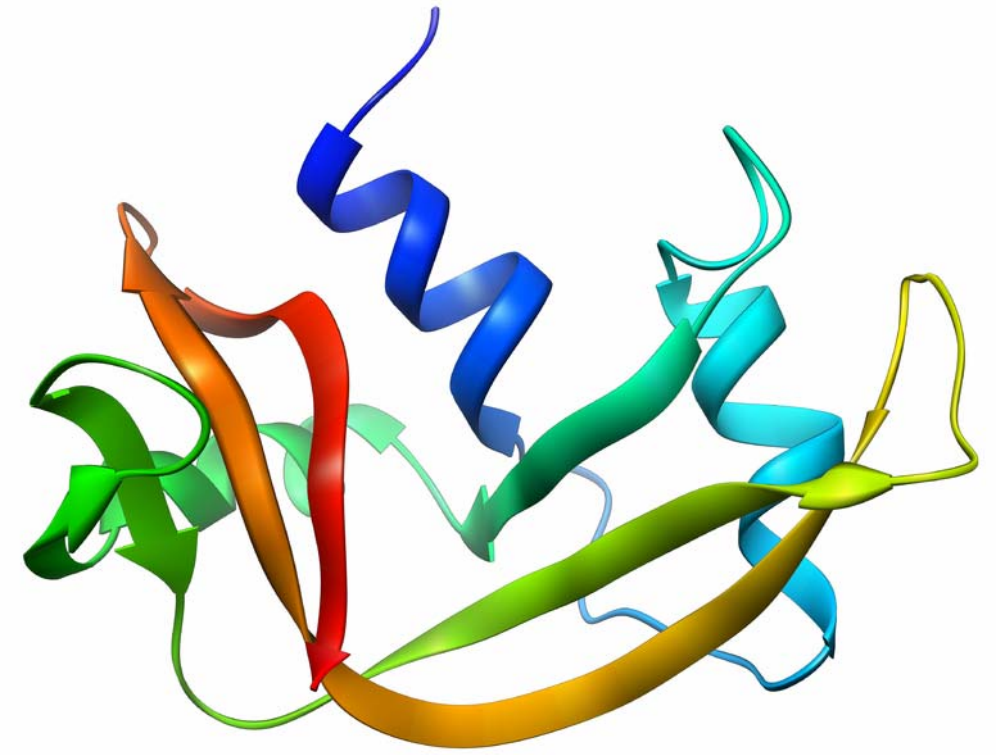


Figure 1. 3D Structure of RNase A in solution by nuclear magnetic resonance, reproduced from Santoro et al. 1993

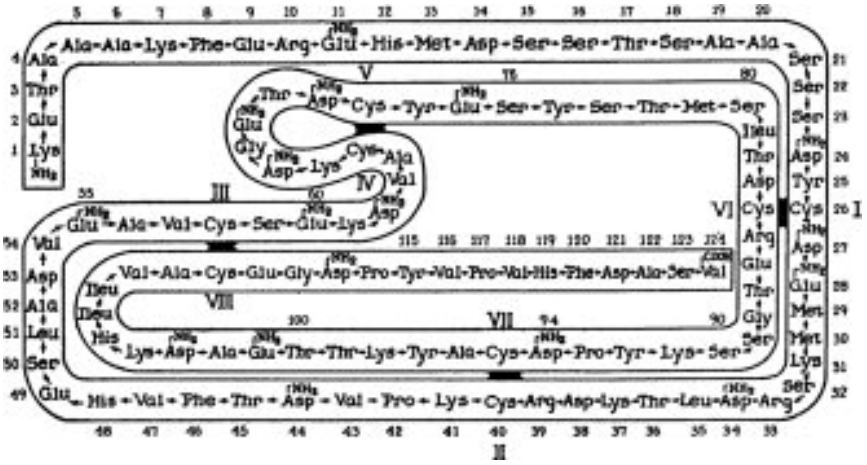


Figure 2. Sequence diagram of bovine pancreatic ribonuclease A, reproduced from Marashall et al. 2007

RNase A belongs to the Ribonuclease A superfamily. The superfamily was named after the bovine pancreatic RNase because it is one of the first enzymes to be characterized. RNases were identified in many vertebrates, including humans, mice and cows. The RNase A superfamily does not share sequence homology with any other RNase families such as RNase H, RNase III, and RNase P [Goo, Cho 2013].

At present 13 human Ribonuclease genes, RNAse 1 to 13, (Figure 3) and 11 mouse EAR genes (EAR 1 to 7, EAR 10 to 12 and EAR 14) are known.

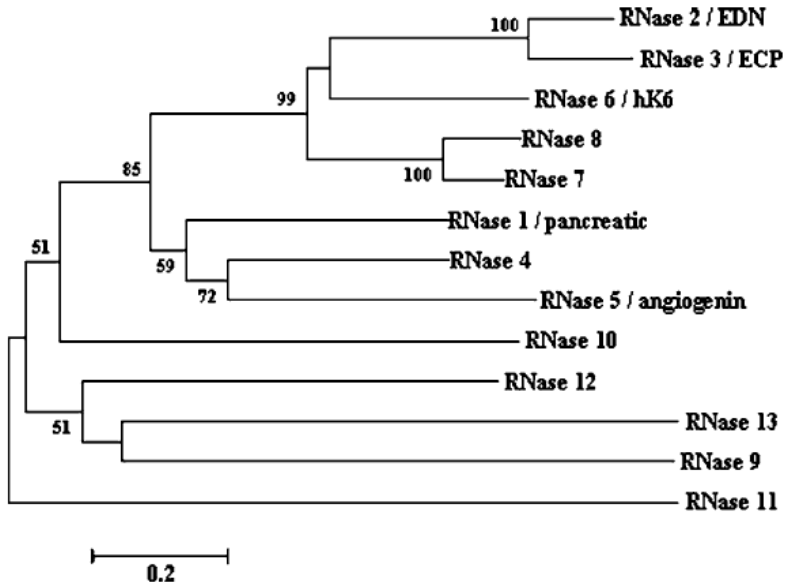


Figure 3. Phylogenetic tree of human RNase A ribonucleases, reproduced from Dyer, Rosenberg 2006

B. Obesity

"Corpulence is not only a disease itself, but the harbinger of others"
Hippocrates (c. 460 BC – c. 370 BC)

Introduction

Obesity is a medical condition where fat accumulates in an abnormal and excessive way that may affect human health.
To classify overweight and obesity several methods exist. The most common and simple one is the Body Mass Index (BMI). BMI is defined as weight in kilograms divided by the square of height in meters (kg/m2).
WHO definition [WHO 2006]:

BMI	**Classification**
< 18.5	underweight
18.5-24.9	normal weight
25.0-29.9	overweight
30.0-34.9	class 1 obesity
35.0-39.9	class 2 obesity
≥ 40.0	class 3 obesity

Although BMI is an accepted method to monitor the weight status of a population, it has often limits for assessing an individual's status. The reason is that BMI does not distinguish between fat and muscle mass.

Another method to measure especially the abdominal obesity is the waist circumference. Women should have a waist circumference < 88cm and men < 102cm, the waist-hip-ratio is considered as healthy under 0.85 for women and 0.9 for men.

Facts according to WHO

- Overweight and obesity are the fifth major risk for death.
- Approximately over 2.8 million persons die each year as a consequence of being overweight or obese.
- 44% of the diabetes burden, 23% of the ischaemic heart disease burden and between 7% and 41% of certain cancer burdens are assignable to overweight and adiposity.
- WHO estimates from 2008
    - More than 1.4 billion adults are overweight.
    - Of these over 200 million men and nearly 300 million women are obese.
    - All together, more than 10% of the world's adult population is obese.
- Obesity is linked to more deaths worldwide than underweight, e.g. 65% of the world's population live in countries where obesity kill more people than underweight [WHO 2013]

Causes and mechanisms

Obesity has not a single cause but is a combination of different conditions with multiple causes. Most important is the interaction between genetics, phsychosocial and environmental factors [Kopelman 2000].

Lifestyle and environment criteria are important but sometimes insufficient factors in obesity.

Different mutated genes can cause obesity. These genes are often part of the energy balance regulation system. For example, the ob gene encodes leptin. Leptin is a hormone which is secreted by adipose tissue with increased body weight. It binds to leptin receptors in the brain, e.g. in the hypothalamus region. As result a sensation of satiety occurs. Body weight reduction decreases leptin production and increases food intake (Figure 4).

Mice with mutations in ob do not produce any leptin and show a significant weight gain compared to normal mice. The situation is similar for humans. There are also neurons in the hypothalamus that are activated by leptin and express the neuropeptide precursor

POMC. Mutations in POMC or in its receptor MC4 can be also responsible for obesity [Friedmann 2009].

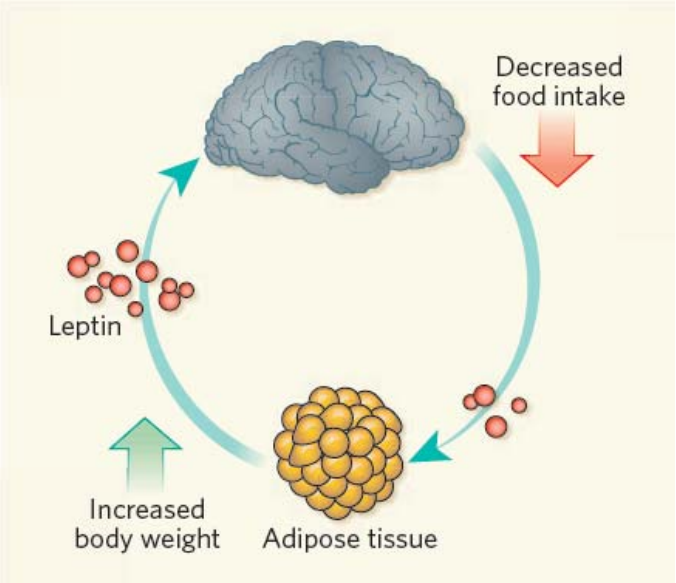


Figure 4. Leptin influence on food intake, reproduced from Friedmann 2009

Obese individuals have an excess of fat cells that are increased in size compared to lean individuals. In addition, almost all obese patients have a higher number of adipocytes as well (Figure 5). However, common therapy methods are capable to decrease adipocyte size but not the adipocyte number [Arner, Spalding 2010]. This can become a new therapeutical target for obesity treatment; see also the section “Ribonuclease A as a new target for treating obesity” below.

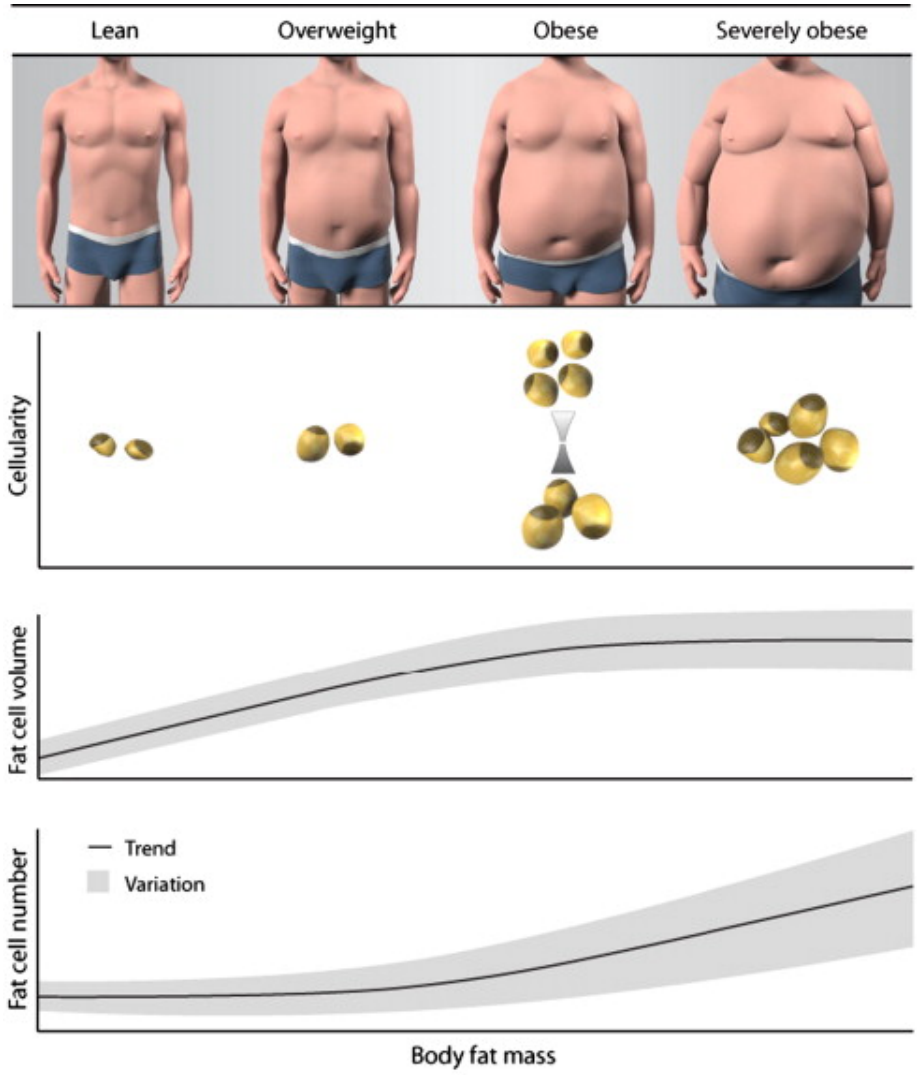


Figure 5. Obesity is characterized by larger adipocyte size as well as by larger adipocyte number, reproduced from Arner, Spalding 2010

Effects on health

Obesity is associated with several health problems like cardiovascular diseases, diabetes mellitus type II, asthma, osteoarthritis, some types of cancer and an increased risk of premature death [Haslam, James 2005].]

Metabolic syndrome

The metabolic syndrome consists of multiple risk factors for cardiovascular disease. Components related to it are:

Abdominal obesity

Atherogenic dyslipidemia

High blood pressure

Insulin resistance, glucose intolerance

Proinflammatory state

Prothrombotic state

[Grundy et al. 2004]

Diagnosis can be made if at least 3 of the 5 compounds from Figure 6 are present.

Risk Factor	Defining Level
Abdominal obesity, given as waist circumference*†	
Men	>102 cm (>40 in)
Women	>88 cm (>35 in)
Triglycerides	≥150 mg/dL
HDL cholesterol	
Men	<40 mg/dL
Women	<50 mg/dL
Blood pressure	≥130/≥85 mm Hg
Fasting glucose	≥110 mg/dL‡

Figure 6. Metabolic syndrome criteria, reproduced from Grundy et al. 2004

Especially abdominal fat is considered to elevate cardiovascular morbidity, including stroke, myocardial infarction, congestive heart failure and cardiovascular death.

In addition to abdominal fat insulin resistance is also a risk factor for cardiovascular diseases in obesity (see Figure 7). Fat tissue secrets adipokines, especially adiponectin and TNF- α, after macrophage recruitment. The adipokines have a direct effect on endothelial dysfunction. TNF-α and IL-6 might as well indirectly influence inflammation and endothelial dysfunction. Furthermore, insulin resistance that is induced by cytokines, NEFA and retinolbinding protein 4 may produce oxidative stress and following endothelial dysfunction. Adipose tissue, insulin resistance, liver-induced inflammation and dyslipidaemia can cause and accelerate the atherosclerotic process [Van Gaal et al. 2006].

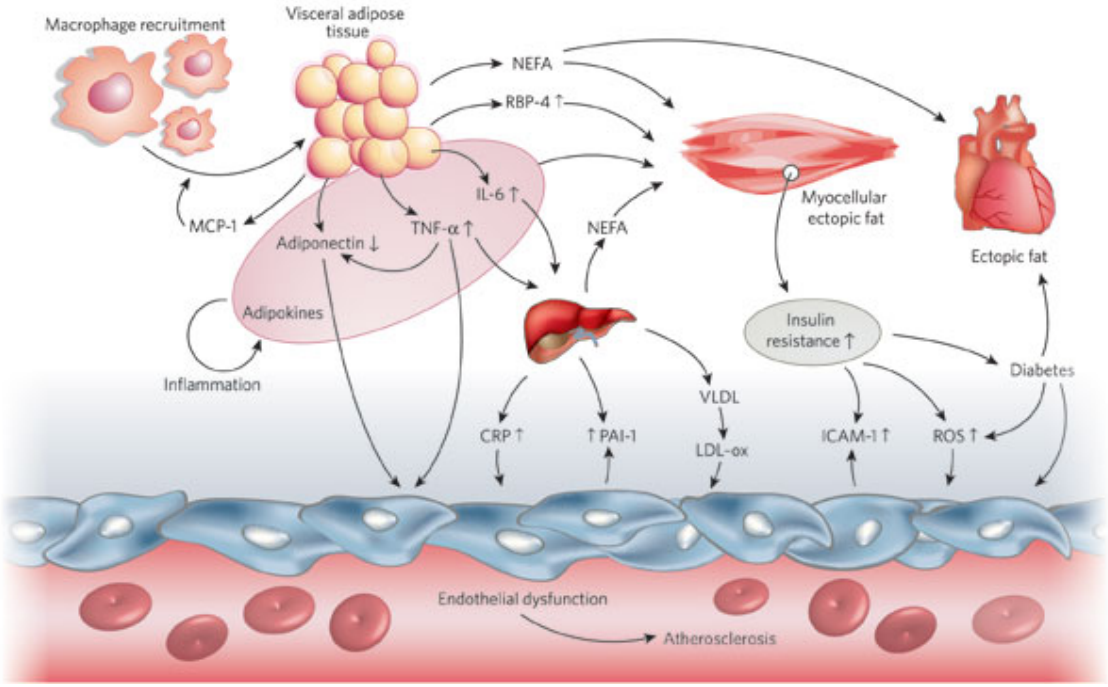


Figure 7. Mechanisms leading to cardiovascular diseases, reproduced from Van Gaal et al. 2006

Excess weight leads to bad health. Figure 8 shows detailed estimates of the years of illness and lost lives due to obesity for several regions of the word. Cardiovascular disease is on top, followed by diabetes and cancer. Surprisingly, Asian countries have a higher disease burden comparing to their lower rate of adiposity. The highest number of disability-adjusted life-years lost can be found in central and eastern Europe [Haslan, James 2005].

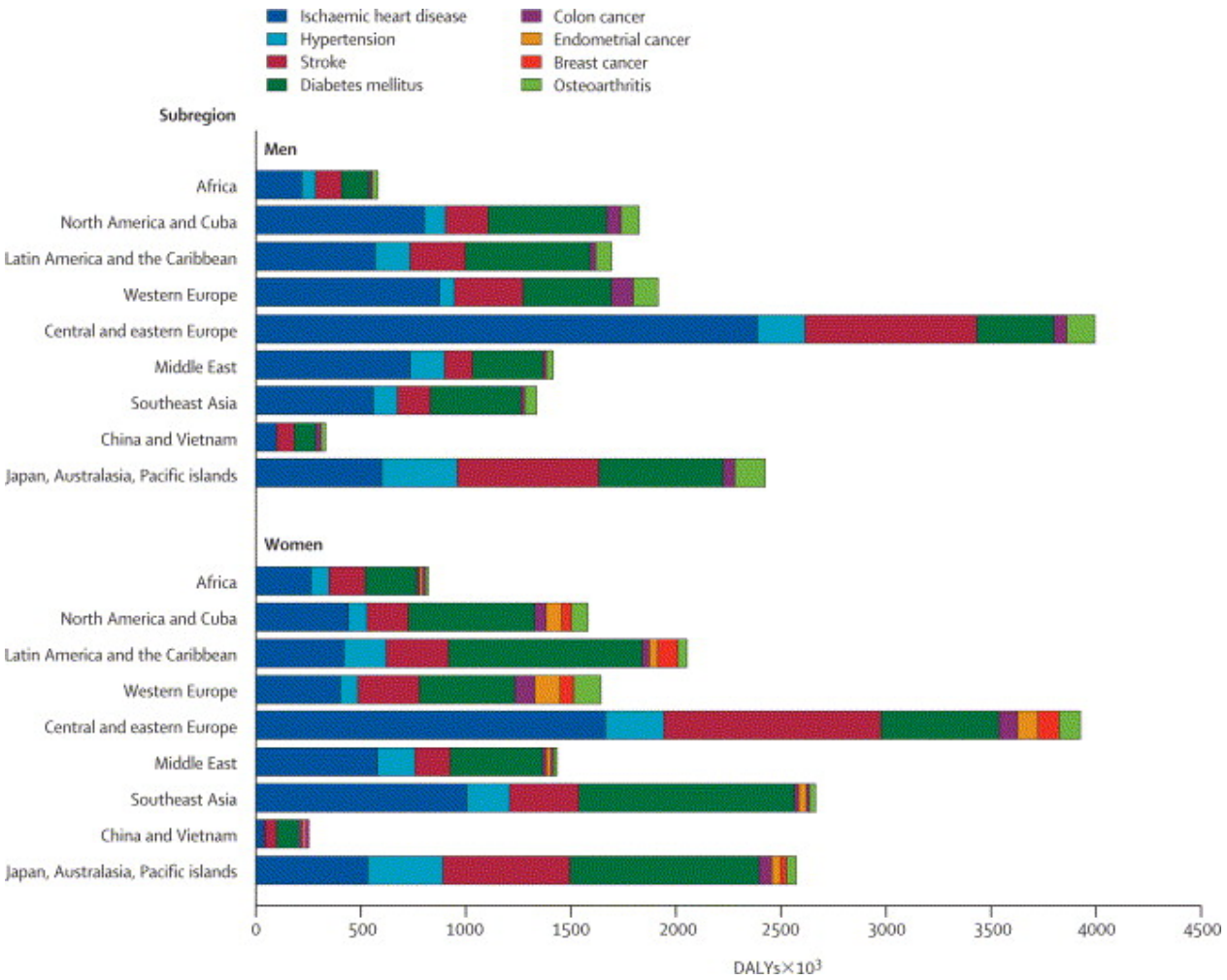


Figure 8. Disability-adjusted life-years (DALYs) lost as a result of obesity, reproduced from Haslan, James 2005

Treatment

Lifestyle change

The main treatment of obesity today is a lifestyle change including diet and physical exercises. The approach is to reduce the calorie intake by 500-1000 kcal per day together with portion sizes and fatty and high-sugar foods. In addition, exercise is recommended, for approximately 30 min per day on a regular basis [Carvajal et al. 2013].

However, the weight loss and weight maintenance success is not really big because of several reasons, for example because of the counseling difficulty and patient compliance [Yanovski et al. 2013]

Bariatric surgery

Bariatric surgery is another treatment option. It has shown meaningful success in treating obesity. However, this is an invasive surgery and may lead to dangerous complications.

Pharmacotherapy

In some cases lifestyle change is combined with pharmacotherapy. For some drugs allowed for obesity treatment see Figure 9. Some of the commonly used drugs are for example orlistat, phentermine, and lorcaserin [Yanovski et al. 2013]. Pharmacotherapy is associated with modest weight loss. However, the weight loss is not enormous and many drugs have side effects that decrease the compliance (e.g. like steatorrhea). Others turn out to have a serious influence on health, like sibutramine, a serotonin–norepinephrine reuptake inhibitor, which was removed from the market in 2010 because of the risk of cardiovascular events [Carvajal et al 2013]

Generic Drug (Proprietary Name[s] Dose Frequency/d)	Mechanism of Action	Wholesale Price/mo, $[a]	1-y Weight Change Relative to Placebo, Mean (95% CI), kg[b]	Common Adverse Effects
Short-term approval[c]				
Phentermine 15-37.5 mg (Adipex-P, Fastin, Oby-Cap, Ionamin, Others; 1×)[d]	Noradrenergic causing appetite suppression	6-45	Not included	Insomnia, elevation in heart rate, dry mouth, taste alterations, dizziness, tremors, headache, diarrhea, constipation, vomiting, gastrointestinal distress, anxiety, and restlessness[e]
Diethylpropion 25 mg or 75 mg, SR (Tenuate, Tenuate Dospan, Tepanil; low dose, 3×; SR dose, 1×)[d]	Noradrenergic causing appetite suppression	47-120	Not included	Same as phentermine[e]
Phendimetrazine 17.5-70 mg or 105 mg, SR (Bontril; lower doses, 2-3×; SR dose, 1×)[f]	Noradrenergic causing appetite suppression	6-20	Not included	Same as phentermine[e]
Benzphetamine 25-50 mg (Didrex; 1-3×)[f]	Noradrenergic causing appetite suppression	20-50	Not included	Same as phentermine[e]
Long-term approval[c]				
Orlistat 60 mg (Alli) or 120 mg (Xenical; 3× within 1 h of a fat-containing meal)[g]	Lipase inhibitor causing excretion of approximately 30% of ingested triglycerides in stool	60 mg, 45 120 mg, 207	60 mg, −2.5 kg (−1.5 to −3.5) 120 mg, −3.4 kg (−3.2 to −3.6)	Oily spotting, flatus with discharge, fecal urgency, fatty oily stool, increased defecation, fecal incontinence[h]
Lorcaserin 10 mg (Belviq; 2×)[d]	Highly selective serotonergic 5-HT2C receptor agonist causing appetite suppression	240	−3.2 kg (−2.7 to −3.8)	Headache, dizziness, fatigue, nausea, dry mouth, cough, and constipation; and in patients with type 2 diabetes, back pain, cough, and hypoglycemia[h]
Phentermine plus topiramate-ER (Qsymia; 3.75 mg/23 mg for 2 weeks, increased to 7.5 mg/46 mg, escalating to a max of 15 mg/92 mg; 1×)[d]	Noradrenergic + GABA-receptor activator, kainite /AMPA glutamate receptor inhibitor causing appetite suppression	140-195	7.5 mg/46 mg, −6.7 kg (−5.9 to −7.5) 15 mg/92 mg, −8.9 kg (−8.3 to −9.4)	Paresthesias dizziness, taste alterations, insomnia, constipation, dry mouth, elevation in heart rate, memory or cognitive changes[h]

Abbreviations: AMPA, α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid; ER, extended release; GABA, γ-aminobutyric acid.

[a] Reference prices[9] as of March 8, 2013.

[b] Weight change data are relative to placebo using intent-to-treat analyses for each medication at 1 year. No studies for older noradrenergic agents met inclusion criteria for length of treatment, sample size, and attrition.

[c] Food and Drug Administration–approved for short-term (ie, a few weeks) or long-term use.

[d] Medications listed on Drug Enforcement Administration Schedule IV are associated with a lower risk of abuse than medications on Schedule III.

[e] Common adverse events for noradrenergic agents include those listed as common in Prescription Medications for the Treatment of Obesity[10] because adverse event frequency is not available in drug package inserts for these agents.

[f] Drug Enforcement Administration Schedule III medication.

[g] Orlistat is a non-Drug Enforcement Administration–scheduled drug.

[h] For orlistat, lorcaserin, and phentermine plus topiramate-ER, common adverse events are those listed in the drug package inserts[11-13] that are reported to occur more frequently than placebo and with more than 5% prevalence. See full prescribing information for all adverse effects, cautions, and contraindications.

Figure 9. Drugs with indication for obesity, approved in the US, reproduced from Yanovski et al. 2013.

Ribonuclease A as a new target for treating obesity

According to new findings Ribonuclease A plays a role in formation of adipose tissue. Members of the RNAse A family have been identified as key regulators of preadipocyte differentiation. It was shown that inhibitors of Ribonucleases prevent differentiation of preadipocytes and therefore interfere with the formation of mature adipocytes. Thus, the number of adipocytes has dramatically decreased in the experiment. It was also demonstrated that murine EAR-1, -2 and -10 and human RNAse 1 are strongly expressed in adipose tissue. Finally, it was found that RNases promote adipogeneses as evidenced by the induction of the expression of adipogenes genes and the increased number of adipocytes

The inhibition of Ribonuclease prevents differentiation of preadipocytes, and decreases the number of adipocytes. Proliferation of preadipocytes and increase in the number of adipocytes are important determinants associated with obesity. Therefore, the inhibition of Ribonuclease A is a new strategy for for developing drugs for treating obesity [Bilban et al. 2013].

II. Aim of Study

The treatment of obesity represents a major task for modern medicine, not only in the developed countries but also in the developing world [Friedmann 2009]. As obesity is increasing rapidly and becoming a major risk for premature death, it is of high importance to conduct research in this field.

The commonly used treatment options, such as lifestyle change, including diet with calorie restriction and physical exercise, bariatric surgery and known pharmacotherapy influence adipocyte size only. A therapy method that also decreases adipocyte number is missing up to date.

Ribonuclease A, an enzyme of the ribonuclease A superfamily, is associated with obesity and has an impact on adipocyte size and number. For this reason RNase A is an interesting novel target for drug development for treatment of obesity.

The aim of this study was to identify new ligands that can act as antagonists for Ribonuclease A by *in-silico* methods. The purpose of the work was to enable the beginning of drug development for obesity, as it is a serious health affecting disease with need for new medication options.

III. Computational Background

A. Computer-Aided Drug Design

Computer-Aided Drug Design (CADD) has become an important part of pharmaceutical research and drug development. CADD expedites drug development by using information of existing drugs and biological targets, combined with interdisciplinary information from different fields.

The process of pharmaceutical R&D is expensive and time-consuming. The drug discovery process can be split in different stages (Figure 10). Target identification and validation, hits and lead compounds identification followed by lead optimization. These make up the research phase of the drug development process. The research phase is followed by preclinical development and clinical trial phases I-III. In total it takes over ten years and around $500 million to more than $2,000 million until a new drug gets developed and approved [Gao et al, 2010]

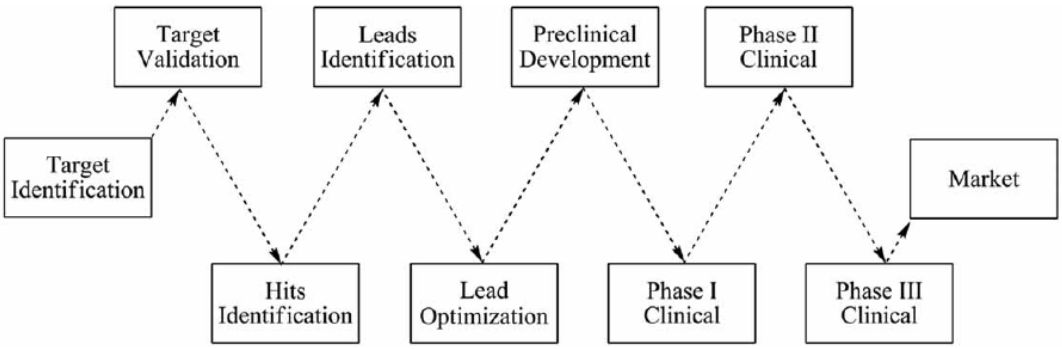


Figure 10. Drug development process: from drug discovery to the market, reproduced from Gao et al. 2010.

Structure-based design

Nowadays a large number of 3D structures of molecular targets are available because of the big medicinal chemistry success in the last decade. Structure based design is based on the search of compounds that fit into the binding site of a protein [Gubernator and Boehm 1998]

B. Sequence Alignment

One of the widely used bioinformatics analyses are nowadays multiple sequence alignments. They are used routinely as parts of more complicated analyses. Multiple sequence alignment is helpful for highlighting areas of similarity that may be associated with specific features that have been more conserved than other regions. Sequence alignment plays also an important role for phylogenetic analysis which models the substitutions that have occurred over evolution. Several software packages are available, e.g. ClustalW, ClustalX, T-Coffee, MUSCLE and MAFT [Larkin et al. 2007].

Clustal W

ClustalW [Larkin et al. 2007] is a web based tool to align multiple sequences together in a computationally efficient manner. The multiple sequence alignment web form is provided by EMBL-EBI at http://www.ebi.ac.uk/Tools/msa/clustalw2/.
There are two different options to utilize this service: either interactively or by email.
In the first case the results are displayed in the browser window. By choosing the email setting instead, a link to the results is sent by email. This is profitable when the submitted amount of data is large and therefore the job would take a long time to run.

The program allows the input of nucleic acid or protein sequences in the following multiple sequence formats:

- NBRF/PIR
- EMBL/UniProt
- Pearson (FASTA)
- GDE
- ALN/ClustalW
- GCG/MSF
- RSF

The sequences can be uploaded to the web form in a file or pasted into the web form. It is necessary that each of the sequences has a unique name. If they do not, the job cannot be done. Other causes for failure are empty lines, white spaces or control characters between sequences at the top of the file.

The input for ClustalW is restricted to a maximum of 500 sequences or to a 1MB file. When the number of sequences is too big, ClustalW can run for a long time and may not finish at all in some cases.

There are several output options for the alignment file, which can be selected during job submission:

- Clustal w/ numbers
- Clustal w/o numbers
- GCG MSF
- PHYLIP
- NEXUS
- NBRF/PIR
- GDE
- Pearson/FASTA

By default the main output is the alignment file. Other outputs can be seen in the results summary tab. The output also contains a Score Table that shows the pair wise scores calculated for every pair of sequences that is to be aligned. Pair wise scores are the number of identities between the two sequences, divided by the length of the alignment, and represented as a percentage. This alignment is only a pioneer to the full multiple alignment [ClustalW2 2007]

C. Pharmacophore Modeling

"A model must be wrong, in some respects, else it would be the thing itself. The trick is to see where it is right."
Henry A. Bent.

Introduction

The concept of pharmacophore was first defined by Paul Ehrlich in 1909 as "a molecular framework that carries (*phoros*) the essential features responsible for a drug's (*pharmacon*) biological activity" [Ehrlich, 1909]. IUPAC defines a pharmacophore as "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response" [Wermuth et al. 1998]. The pharmacophore concept is basically based on the same interactions as in drug-receptor complexes: hydrogen bonding, charge transfer, electrostatic and hydrophobis interactions [Wolber, Langer 2005].

A pharmacophore model can be generated either ligand-based by extracting common chemical features from a set of active molecules, or structure-based by examining possible interaction points between the protein binding site and selected ligands. Pharmacophore models are often used in virtual screening, de novo design and other applications (see Figure 11). [Yang 2010]

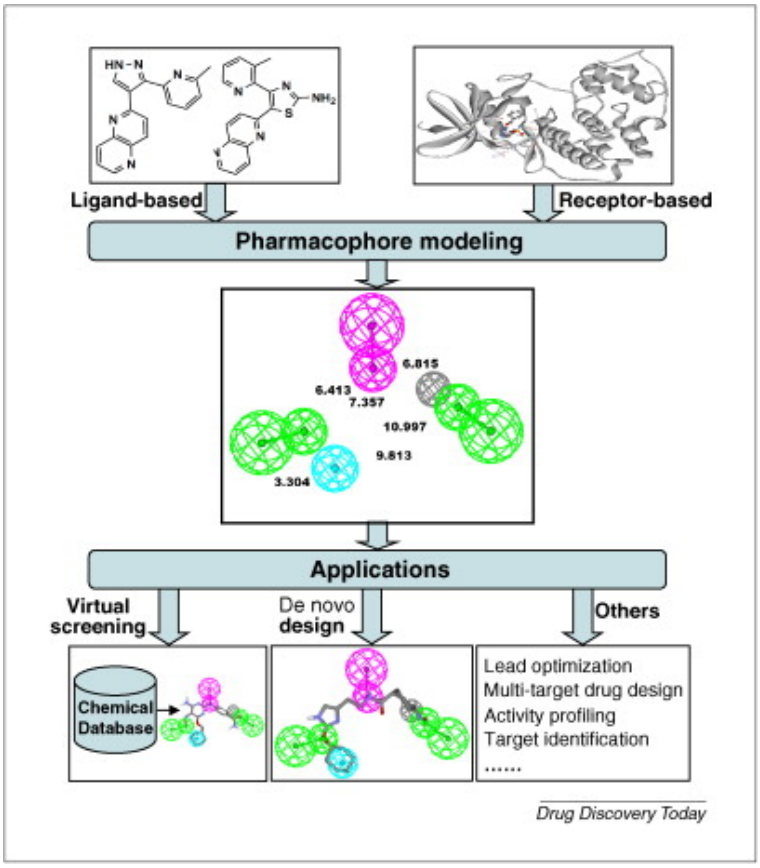


Figure 11. Framework of pharmacophore modeling and applications, reproduced from Yang, 2010.

Structure-based Pharmacophore Generation

The structure-based pharmacophore modeling concept is based on the examination of the 3D structure of a macromolecule-ligand complex or a macromolecule without any ligand. The protocol analyzes the complementary chemical features of the binding site and their spatial relationships to receive a pharmacophore model assembly with selected features. The models are generated based on the set of pharmacophore features mentioned before in addition to exclusion spheres [Yang 2010]. There are several commercial tools provided for structure-based pharmacophore modeling, e.g. LigandScout, DS Structure Based Pharmacophore, MOE, and Phase [Gao et al.2010].

Virtual Screening

Virtual screening is an important part of CADD because it can minimize the costs and time of drug development for pharmaceutical companies. The screening eliminates undesired molecules from compound libraries. There are different stages the virtual screening has to proceed to deliver the desired lead molecules (Figure 12):

Stage 1: Generation of Pharmacophore

For the structure-based design a 3D Structure of the target or a target-ligand-complex is needed. If there is no 3D structure avaible, a homology model can be used as well.

Stage 2: Virtual Screening

Commercial databases are screened with 3D pharmacophore models. Key points are: database selection, involvement and tolerance radius of excluded volumes, method for conformational search, energy windows, and maximum number of conformers in the conformation generation.

Stage 3: Generation of Primary Hits

The aim in this stage is to limit the number of compounds that will proceed the more expensive phase of biological testing. To provide information for false positive hits bioassays are used.

Stage 4: Generation of Lead Compounds

In this stage the identification and optimization process takes place. It is necessary to optimize the compounds to obtain lead-like and drug-like properties.

[Gao et al. 2010]

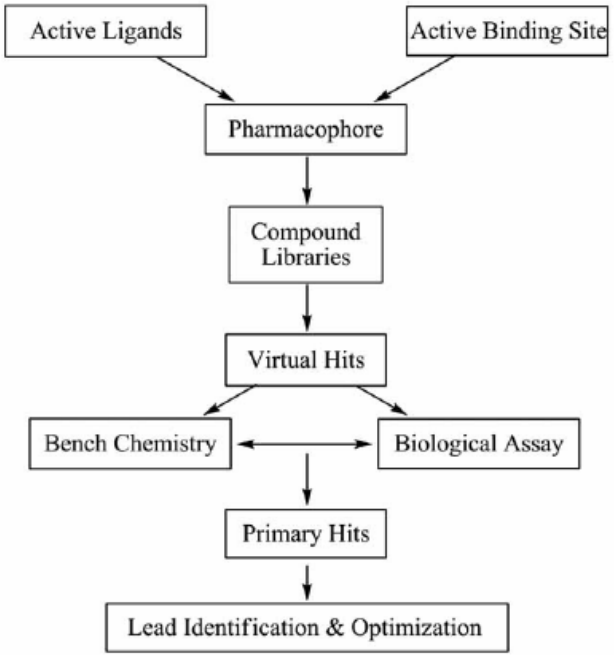


Figure 12. Workflow for pharmacophore based virtual screening, reproduced from Gao et al 2010

D. ChemGPS-NP

ChemGPS-NP_{Web} is a web-based public tool for navigation in biologically relevant chemical space. A chemical space map can be created applying similar principles as the Mercator convention in geography: Structures correspond to objects (e.g. cities) and rules correspond to dimensions (e.g. longitude and latitude).

ChemGPS-NP_{Web} uses a total set of 35 descriptors (Figure 13). The first four PCAs are the most important ones, they explain 77% of the variance [Larsson et al. 2007].

Work-flow

ChemGPS-NP_{Web} consists of different programs and libraries. The main three elements are DragonX for calculation of molecular descriptors, Simca-QP for multivariate data analysis, and the web interface Batchelor (Figure 14).

Structures are uploaded as SMILES-strings by the queue handler. First, an initial Perl script removes erroneous SMILES and the information about stereochemistry and isotopes. Then the SMILES are submitted to DragonX. 40 molecular descriptors from DragonX are used, out of which six are summarized into one descriptor (n_amid). Altogether there is a final set of 35 descriptors. The resulting matrix is used as input data to the client cgpsclt that connects to the server cgpsd and starts Simca-QP. Simca-QP performs the PCA score prediction via the library libchemgps. The result (eight coordinates for each compound) is returned to the user. The coordinates can then be plotted by using any avaible software [Rosen et al. 2009].

number	abbreviation	description
1	MW	molecular weight
2	Sv	sum of atomic van der Waals volumes (scaled on C atom)
3	Se	sum of atomic Sanderson electro-negativities (scaled on C atom)
4	Sp	sum of atomic polarizabilites (scaled on C atom)
5	Mv	mean atomic van der Waals volume (scaled on C atom)
6	Me	mean atomic Sanderson electro-negativity (scaled on C atom)
7	nAT	number of atoms
8	nSK	number of non-hydrogen atoms
9	nBT	number of bonds
10	nBO	number of non-hydrogen bonds
11	nBM	number of multiple bonds
12	ARR	aromatic ratio
13	nCIC	number of rings
14	RBN	number of rotatable bonds
15	RBF	rotatable bond fraction
16	nDB	number of double bonds
17	nAB	number of aromatic bonds
18	nC	number of carbon atoms
19	nN	number of nitrogen atoms
20	nO	number of oxygen atoms
21	nX	number of halogens
22	nBnz	number of benzene-like rings
23	nCar	number of aromatic carbon atoms (sp^2)
24	n_amid	number of amides
25	nROH	number of aliphatic hydroxy groups
26	nArOH	number of aromatic hydroxy groups
27	nHDon	number of donor atoms for hydrogen bonds (N and O)
28	nHAcc	number of acceptor atoms for hydrogen bonds (N, O, and F)
29	Ui	unsaturation index
30	Hy	hydrophilic factor
31	AMR	Ghose−Crippen molar refractivity
32	TPSA(NO)	topological polar surface area using N and O
33	TPSA(Tot)	topological polar surface area using N, O , S, and P
34	ALOGP	Ghose−Crippen octanol−water partition coefficient
35	LAI	Lipinski alert index (drug-like index)

Figure 13. ChemGPS-NP descriptors, reproduced from Larsson et al. 2007

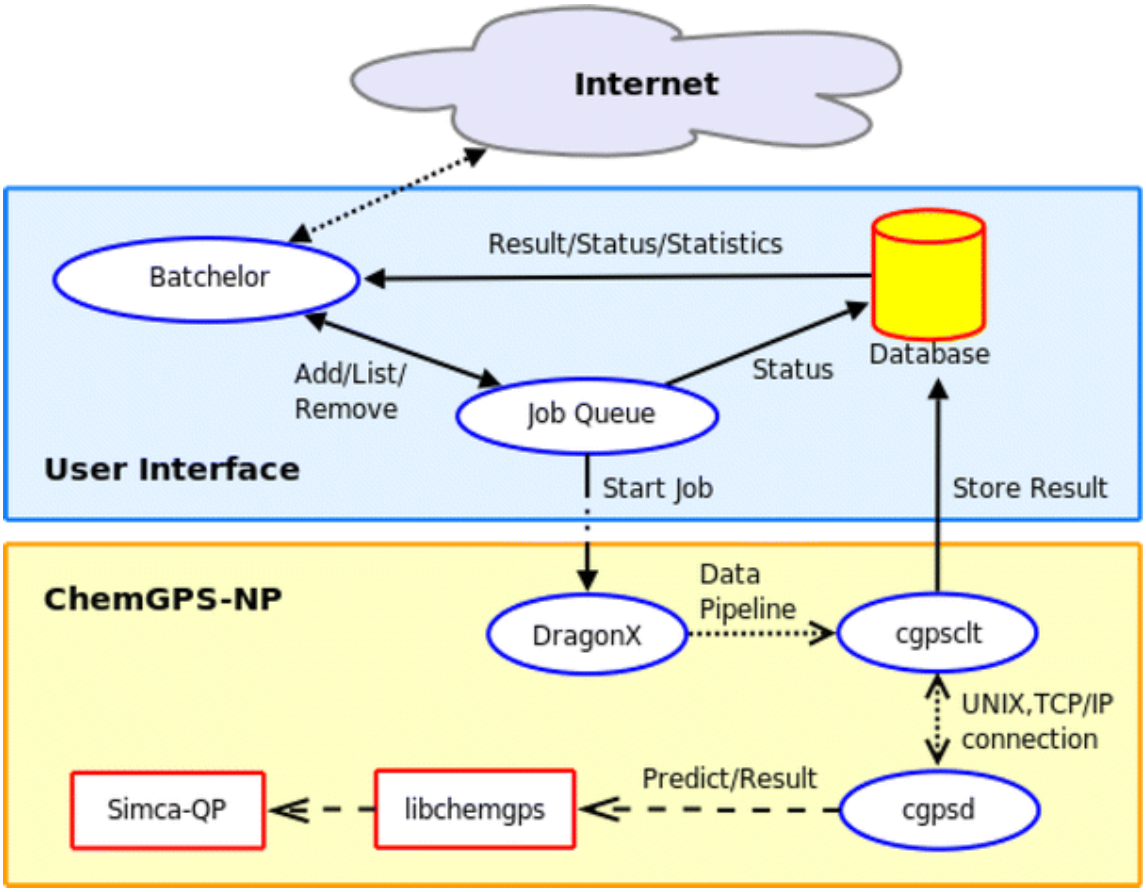


Figure 14. Interactions between the different elements of ChemGPS-NP$_{Web}$, reproduced from Rosen et al. 2009

E. Docking

Docking is a computational method to predict the docking pose for a compound in a binding site of another molecule. The ligand and the protein form a complex (Figure 15). It is also possible to dock big molecules like proteins into proteins.

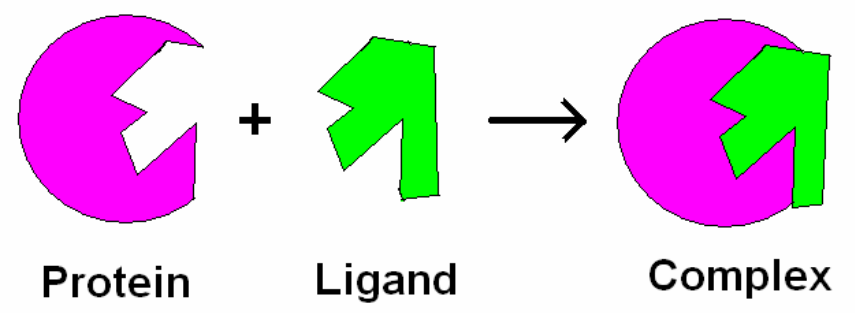


Figure 15. Schema of complex formation

According to Leach et al. 2006, there are three main aims for docking studies:

1. To predict the binding mode of known active ligands.
2. Discovery of new ligands using virtual screening techniques.
3. To predict binding affinities of active compound in the binding site.

Today several algorithms are avaible for protein-ligand docking. Some of the programs are DOCK, FlexX, PRO_LEADS and GOLD.

GOLD (Generic Optimization for Ligand Docking) is based on a genetic logarithm and is provided by GOLD Suite 5.1 (Verdonk et al. 2003). Similar to other docking programs, GOLD consists of three main parts:

1. The scoring function has the purpose to rank different binding modes.
    - GOLDScore is the original scoring function and includes aspects like hydrogen bonding energy, van der Waals energy, metal interaction and ligand torsion strain.
    - ChemScore includes the term dG, which represents the total free energy change that occurs on ligand binding.
    - ChemPLP uses Chemscore terms for hydrogen bonding and internal energy, and a piece-wise linear potential for hydrophobic and non-complementary interactions.

- ASP (Astex Statistical Potential) is based on atom-atom distance potential considering protein-ligand complexes from a database.

2. Mechanism for placing the ligand in the binding site: GOLD uses a method that is based on fitting points.

   In general it is important that the algorithm considers flexibility of protein and ligand. However, because of speed issues, one of the bodies, e.g. protein or ligand, is often considered rigid.

   3. Search algorithm to explore possible binding modes, a genetic algorithm in GOLD [Verdonk et al. 2003]

IV. Material and Methods

A. Sequence Alignment

First of all a crystal structure of Ribonuclease A was required because of the aim of this study to discover new possible inhibitors of RNase A via structure-based design. However, at the moment a crystal structure of human RNase A has not been published yet. Furthermore the experimental testing on inhibition activity will be first carried out in mice before going on testing it on humans. For these reasons sequence alignment has to be done to find out if the human, murine and bovine Ribonuclease A proteins are comparable with each other.

The protein sequences of the three different species (human, bovine and murine) RNase A were aligned using ClustalW2 provided by EBI [Larkin et al. 2007] with default settings.
The human protein sequence of RNase A was also blasted against the Protein Data Base [Bernstein et al. 1977] in UniProt [UniProt ©2002-2013] to obtain the protein sequence identity.

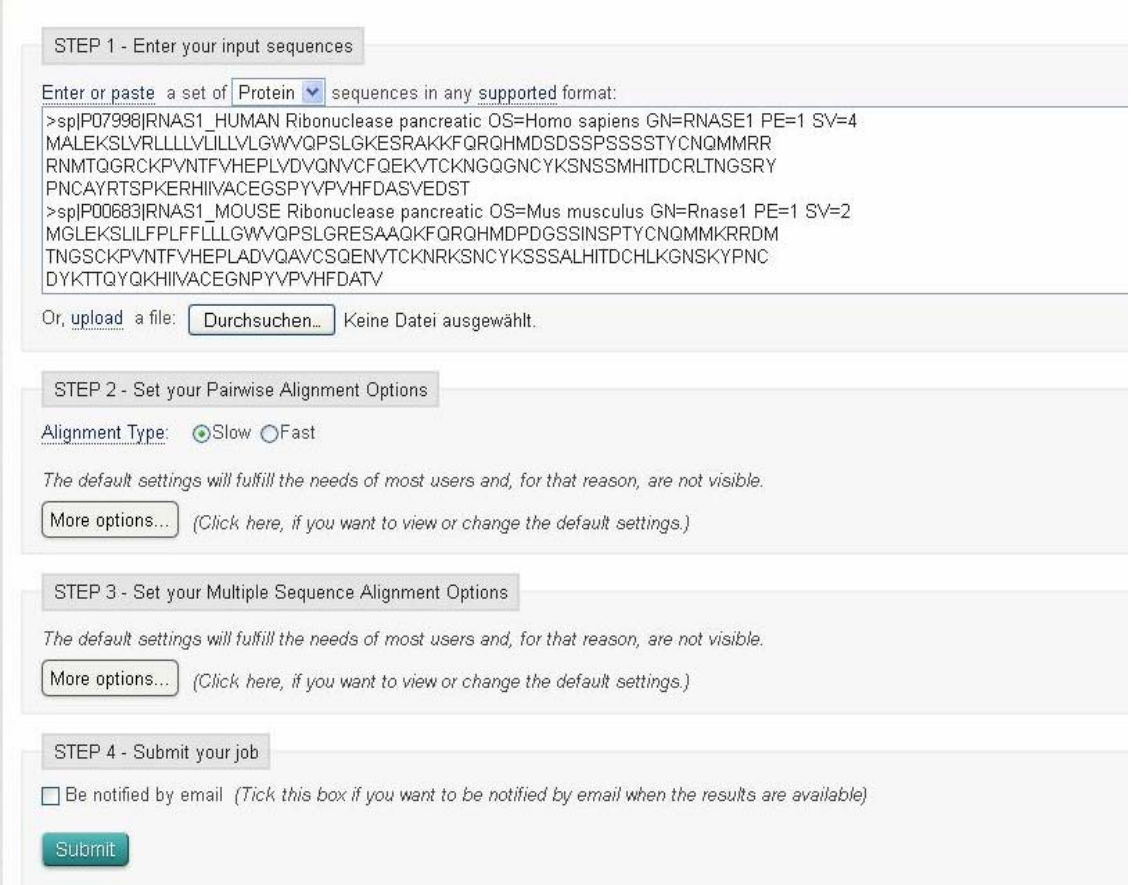


Figure 16. ClustalW2 web input form (Picture taken from www.ebi.ac.uk/Tools/msa/clustalw2/)

B. Database search

To generate the pharmacophore model a crystal structure of a complex of Ribonuclease A with an inhibitor was needed. In addition, the ChemGPS-NP method [Larsson et al. 2007] also required a database search for known antagonists. This task for was carried out by using the ChEMBL database [Gaulton et al. 2011], Protein Data Base (PDB) and PubMed [http://www.ncbi.nlm.nih.gov/pubmed].

C. Pharmacophore Model and Screening

Generation of Pharmacophore Model

To create a structure based pharmacophore model the crystal structure of bovine RNase A in complex with Thymidine-3-phosphate from PDB entry 3LXO was used. This complex made a good staring point for the pharmacophore creation as Thymidine-3-phosphate is described as a competitive inhibitor of Ribonuclease A, [Doucet et al. 2010] and the crystal structure possesses a good resolution of 1,7A.

First the ligand interactions were analyzed in MOE [Molecular Operating Environment Software 2012.10]. Then a three dimensional pharmacophore model was created by using the LigandScout program package of Inte:Ligand [Wolber, Langer 2005]. Five Hydrogen Bond Acceptor features, one Hydrogen Bond Donor feature and one ionizable area were included in the model. In addition to the chemical features exclusion volumes were added to achieve a more selective model.

Screening a Commercial Database

To receive compounds for the future experimental testing a commercial database had to be screened. For this, the LifeChemicals database [www.lifechemicals.com] containing 366000 compounds was screened with the generated Pharmacophore model using LigandScout.

D. ChemGPS-NP

To analyze the chemical properties of the received hits from the pharmacophore screening and to compare them with each other and with the known inhibitors from literature, ChemGPS-NP [Larsson et al. 2007] was used.

The compounds found through the pharmacophore screening were converted into a list of SMILES (Simplified Molecular Input Line Entry Specification) using MOE. The inhibitors found in the literature were also converted into SMILES using Chemical Identifier Resolver [http://cactus.nci.nih.gov/chemical/structure]. Both SMILES lists were processed with ChemGPS-NP Web. The output consisted of eight dimensions of chemical properties for each compound. The first three PCA were plotted in a 3D- diagram using MOE. The main influence in PCA1 was size, in PCA2 aromaticity, and in PCA3 lipophilicity.

E. Docking

The docking studies were performed with the genetic-algorithm based program GOLD (Genetic Optimization for Ligand Docking) provided by GOLD Suite 5.1 [Verdonk et al. 2003].

Ligand and Protein Preparation

The compounds received from pharmacophore screening were analyzed concerning the Lipinski rule of five. The compounds that didn't fulfill the requirements were deleted from the dataset, in particular compounds with molecular weight over 500 and log*P* value over 5 [Lipinski et al. 1997].

The other compounds were further processed in MOE. First the function "Protonate 3D" was used to add hydrogen atoms and calculate the ionization states and partial charges with the default settings. Then energy minimisation was performed. Also the protein from the PDB entry 3LXO was prepared the same way, protonated and energy minimized with MOE.

Docking workflow

In the next step the compounds were docked into the binding site that was defined with the Ligand T3P. Exact settings concerning ligand flexibility are shown in Fig 17.

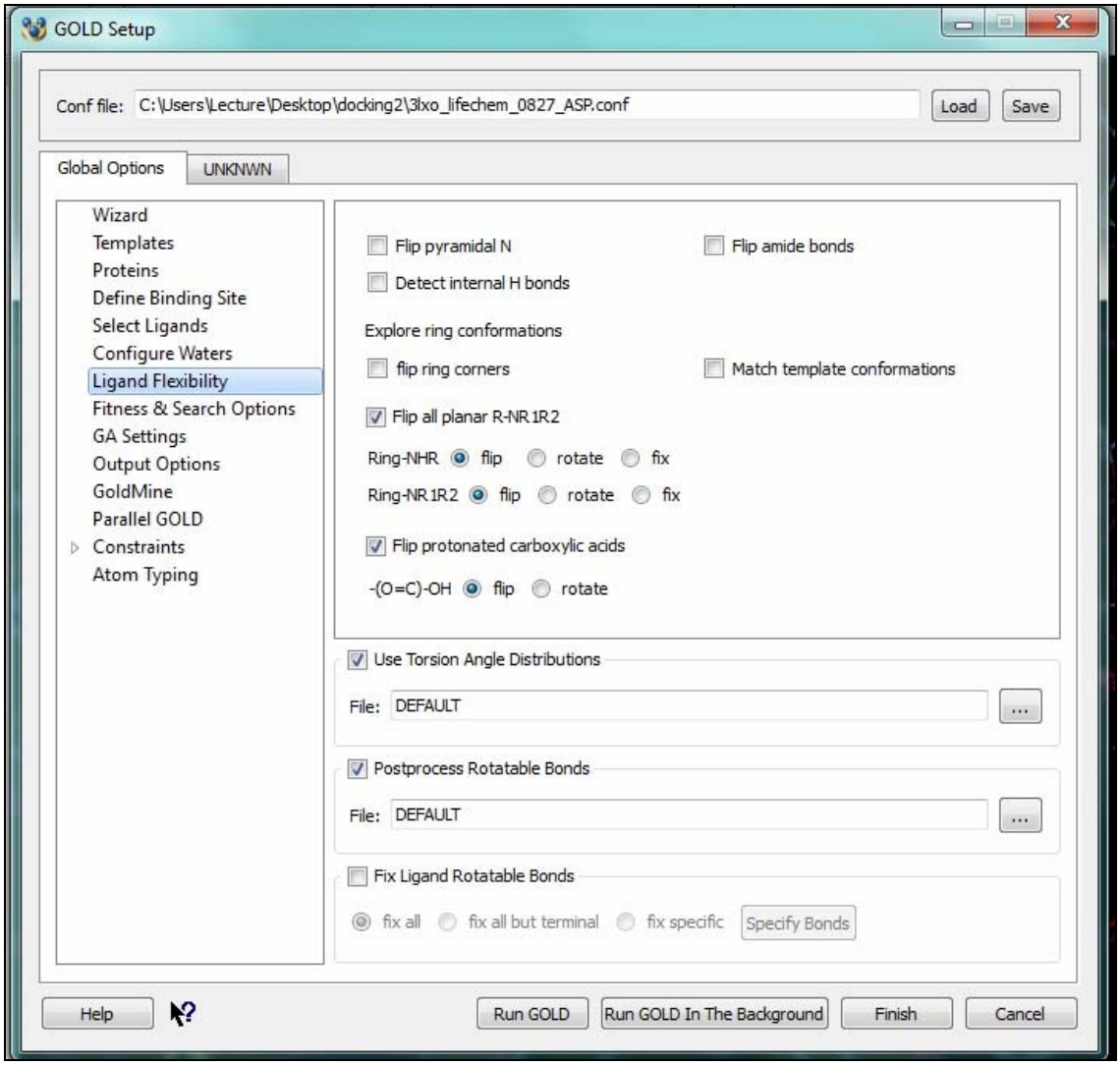


Figure 17. Ligand flexibility settings

Redocking - Validation of the method

To estimate the best suitable scoring function and to validate the method a redocking was performed.

The crystal structure of RNase A in complex with Thymidine-3-phosphate was used (PDB entry 3LXO). The ligand T3P was taken out of the structure of the complex, docked into the binding site, and scored with different functions:

- ChemScore
- GoldScore
- ASP (Astex Statistical Potential)
- ChemPLP (Piecewise Linear Potential)

The binding poses were analyzed and compared to the original binding pose from the PDB entry. The scoring function leading to the most similar binding pose was used for further scoring.

Scoring

Based on the results from redocking the Astex Statistical Potential (ASP) was used as scoring function.

Outcome Processing

The docking outcome complexes were exported from the GOLD output sdf file to an mdb file in MOE. There they were protonated and energy minimized.

V. Results and Discussion

A. Sequence Alignment

The human protein sequence of RNase A was blasted against PDB in Uniprot (Figure 18). The protein sequences of human, mouse and bovine RNase A were also aligned using ClustalW2. Figure 19 shows the results of the alignment.

As you can see the protein sequences of mouse and bovine RNase A show 69% and 70% identity to the human protein. This makes the crystal structure of the murine Ribonuclease A a good starting point for identification of inhibitors that are very likely to inhibit the human Ribonuclease as well.

	Alignments	Entry	Entry name	Status	Protein names	Organism	Length	Identity
☐		P07998	RNAS1_HUMAN	★	**Ribonuclease pancreatic**	Homo sapiens (Human)	156	100.0%
☐		P00683	RNAS1_MOUSE	★	**Ribonuclease pancreatic**	Mus musculus (Mouse)	149	69.0%
☐		P61823	RNAS1_BOVIN	★	**Ribonuclease pancreatic**	Bos taurus (Bovine)	150	70.0%

Figure 18. BLAST in Uniprot results

Results for job clustalw2-I20130902-130745-0267-95136762-pg

Alignments | Result Summary | Guide Tree | Phylogenetic Tree | Submission Details

Download Alignment File | Hide Colors | Send to ClustalW2_Phylogeny

```
CLUSTAL 2.1 multiple sequence alignment


sp|P07998|RNAS1_HUMAN      MALEKSLVRLLLLVLILLVLGWVQPSLGKESRAKKFQRQHMDSDSSPSSS 50
sp|P00683|RNAS1_MOUSE      MGLEKSLI---LFPLFFLLLGWVQPSLGRESAAQKFQRQHMDPDGSSINS 47
sp|P61823|RNAS1_BOVIN      MAL-KSLVLLSLLVLVLLLV-RVQPSLGKETAAAKFERQHMDSSTSAASS 48
                           *.* ***:   *: *.:*::  ******:*: * **:*****.. *. .*

sp|P07998|RNAS1_HUMAN      STYCNQMMRRRNMTQGRCKPVNTFVHEPLVDVQNVCFQEKVTCKNGQGNC 100
sp|P00683|RNAS1_MOUSE      PTYCNQMMKRRDMTNGSCKPVNTFVHEPLADVQAVCSQENVTCKNRKSNC 97
sp|P61823|RNAS1_BOVIN      SNYCNQMMKSRNLTKDRCKPVNTFVHESLADVQAVCSQKNVACKNGQTNC 98
                           ..******: *::*:. **********.*.*** ** *::*:*** : **

sp|P07998|RNAS1_HUMAN      YKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYVPVHFDA 150
sp|P00683|RNAS1_MOUSE      YKSSSALHITDCHLKGNSKYPNCDYKTTQYQKHIIVACEGNPYVPVHFDA 147
sp|P61823|RNAS1_BOVIN      YQSYSTMSITDCRETGSSKYPNCAYKTTQANKHIIVACEGNPYVPVHFDA 148
                           *:* *:: ****: ...*:**** *:*:  ::********.*********

sp|P07998|RNAS1_HUMAN      SVEDST 156
sp|P00683|RNAS1_MOUSE      TV---- 149
sp|P61823|RNAS1_BOVIN      SV---- 150
                           :*
```

Residue	Colour	Property
AVFPMILW	RED	Small (small+ hydrophobic (incl.aromatic -Y))
DE	BLUE	Acidic
RK	MAGENTA	Basic - H
STYHCNGQ	GREEN	Hydroxyl + sulfhydryl + amine + G
Others	Grey	Unusual amino/imino acids etc

An * (asterisk) indicates positions which have a single, fully conserved residue.
A : (colon) indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix.
A . (period) indicates conservation between groups of weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix.

Figure 19. Sequence alignment results with ClustalW2

B. Database Search

The next step was to search for already known and tested inhibitors for Ribonuclease A. This was first carried out in the ChEMBL database and PDB. Because of the few compounds found in these databases the search had to be expanded to PubMed articles.

Finally 25 inhibitors could be identified for further use. Most of the published inhibitors are nucleosides, but there are also a few studies showing polyphenols (for example from green tea) acting as inhibitors. However, there is not much information avaible on this topic yet.

In Table 1 you can see the list of the inhibitors from literature and their SMILES code.

Table 1: Inhibitors and their SMILES

[P](=O)(O)(O)O[C@@H]1[C@H](O[C@H](C1)N2C=C(C(NC2=O)=O)C)CO	Thymidine-3'-monophosphate
[P](=O)(O)(O)O[C@@H]1[C@H](O[C@H]([C@@H]1O)N2C=CC(=NC2=O)N)CO	Cytidine 3'-phosphate
C(C1=CC=CC=C1)(=O)O[C@H]2C[C@@H](O[C@@H]2CO)N3C=C(C(NC3=O)=O)C	3'-O-Benzoylthymidine
C(C)(=O)O[C@@H]1[C@H](O[C@H](C1)N2C=C(C(NC2=O)=O)C)CO	3'-O-Acetylthymidine
[P](=O)(O)(O)OC[C@H]1O[C@H]([C@@H]([C@@H]1O)O)N2C=CC(NC2=O)=O	Uridine 5'-phosphate (U5P)
C1=NC2=C(C(=N1)O)N=CN2[C@H]3[C@@H]([C@@H]([C@H](O3)COP(=O)(O)O)O)O	Inosinic acid
[C@@H]1(OC3=C(C[C@H]1OC(=O)C2=CC(=C(O)C(=C2)O)O)C(=CC(=C3)O)O)C4=CC(=C(O)C=C4)O	(-)-Epicatechin gallate (->Polyphenol)
[C@@H]1(OC3=C(C[C@H]1OC(=O)C2=CC(=C(O)C(=C2)O)O)C(=CC(=C3)O)O)C4=CC(=C(O)C(=C4)O)O	(-)-Epigallocatechin gallate (->Polyphenol)
O=P(O)(O)OP(=O)(O)OC[C@H]3O[C@@H](n2cnc1c(ncnc12)N)[C@H](O)[C@@H]3OP(=O)(O)O	5'-Diphosphoadenosine 3'-phosphate
[P](=O)(O)(O)O[C@@H]1[C@H](O[C@H]([C@@H]1O)N2C=CC(NC2=O)=O)CO	3'-[(Carboxycarbonyl)amino]-2',3'-dideoxy-3,4-dihydrothymidine
c1c(nnn1[C@H]2[C@@H]([C@@H]([C@H](O2)CO)O)O)Cn3cc(c(=O)[nH]c3=O)Br	5-Bromo-2,4-dioxo-1-{[1-(β-D-ribofuranosyl)-1H-1,2,3-triazol-4-yl]methyl}-1,2,3,4-tetrahydropyrimidine
[P](=O)(O)(O)O[C@@H]1[C@H](O[C@H]([C@@H]1O)N2C=CC(NC2=O)=O)CO	Uridine 3'-monophosphate
Nc1ncnc2n(cnc12)[C@@H]1O[C@H](COP(O)(O)=O)[C@@H](OP(O)(O)=O)[C@H]1O	Adenosine-3',5'-bismonophosphate

OC[C@H]1O[C@H]([C@H](OP(O)(O)=O)[C@@H]1O)n1ccc(=O)[nH]c1=O	5'-Diphosphoadenosine 2'-phosphate
OC[C@H]1O[C@H]([C@H](O)[C@@H]1OP(O)(O)=O)n1ccc(=O)[nH]c1=O	Adenosine-5'-diphosphate
Nc1ncnc2n(cnc12)[C@@H]1O[C@H](COP(O)(O)=O)[C@@H](O)[C@H]1OP(O)(O)=O	Adenosine-3-5'-diphosphate
Cc1cn(c(=O)[nH]c1=O)[C@H]2C[C@@H]([C@H](O2)CO)NC(=O)C(=O)O	Uridine 2'-monophosphate
[P](=O)(O)(O)O[C@@H]1[C@H](O[C@H]([C@@H]1O)[N]2C3=C(N=C2)C(=NC=N3)N)C(O)[P](=O)(O)O[P](=O)(O)O	3'-Uridinemonophosphate
OC[C@H]1O[C@H]([C@H](F)[C@@H]1OP(O)(O)=O)n1ccc(=O)[nH]c1=O	Adenosine-2-5'-diphosphate
OC[C@H]1O[C@H]([C@@H](O)[C@@H]1OP(O)(O)=O)n1ccc(=O)[nH]c1=O	1-(3-O-Phosphono-β-D-arabinofuranosyl)-2,4(1H,3H)-pyrimidinedione
OC[C@H]1O[C@H](C[C@@H]1OP(O)(O)=O)n1ccc(=O)[nH]c1=O	2'-Fluoro-2'-deoxyuridine 3'-monophosphate
Nc1ncnc2n(cnc12)[C@@H]1O[C@H](CO[P@@](O)(=O)O[P@@](O)(=O)O[C@H]2C[C@@H](O[C@@H]2COP(O)(O)=O)n2ccc(=O)[nH]c2=O)[C@@H](OP(O)(O)=O)[C@H]1O	Uracil arabinose-3'-phosphate
[P](=O)(O)(O)O[C@@H]1[C@H](O[C@H]([C@@H]1F)N2C=CC(NC2=O)=O)CO	2'-Deoxyuridine 3'-monophosphate
Nc1ncnc2n(cnc12)[C@@H]1O[C@H](CO[P@@](O)(=O)OP(O)(O)=O)[C@@H](O)[C@H]1O	5'-Phospho-2'-deoxyuridine-3'-pyroposphate adenosine 3'phosphate
[C@@H]1(O)[C@H](O)[C@H](O[C@H]1[N]2C3=C(N=C2)C(=O)N=CN3)CO[P](O)(O)=O	Uridine 5′-diphosphate (UDP)

C. Pharmacophore Model and Screening

Generation of a structure based Pharmacophore Model

From the compounds found through database search the crystal structure of Thymidine-3-monophosphate with murine Ribonuclease A (PDB entry 3LXO) was selected for the generation of the pharmacophore model. This complex was taken because of the good resolution of the crystal structure (1,7). Thymidine-3-monophosphate is described as a competitive inhibitor of Ribonuclease A [Doucet et al. 2010]. The overview of the complex is shown in Figure 20. In Figure 21 you can see Thymidine-3-monophosphate in the isolated binding site of Ribonuclease A.

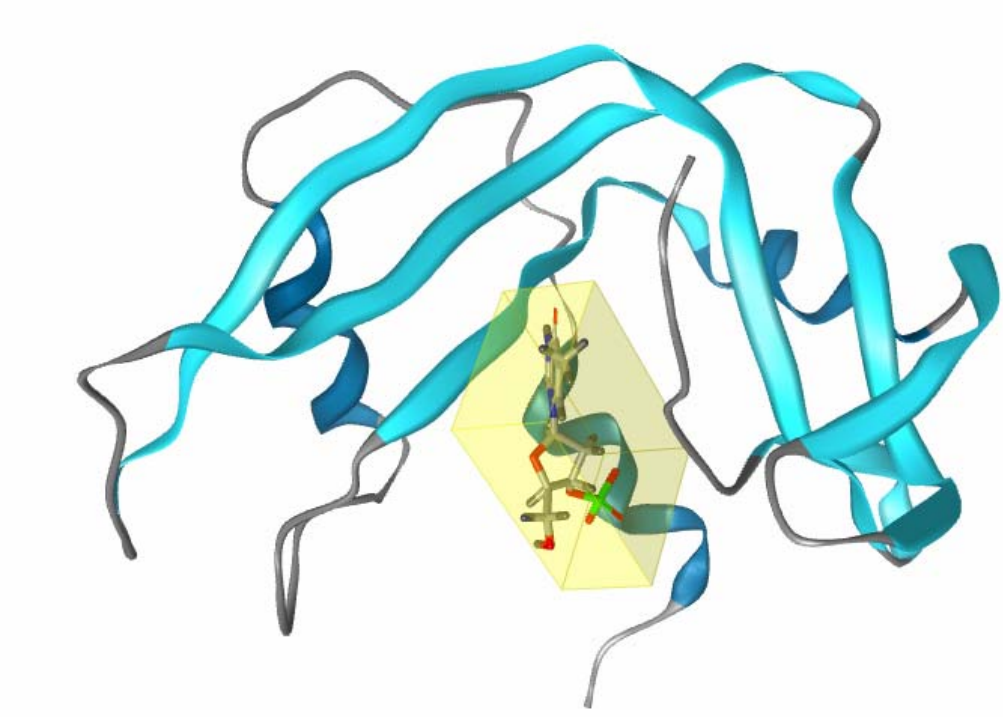


Figure 20. Overview of the complex of Ribonuclease A and Thymidine-3-monophosphate.

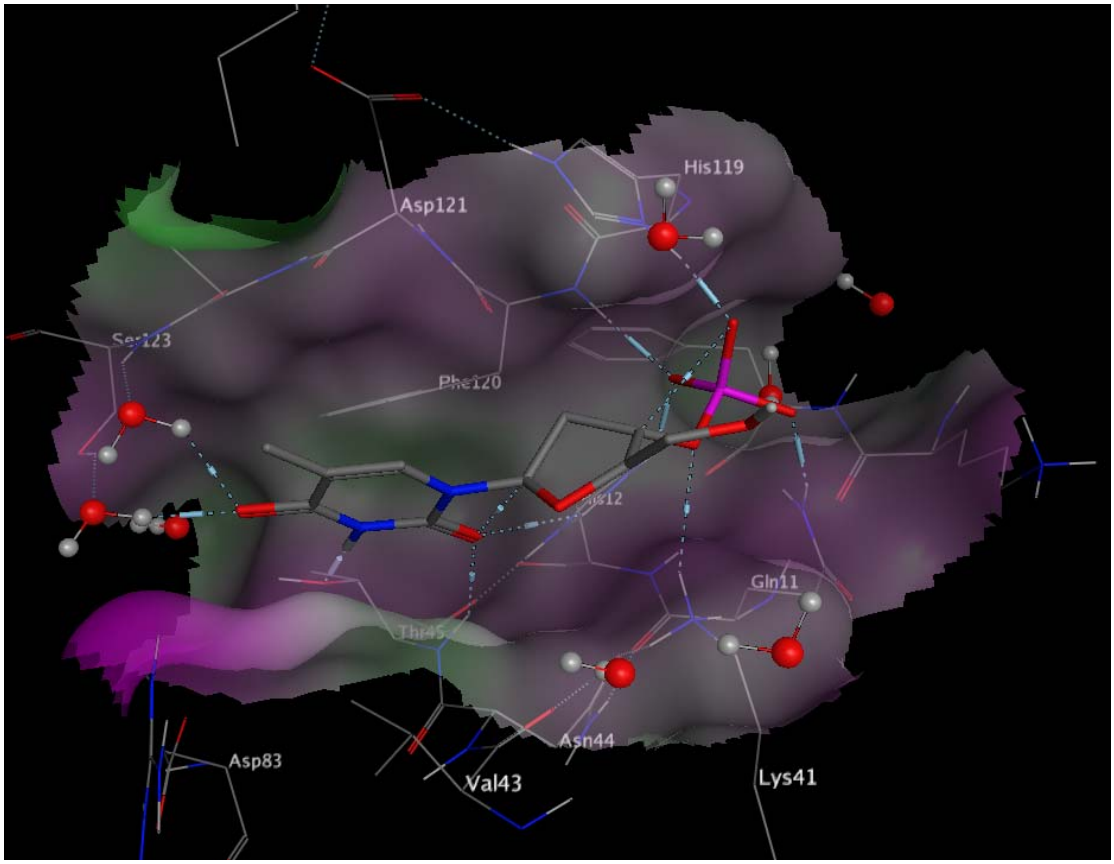


Figure 21. Isolated binding site of Ribonuclease A in complex with Thymidine-3-monophosphate. The surface of the binding site is coloured as following: green represents the lipophilic features, pink the hydrophilic.

Figure 22 represents the exact interactions of the ligand with the protein. Based on these interactions a structure based pharmacophore model was generated using Ligand Scout (Wolber, Langer, 2005). The model, which is shown in Figure 23, is composed of a total of 6 features. The area around the phosphate was defined as negative ionizable feature. As the nitrogen atom is capable to act as a hydrogen donor for Thr 45, a donor feature was placed there. The oxygens from the phosphate and amide moiety are capable of accepting hydrogen bonds, so four acceptor features were considered there. In addition, 12 exclusion volumes were added in order to define the shape of the binding site.

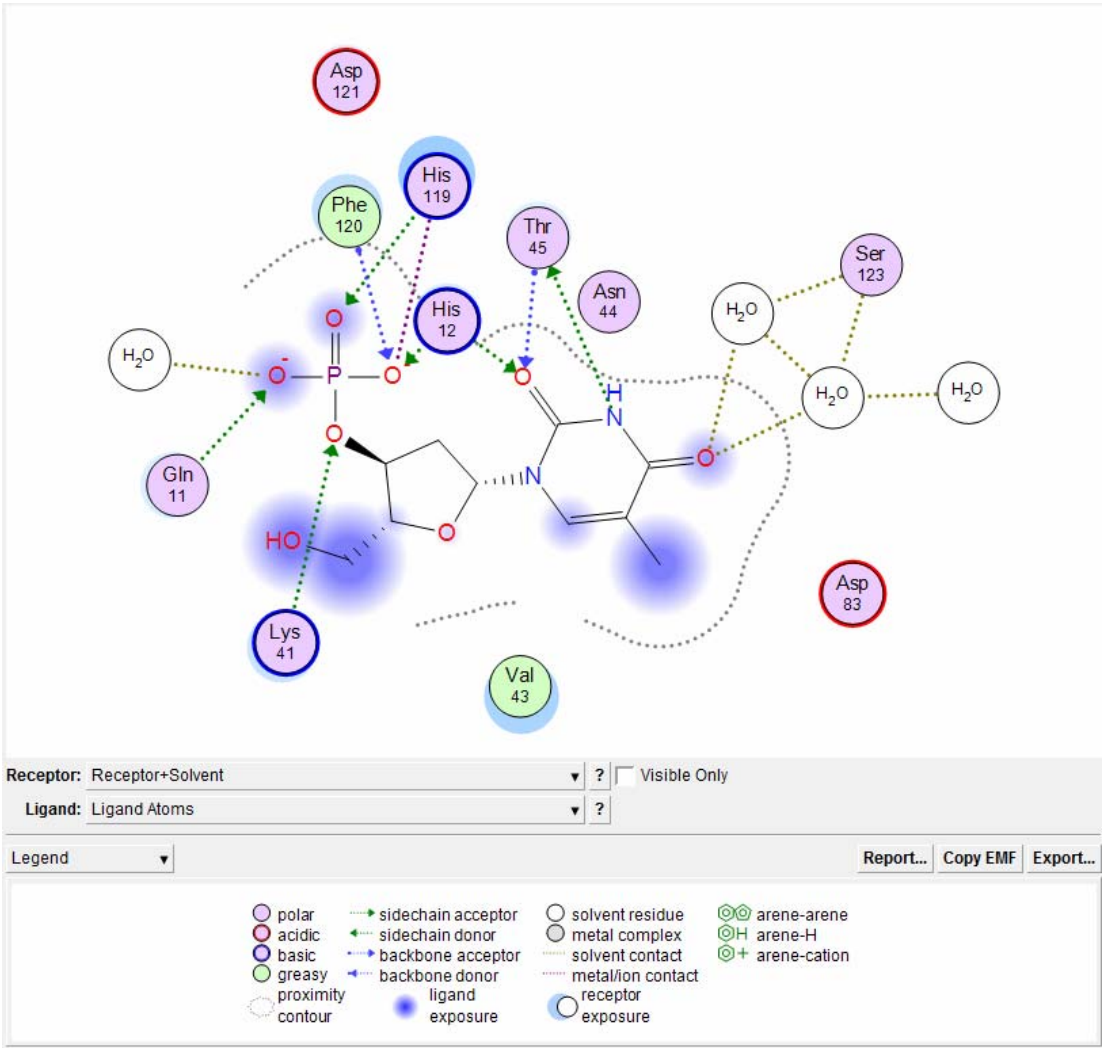


Figure 22. Ligand interactions shown in 2D plot, T3P in binding site of RNase A

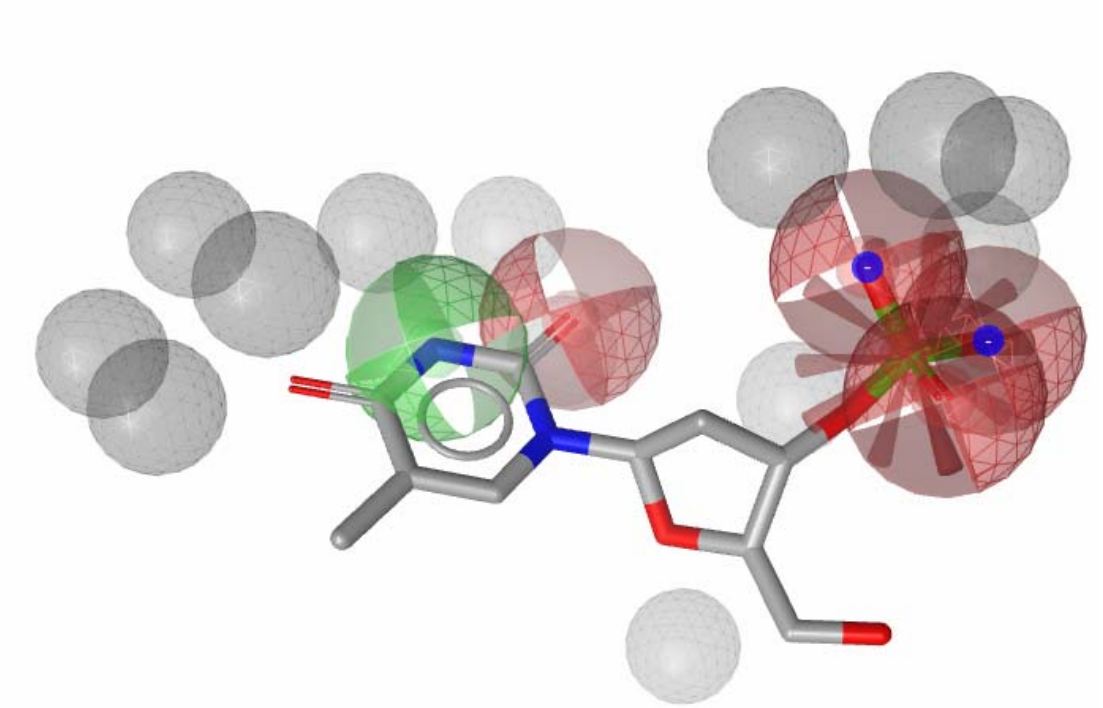


Figure 23. Illustration of the pharmacophore model. The green sphere represents the hydrogen bond donor feature, red spheres represent the hydrogen bon acceptor features. The negative ionizable area is shown through red circular bars. The exclusion volumes are represented by grey spheres.

Screening a Commercial Database

The pharmacophore model was used to screen the LifeChemicals database [Life Chemicals Inc. 2003-2013]. The screening led to identification of 208 compounds. The compounds were further analyzid concerning the Lipinski rule of five. The log P values, the numbers of hydrogen bond acceptors and donors, and the molecular mass were calculated and evaluated. Fifteen compounds did not fullfill the requirements and were exluded from the hit list for the docking experiment. The hits are shown in Annex A.

D. ChemGPS-NP

The compounds found through pharmacophore screening were further analyzed with ChemGPS-NP_{Web} [Rosen et al. 2009]. The hits were compared to the 25 known inhibitors from literature (see section V. B. Results: Database search). For each compound the prediction scores were calculated. The most significant dimensions (the first three of eight) were plotted in a 3D-diagram using MOE [Molecular Operating Environment, 2012.10], see Figure 24.

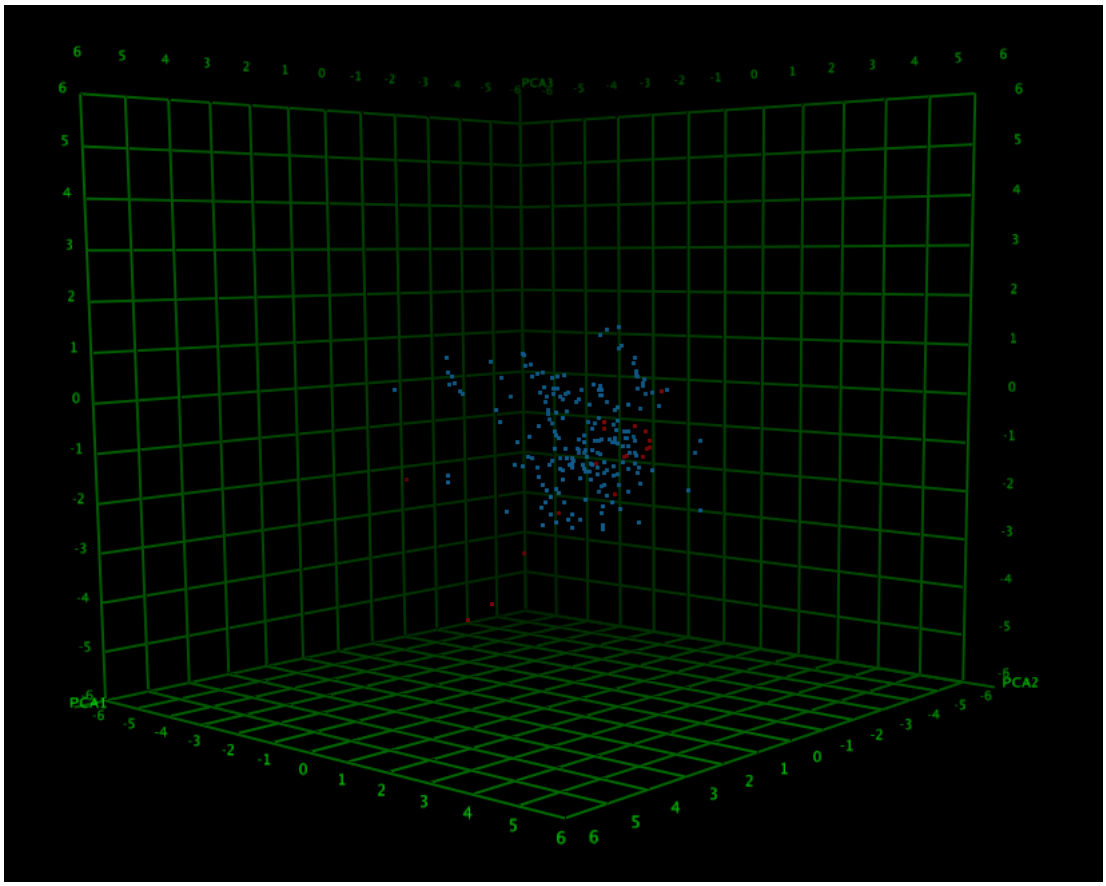


Figure 24. ChemGPS-NP mapping (three first dimensions). Main influence in PCA1 is size, in PCA2 aromaticity and in PCA3 lipophilicity. Blue represents the hit compounds from pharmacophore screening, red represents the reference inhibitors from database search.

What the diagram tells us, is that the physico-chemical properties of the hits from screening and the reference inhibitors are overlapping and are localized in one sector of chemical space, at least in the first three dimensions. There are two outliers that are set aside by a higher prediction score in PCA 1 and a lower prediction score in PCA 3. However, these compounds are two polyphenols. That explains why their physico-chemical properties are different compared to the hit compounds that were found based on a pharmacaphore model from a nucleoside. The detailed results – all coordinates of PCA1-PCA 3, are shown in Table 2.

Table 2. ChemGPS Results: Coordinates of PCA1 – PCA3

MOL	PCA1	PCA2	PCA3
1	-0,84976	0,753497	0,155294
2	-1,530664	-0,999319	-1,289755
3	-1,015383	-0,93725	-2,718328
4	-1,289744	-1,187465	-0,931961
5	1,08266	-0,151633	-1,64872
6	-2,155506	1,243638	-1,10647
7	-0,380016	3,252753	0,760047
8	-0,168353	2,286296	0,209794
9	-1,193897	0,748834	-0,203209
10	-0,519743	0,901522	-0,868427
11	-1,498361	-0,892788	-1,320923
12	-1,472081	-1,053408	-1,701027
13	-1,74007	-0,913367	-1,511952
14	-1,910016	-0,820836	-1,74896
15	-0,878197	-1,402339	-0,429766
16	-0,519743	0,901522	-0,868427
17	-0,523633	0,681811	-0,651345
18	-1,469407	-0,959569	-1,340551
19	-1,50381	-0,890198	-1,316962
20	-1,297246	-0,988938	-1,168987
21	-1,250739	3,131999	-0,477887
22	0,996093	0,507085	0,317345
23	0,823353	0,578811	0,070842
24	1,175243	0,42386	0,476674
25	1,175237	0,421431	0,465648
26	1,35223	0,373383	0,631145
27	1,015994	0,515064	-0,123817
28	2,826184	1,645136	0,946706
29	1,202753	0,394281	0,141203
30	2,864619	1,696821	1,070035
31	2,277189	0,063196	-0,287838
32	1,830884	1,824788	0,799376
33	3,229384	1,502148	1,013751
34	2,531232	4,215555	1,543781
35	2,567757	4,038863	1,534601
36	-1,83418	1,382572	-1,339366
37	-0,767967	1,356945	-1,46162
38	-2,81642	-0,661319	-2,747173
39	0,041554	4,128739	0,563105
40	-3,073197	1,099438	-1,716742
41	-0,193725	2,090257	0,97581
42	-0,312475	2,43453	-1,188911
43	-1,267111	-1,19685	-1,42744

MOL	PCA1	PCA2	PCA3
44	-1,70831	1,120705	-1,366795
45	0,030075	0,832019	-1,20632
46	1,65926	2,083202	0,763419
47	0,675713	1,83306	0,454559
48	2,037822	0,264054	-0,557001
49	-0,670518	0,673825	0,335133
50	-0,837637	0,617328	0,152168
51	-0,41269	-1,504986	-3,948118
52	0,216058	0,393064	-1,779172
53	3,033115	1,544607	1,092308
54	3,205526	1,406087	1,591341
55	3,23848	1,408222	1,215014
56	3,06668	1,56909	1,28459
57	4,37006	2,644889	2,047089
58	2,500085	0,152105	0,18069
59	0,077403	2,165591	0,726682
60	0,044144	2,20996	0,679027
61	-0,737297	0,397102	0,006373
62	0,986502	3,102018	0,866463
63	0,129439	1,778741	0,762668
64	-0,841101	1,142902	-1,312595
65	-0,753659	1,195579	-1,358362
66	-0,999041	1,06581	-0,612246
67	0,812366	0,36105	-1,979873
68	0,808162	0,676354	-0,400872
69	0,516508	0,67159	0,168601
70	0,376543	0,285687	-0,374529
71	0,545009	1,564138	-0,722182
72	-0,470803	0,784555	-2,176078
73	-0,384561	0,629151	-1,876169
74	0,049738	0,384994	-1,774015
75	-0,7846	0,27623	-2,672906
76	-0,378319	0,831996	-2,282135
77	-0,24794	0,126231	-2,043075
78	-0,137208	3,130759	0,270415
79	-1,281668	1,735207	-0,519681
80	1,410344	0,294557	0,288899
81	1,383411	0,330212	0,646536
82	0,439321	2,908473	0,363892
83	0,258572	3,003192	0,175017
84	0,272694	3,201592	0,07664
85	-0,566145	1,771247	-0,305674
86	-0,327176	0,390583	-0,485587

MOL	PCA1	PCA2	PCA3
87	-1,197977	0,67418	-0,544088
88	-0,363489	1,633975	-0,037369
89	-0,118441	0,264515	-0,285176
90	0,062213	1,981866	0,370108
91	-0,117726	2,070925	0,182528
92	-1,721426	1,088582	-0,946762
93	-0,88427	2,025618	-0,43042
94	-0,350166	2,325995	0,140892
95	-0,560817	2,473917	-0,059803
96	-0,640532	0,592264	-0,598862
97	-0,536601	2,481352	-0,05113
98	-0,88427	2,025618	-0,43042
99	-0,671247	2,639289	-0,433076
100	-3,013551	-0,63086	-2,378834
101	-3,261984	0,890195	-2,486406
102	-1,365135	0,233135	-1,851724
103	-1,008494	2,483396	-0,898477
104	0,162011	0,855255	-0,202035
105	0,334718	0,75168	0,046796
106	0,966176	0,290737	-0,907474
107	-0,850954	0,174516	-1,222572
108	-0,284328	1,623466	-0,286833
109	-2,179235	-0,939801	-1,444878
110	-0,106986	0,560826	-0,207176
111	-0,017189	0,585718	-0,248588
112	0,506565	0,293577	0,05902
113	0,247205	0,414213	0,228532
114	0,958029	0,204439	0,056191
115	-1,325219	-0,949126	-1,484766
116	0,703602	0,212246	-0,064499
117	-0,337158	-1,297379	-0,542109
118	-1,312366	-1,000703	-1,087174
119	0,942993	0,232838	-0,545635
120	0,506565	0,293577	0,05902
121	-0,568861	-1,182857	-0,274273
122	0,521973	0,402097	-0,130075
123	-0,875804	-1,093646	-0,64423
124	0,659945	0,438916	-0,239225
125	0,18989	0,480825	-0,110323
126	-0,932549	-1,177588	-0,656387
127	0,340628	0,590794	-0,177863
128	0,935095	0,328057	-0,998632
129	0,703602	0,212246	-0,064499

MOL	PCA1	PCA2	PCA3
130	0,942993	0,232838	-0,545635
131	-0,665664	1,035257	-1,055622
132	-1,595687	0,78694	0,002705
133	-0,626343	1,954973	0,282059
134	-1,342336	0,979162	0,106016
135	-0,54326	1,99767	-0,43476
136	-1,13728	0,7939	-0,670503
137	-1,329998	0,660087	-0,82793
138	-0,323079	2,367566	0,052146
139	-0,313678	2,204885	0,124174
140	-0,736419	1,887379	-0,356511
141	-1,352527	2,530839	-0,564992
142	-1,062072	2,7577	-1,024927
143	-0,611171	1,860773	-0,583063
144	0,894733	2,096336	-0,081913
145	1,329197	3,503179	0,101369
146	0,867365	1,948447	-0,491286
147	-0,60923	0,8099	-0,94853
148	-0,261436	0,555455	-0,779933
149	-1,513906	2,415802	-0,857257
150	-0,254327	2,326944	0,076508
151	0,036916	3,36037	0,398084
152	-0,705326	2,483712	0,300745
153	-0,710523	2,442447	0,141234
154	-0,917594	2,6376	0,093667
155	-1,238453	2,892228	-0,26326
156	0,217056	3,264834	0,584905
157	-1,929506	2,272327	-1,322895
158	-1,121719	1,992836	-0,157306
159	0,782189	3,049787	0,790058
160	-0,771754	1,841912	-0,037448
161	-0,538135	1,714338	-0,210652
162	-0,755362	2,072771	-0,137074
163	-0,771754	1,841912	-0,037448
164	0,778794	2,664336	0,592569
165	1,093631	2,451994	0,706516
166	-1,23901	-1,116766	-2,169864
167	0,880228	2,078239	0,452936
168	-0,764663	0,959033	-0,732599
169	-0,401557	3,126929	-0,2679
170	0,352888	1,755873	-0,187053
171	-1,054768	1,999057	-1,002854
172	-0,039684	3,461987	-0,124398

MOL	PCA1	PCA2	PCA3
173	-1,267378	-1,09514	-0,997782
174	-0,516967	-0,093703	-0,760932
175	-0,709722	0,017763	-0,815198
176	-0,333482	3,086867	-0,648126
177	-0,55035	0,995156	-0,439316
178	0,836976	0,628703	-0,253626
179	-0,311384	3,201009	-0,328728
180	-0,972672	0,769781	-1,022012
181	-1,305799	0,522471	-1,140558
182	-1,431269	0,647685	-0,831412
183	-1,606899	0,744561	-0,973413
184	-1,344271	0,946004	-0,894074
185	0,120191	2,284131	-0,946048
186	-1,486371	0,907979	-1,092336
187	-1,014871	0,828276	-1,782713
188	-1,45712	0,931709	-0,915294
189	-1,607697	0,938289	-1,314897
190	-1,332129	1,080013	-1,912588
191	-1,389956	0,813989	-0,818398
192	-0,186937	2,584214	-0,885136
193	-1,791702	1,23457	-1,10402
194	-0,346216	2,248436	0,111615
195	-1,337587	1,186207	-1,051775
196	-0,051022	2,43037	0,054519
197	-1,098036	0,911106	-0,83823
198	-1,047821	0,750931	-0,929185
199	0,205111	3,169027	-0,265926
200	-0,034725	2,499231	0,17342
201	-1,445617	1,165099	-1,214569
202	-1,148228	0,858366	-0,918006
203	0,444583	3,134696	-0,161868
204	0,308641	2,399958	0,130779
205	-0,209172	0,794752	-0,854146
206	-0,749934	0,614795	0,129416
207	0,190395	2,465468	-0,404263
208	-0,938453	0,957618	-1,679628

Reference Inhibitors from literature

MOL	PCA1	PCA2	PCA3
1	-1,261027	-1,293171	-2,947422
2	-0,41269	-1,504986	-3,948118
3	-0,739898	0,564653	-0,727449
4	-1,973222	-1,075141	-1,638865
5	-1,026708	-1,390621	-3,763394
6	0,590169	-1,719376	-4,418005
7	1,984708	4,158894	-1,681825
8	2,837957	4,152544	-2,220777
9	-0,016935	-1,257944	-2,939174
10	-1,026708	-1,390621	-3,763394
11	2,777355	-0,274119	-5,530574
12	2,777355	-0,274119	-5,530574
13	1,538824	-0,522412	-4,550096
14	-0,488243	-1,490138	-3,889643
15	-0,488243	-1,490138	-3,889643
16	1,538824	-0,522412	-4,550096
17	-0,983263	-1,233567	-2,757886
18	2,777355	-0,274119	-5,530574
19	-1,248947	-1,204745	-3,338929
20	-0,488243	-1,490138	-3,889643
21	-1,337026	-1,298785	-3,222661
22	7,112151	-1,871694	-5,537524
23	-1,343671	-1,167648	-3,241688
24	1,538824	-0,522412	-4,550096
25	-0,434203	-0,234722	-3,525882

E. Docking

The docking was performed with the software GOLD Suite 5.1 [Verdonk et al. 2003]. The detailed workflow is described in the methods section. To estimate the best scoring function a redocking was performed.

Redocking

For validation of the method, the ligand Thymidine-3-monophosphate was taken out of the crystal structure from the complex with Ribonuclease A, PDB entry 3LX), docked into the binding site and scored with different functions. The binding poses were analyzed and compared to the original binding pose from the PDB entry.

- GoldScore, see Figure 25
- ChemScore, see Figure 26
- ChemPLP (Piecewise Linear Potential), see Figure 27
- ASP (Astex Statistical Potential), see Figure 28

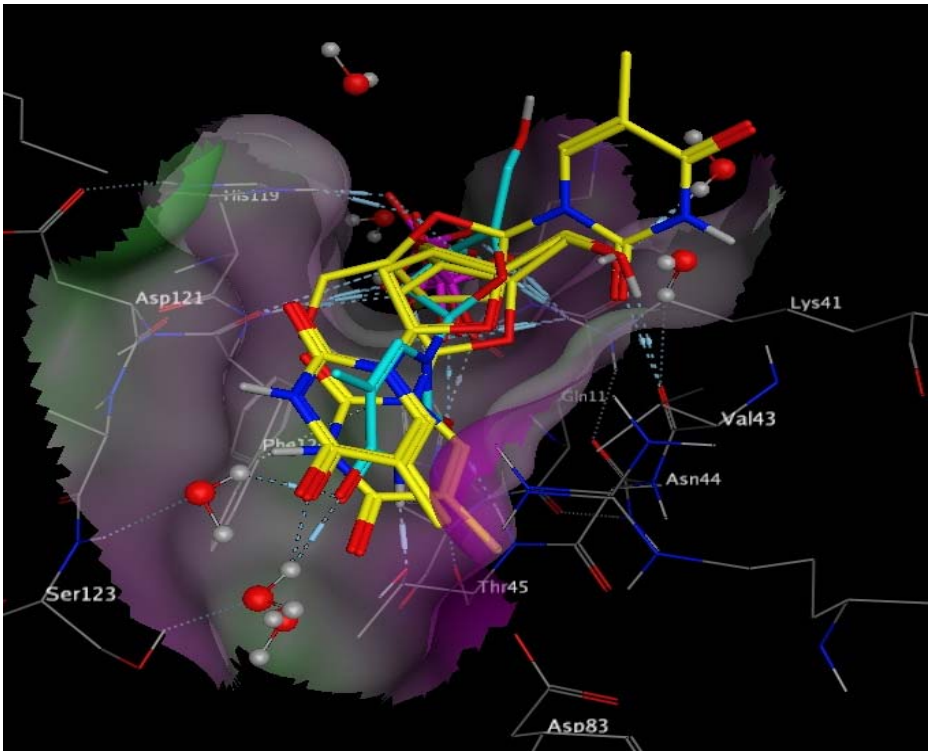


Figure 25. Redocking results, scored with Goldscore. Original binding pose is colored blue.

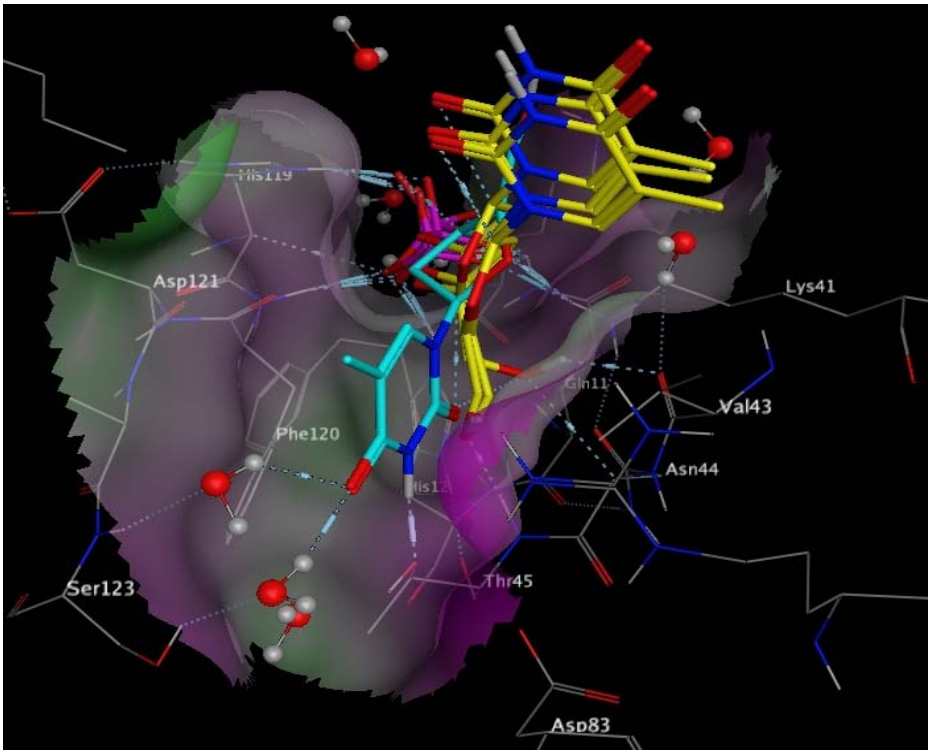


Figure 26. Redocking results, scored with Chemscore. Original binding pose is colored blue.

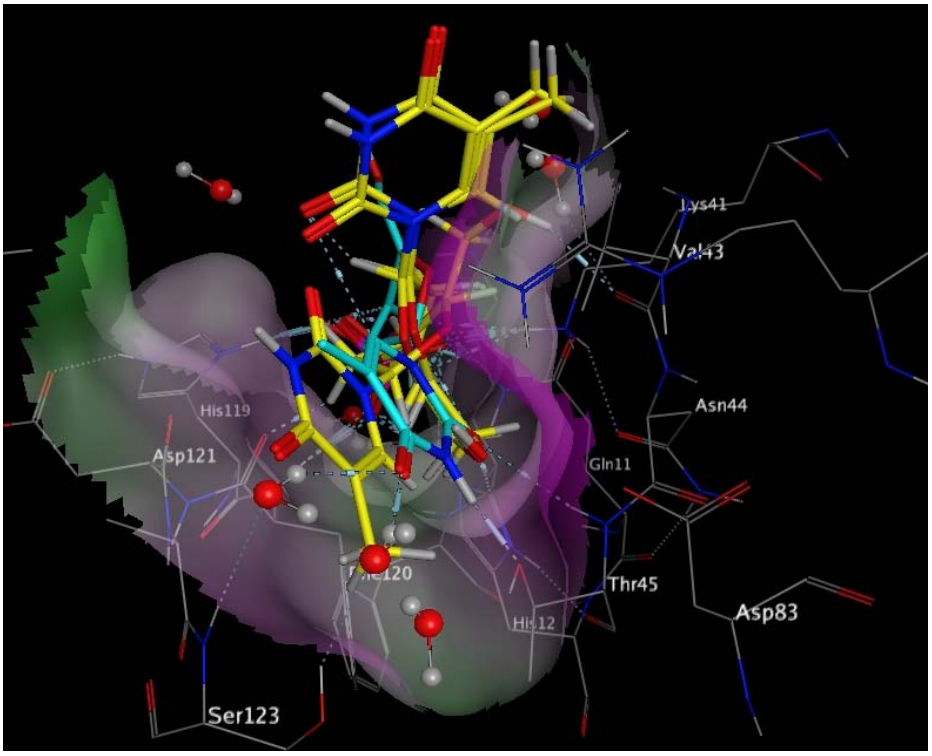


Figure 27. Redocking results, scored with ChemPLP. Original binding pose is colored blue.

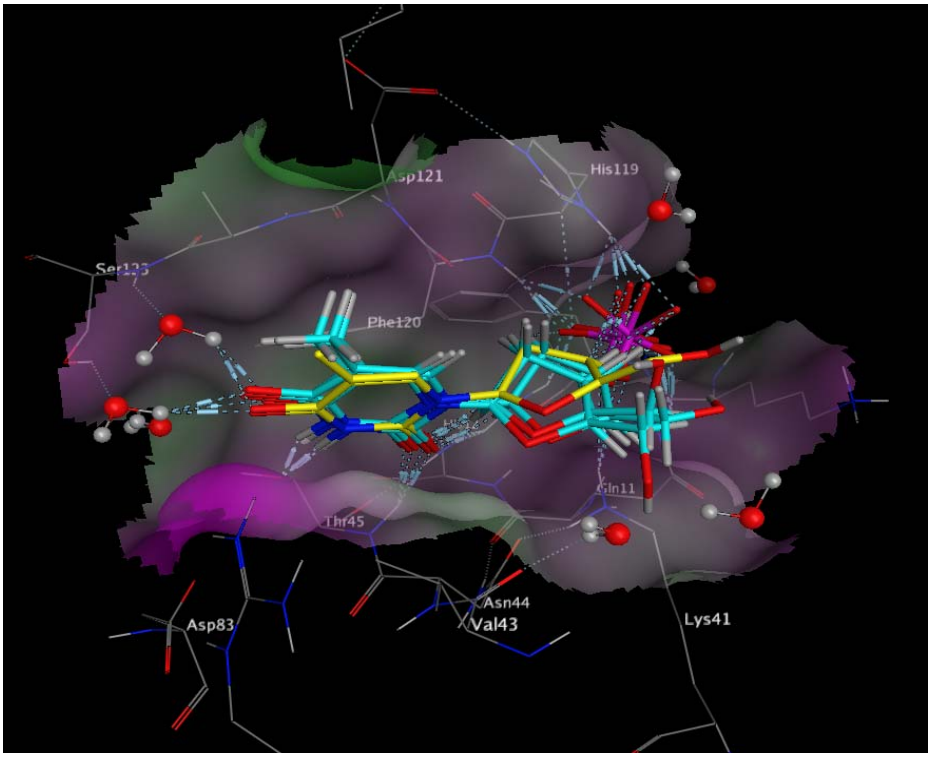


Figure 28. Redocking results, scored with ASP. Original binding pose is colored blue.

As can be seen in Figure 28, the docking poses that matched best with the original binding pose were generated with the Astex Statistical Potential (ASP) scoring function. Therefore ASP was used as scoring function for the docking of the 193 compound (the outcome of the pharmacophore screening).

Docking results

The run produced 1544 poses in total. The compounds were visually analyzed and classified into 42 scaffolds. The scaffold numbering was assigned considering the best docking pose for each compound, for example: the compound with the best docking pose was determined as scaffold 1, the next different scaffold compound with the next best docking pose as scaffold 2 and so on.

The 193 compounds are shown in Annex A ranked downwards by the docking results.

Compound Selection for Experimental Testing

For the experimental testing 51 compounds were selected, 42 from each scaffold one and 9 in addition from the scaffolds ranked best through docking. At the time of the thesis writing the compounds were being tested by the Medical University of Vienna.

VI. Conclusion and Outlook

As result of this study, compounds could be identified by in silico methods, that are considered to act as Ribonuclease A inhibitors. Ribonuclease A is associated with obesity and findings show that the inhibition of RNase A leads to a decrease of the number of adipocytes.

After sequence alignment a structure-based pharmacophore model was generated. With this model the LifeChemicals database was screened. The hit compounds were further analyzed with ChemGPS-NP to compare them with already known inhibitors of Ribonuclease A. This was followed by docking studies that were performed in GOLD. Out of 194 hits, 51 compounds were selected taking into consideration their docking scores and diversity with respect to the chemical scaffold.

The identified compounds are currently tested in-vitro. Depending on the results, several follow up studies are possible. The compounds tested positive for inhibition should be further analyzed, to retrieve a potential lead candidate. This could be further tested in mice for a proof of concept.

VII. References

Arner P, Spalding KL. Fat cell turnover in humans. Biochemical and Biophysical Research Communications 2010;396,101-104

Bernstein FC, Koetzle TF, Williams GJ, Meyer EE, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: A Computer-based Archival File For Macromolecular Structures. J. of. Mol. Biol. 1977; 112: 535.

Bilban M, Esterbauer H, Wagner O. Antagonists of Ribonucleases for Treating Obesity. Patent WO2013017656 (A1) 2013

Carvajal R, Wadden TA, Tsai AG, Peck K, Moran CH. Managing obesity in primary care practice: a narrative review. Ann. N.Y. Acad. Sci. 2013;1281,191–206

ClustalW2 2007 http://www.ebi.ac.uk/Tools/msa/clustalw2/ accessed August 2013

Després JP, Lemieux I. Abdominal obesity and metabolic syndrome, Nature 2006;444,881-887

Doucet N, Jayasundera TB, Simonović M, Loria JP. The crystal structure of ribonuclease A in complex with thymidine-3'-monophosphate provides further insight into ligand binding. Proteins 2010;78(11):2459-68.

Dyer KD, Rosenberg HF. The RNase a superfamily: Generation of diversity and innate host defense. Molecular Diversity 2006;10: 585–597

Ehrlich P. Ueber den jetzigen Stand der Chemotherapie.Ber. Dtsch. Chem. Ges. 1909;42, 17–47

Friedman JM. Causes and control of excess body fat. Nature 2009;459(21),340-342

Gao Q, Yang L, and Zhu Y. Pharmacophore Based Drug Design Approach as a Practical Process in Drug Discovery, Current Computer Aided-Drug Design 2010;6, 37-49

Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Research 2011; 40(D1):D1100-D1107.

Goo SM, Cho S. The Expansion and Functional Diversification of the Mammalian Ribonuclease A Superfamily Epitomizes the Efficiency of Multigene Families at Generating Biological Novelty. Genome Biol Evol 2013;5 (11): 2124-2140.

Haslam D, James P. Obesity, The Lancet 2005;366, 1197-1209

http://cactus.nci.nih.gov/chemical/structure accessed July 2013

http://www.ncbi.nlm.nih.gov/pubmed accessed July 2013

Koppelman PG. Obesity as a medical problem. Nature 2000;404,635-643

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG. ClustalW and ClustalX version 2. Bioinformatics 2007;23(21): 2947-2948.

Larsson J, Gottfries J, Muresan S, Backlund A. ChemGPS-NP: tuned for navigation in biologically relevant chemical space. J Nat Prod. 2007;70(5):789-794.

Leach AR, Shoichet BR, Peishoff C. Prediction of Protein−Ligand Interactions. Docking and Scoring: Successes and Gaps. J. Med. Chem. 2006, 49 (20), 5851–5855

Life Chemicals Inc. 2003-2013 http://www.lifechemicals.com/

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Delivery Rev 1997;23, 3–25.

Marshall GR, Feng JA, Kuster DJ. Back to the future: Ribonuclease A. Peptide Science 2008; 90-3, 259–277

Molecular Operating Environment (MOE), 2012.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2012.

Moouj WTM, Verdonk ML. General and targeted statistical potentials for protein–ligand interactions. Proteins: Structure, Function and Bioinformatics 2005;61(2);272-287

Rosen J, Loevgren A, Kogej T, Muresan S, Gottfries J, Backlund A. ChemGPS-NPweb: chemical space navigation online. J Comput Aided Mol Des 2009;23,253-259

Samanta A, Dasgupta S, Pathak T. 5′-Modified pyrimidine nucleosides as inhibitors of ribonuclease A. Bioorganic & Medicinal Chemistry 2011;19-7,2478–2484

Santoro, J., Gonzalez, C., Bruix, M., Neira, JL., Nieto, JL., Herranz, J., Rico, M. High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy. Journal: J.Mol.Biol. 1993;229: 722-734

Uniprot © 2002–2013. The UniProt Consortium. Update on activities at the Universal Protein Resource (UniProt) in 2013. Nucleic Acids Res. 41: D43-D47 (2013). http://www.uniprot.org/ accessed May 2013

Van Gaal LF, Mertens I, De Block C. Mechanisms linking obesity with cardiovascular disease, Nature 2006;444, 875-880

Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved Protein-Ligand Docking Using GOLD. Proteins 2003;52, 609-623.

Wermuth CG et al., Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997), Annu. Rep. Med. Chem. 1998; 33, 385–395

WHO, Obesity and overweight Fact sheet N°311, Updated March 2013 http://www.who.int/mediacentre/factsheets/fs311/en/index.html

WHO, Obesity: preventing and managing the global epidemic. Report of a WHO Consultation*. In:* WHO Technical Report Series*. 894, 2000*

Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 2005;45(1):160–169.

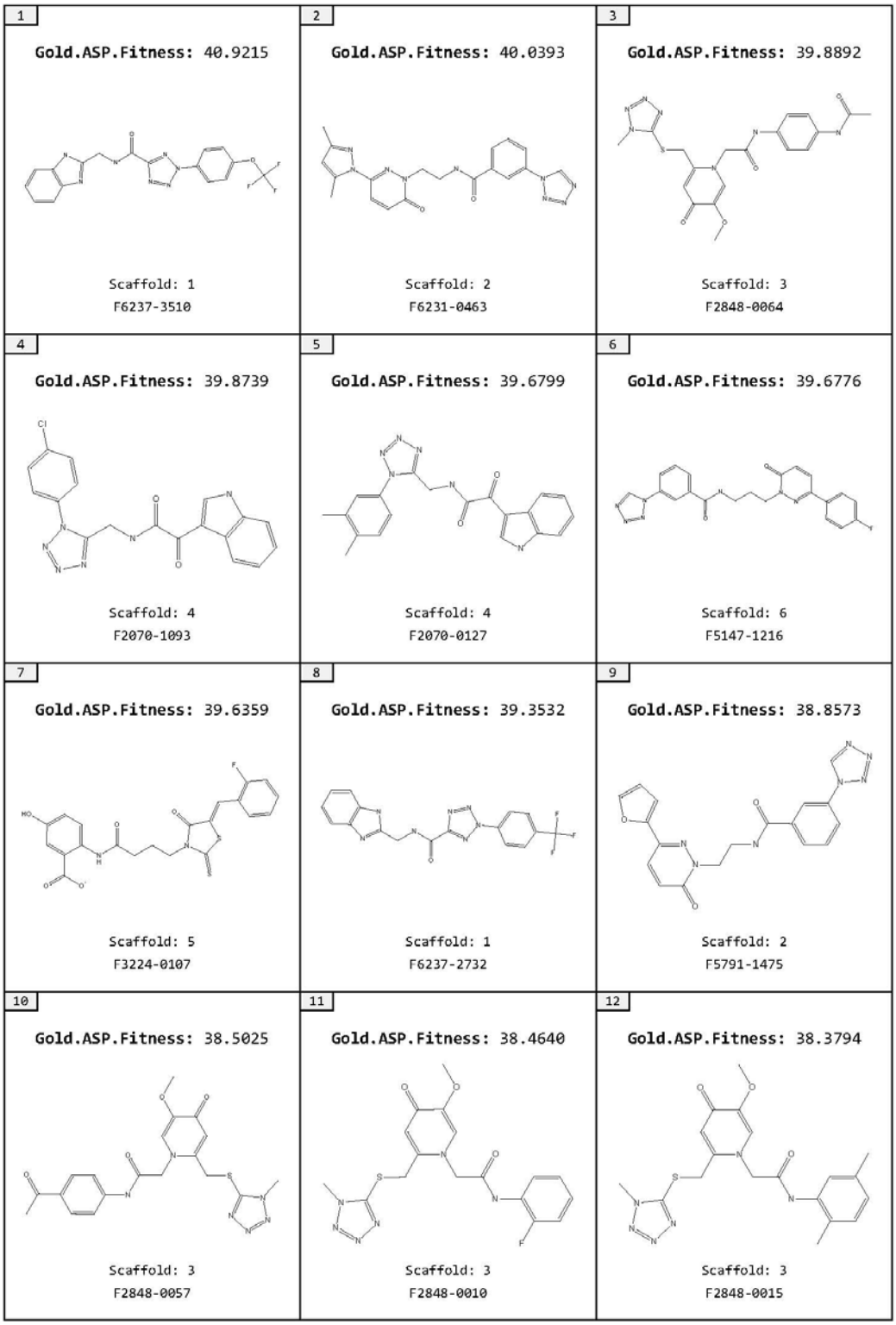
Wolber, G; Kirchmair H; Langer T. Structure-Based 3D Pharmacophores: An Alternative to Docking? Plenary talk at the 7th International Conference on Chemical Structures in Noordwijkerhout, 2005

Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discovery Today 2010;15, 444–450

Yanovski SZ, Yanovski JA. Long-term Drug Treatment for Obesity. A Systematic and Clinical Review, JAMA Published online November 14, 2013

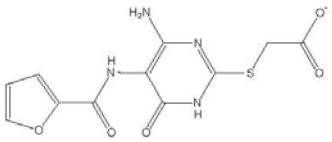
VIII. Annex

A. Hit compounds (ID Numbers from LifeChemicals)



13

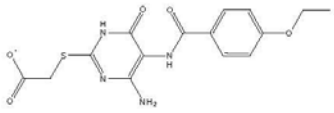
Gold.ASP.Fitness: 38.2943



Scaffold: 7
F1880-1727

14

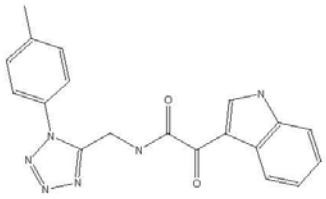
Gold.ASP.Fitness: 38.2779



Scaffold: 7
F1880-1183

15

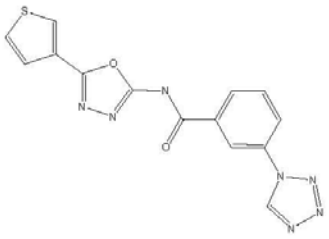
Gold.ASP.Fitness: 38.1420



Scaffold: 4
F2070-0610

16

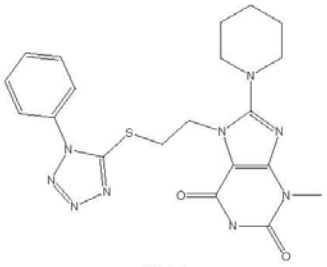
Gold.ASP.Fitness: 37.9350



Scaffold: 8
F5945-0001

17

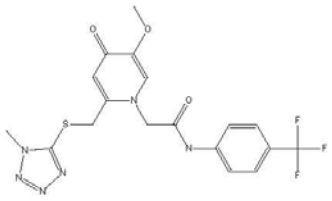
Gold.ASP.Fitness: 37.6627



Scaffold: 9
F0570-0018

18

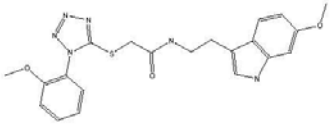
Gold.ASP.Fitness: 37.5817



Scaffold: 3
F2848-0056

19

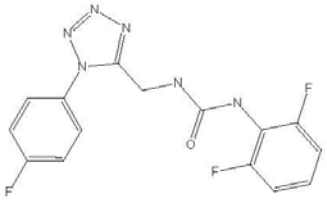
Gold.ASP.Fitness: 37.3015



Scaffold: 10
F3375-2595

20

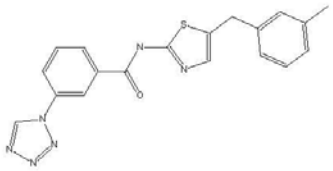
Gold.ASP.Fitness: 37.2392



Scaffold: 11
F2072-0373

21

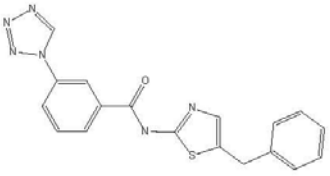
Gold.ASP.Fitness: 37.2168



Scaffold: 8
F3296-0550

22

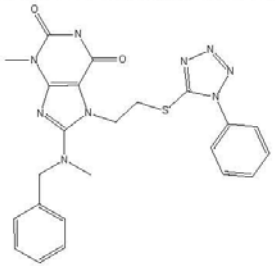
Gold.ASP.Fitness: 37.1834



Scaffold: 8
F3296-0281

23

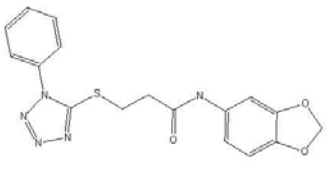
Gold.ASP.Fitness: 37.1635



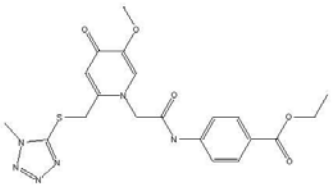
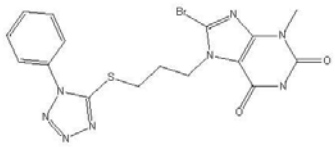
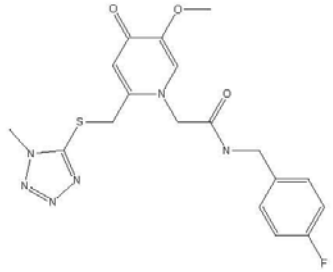
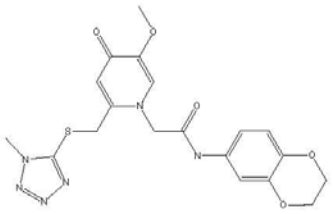
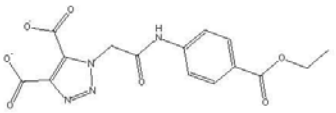
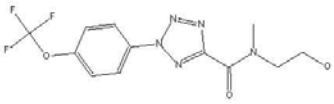
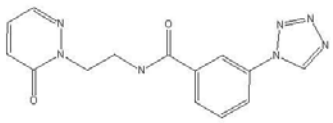
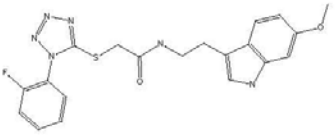
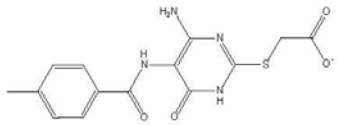
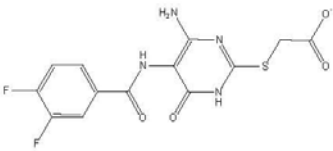
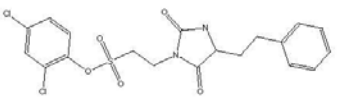
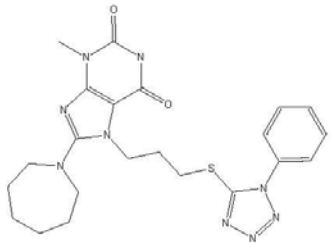
Scaffold: 9
F0570-0144

24

Gold.ASP.Fitness: 37.1070

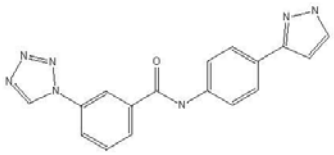


Scaffold: 10
F0307-0544

25 **Gold.ASP.Fitness:** 37.0969  Scaffold: 3 F2848-0007	26 **Gold.ASP.Fitness:** 37.0704  Scaffold: 9 F2582-0095	27 **Gold.ASP.Fitness:** 36.9904  Scaffold: 3 F2848-0035
28 **Gold.ASP.Fitness:** 36.9544  Scaffold: 3 F2848-0026	29 **Gold.ASP.Fitness:** 36.8980  Scaffold: 12 F1604-0153	30 **Gold.ASP.Fitness:** 36.8662  Scaffold: 1 F6237-3529
31 **Gold.ASP.Fitness:** 36.7491  Scaffold: 13 F5283-0183	32 **Gold.ASP.Fitness:** 36.7434  Scaffold: 10 F3375-1988	33 **Gold.ASP.Fitness:** 36.5488  Scaffold: 7 F1880-0881
34 **Gold.ASP.Fitness:** 36.5397  Scaffold: 7 F1880-1373	35 **Gold.ASP.Fitness:** 36.5334  Scaffold: 14 F1678-0103	36 **Gold.ASP.Fitness:** 36.5278  Scaffold: 9 F1889-0414

37

Gold.ASP.Fitness: 36.4529

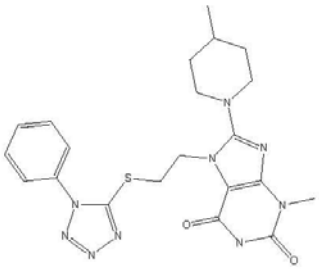


Scaffold: 8

F5835-0571

38

Gold.ASP.Fitness: 36.4021

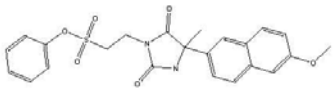


Scaffold: 9

F0570-0046

39

Gold.ASP.Fitness: 36.3265

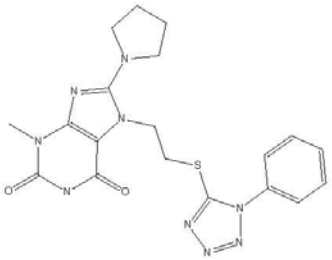


Scaffold: 14

F1678-0206

40

Gold.ASP.Fitness: 36.2465

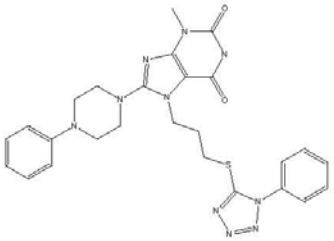


Scaffold: 9

F0570-0025

41

Gold.ASP.Fitness: 36.2045

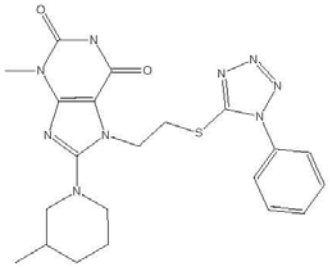


Scaffold: 9

F1614-0056

42

Gold.ASP.Fitness: 36.1970

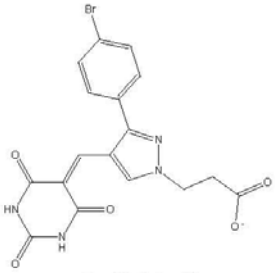


Scaffold: 9

F0570-0039

43

Gold.ASP.Fitness: 36.1458

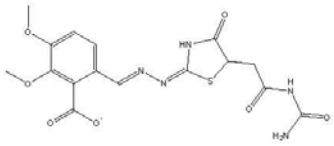


Scaffold: 15

F1218-0127

44

Gold.ASP.Fitness: 36.1278

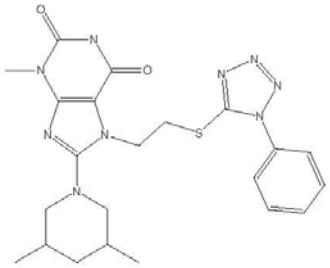


Scaffold: 16

F1726-0193

45

Gold.ASP.Fitness: 36.0294

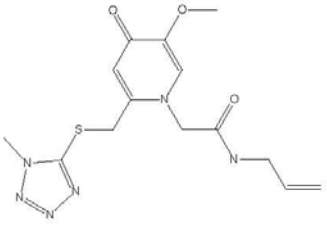


Scaffold: 9

F0570-0060

46

Gold.ASP.Fitness: 35.9908

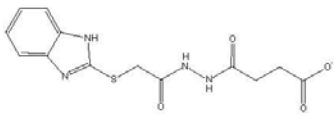


Scaffold: 17

F2848-0072

47

Gold.ASP.Fitness: 35.9535

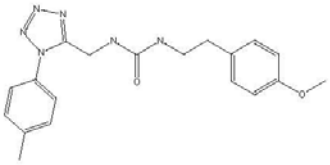


Scaffold: 18

F0856-0070

48

Gold.ASP.Fitness: 35.9178

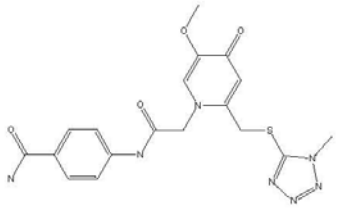


Scaffold: 11

F2072-0139

49

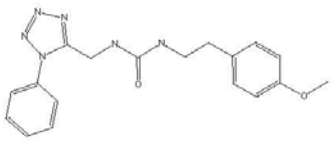
Gold.ASP.Fitness: 35.8851



Scaffold: 3
F2848-0065

50

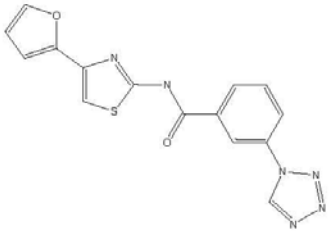
Gold.ASP.Fitness: 35.8373



Scaffold: 11
F2072-0174

51

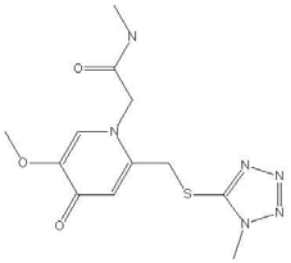
Gold.ASP.Fitness: 35.7507



Scaffold: 8
F5791-1447

52

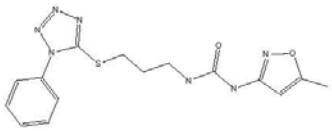
Gold.ASP.Fitness: 35.6792



Scaffold: 17
F2848-0067

53

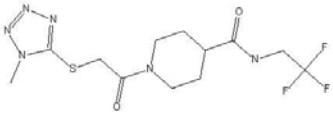
Gold.ASP.Fitness: 35.6766



Scaffold: 11
F2778-0029

54

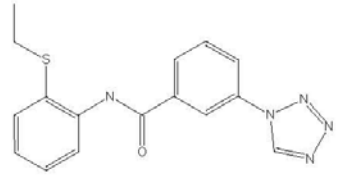
Gold.ASP.Fitness: 35.6609



Scaffold: 19
F5871-1502

55

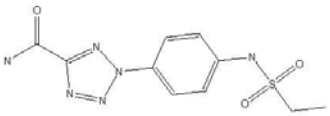
Gold.ASP.Fitness: 35.5020



Scaffold: 13
F3296-0285

56

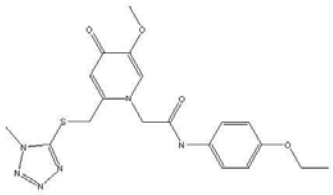
Gold.ASP.Fitness: 35.4864



Scaffold: 20
F6242-0250

57

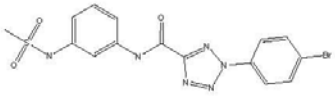
Gold.ASP.Fitness: 35.4203



Scaffold: 3
F2848-0013

58

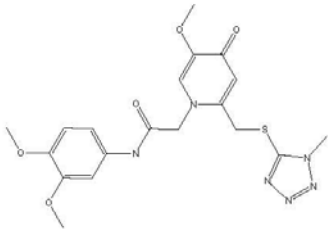
Gold.ASP.Fitness: 35.4039



Scaffold: 20
F6237-4230

59

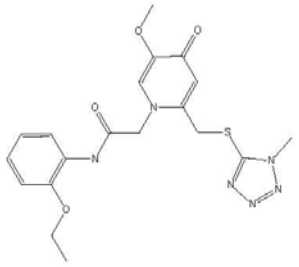
Gold.ASP.Fitness: 35.3974



Scaffold: 3
F2848-0062

60

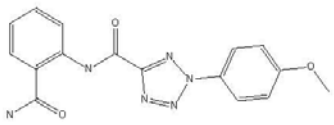
Gold.ASP.Fitness: 35.3524



Scaffold: 3
F2848-0046

61

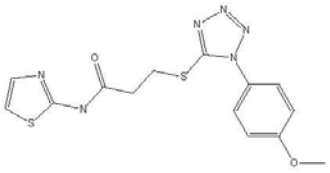
Gold.ASP.Fitness: 35.3290



Scaffold: 20
F6237-0982

62

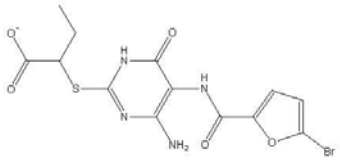
Gold.ASP.Fitness: 35.2255



Scaffold: 10
F3375-2046

63

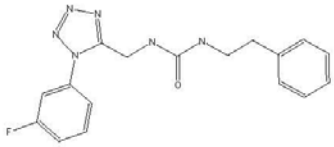
Gold.ASP.Fitness: 35.2194



Scaffold: 7
F1880-1867

64

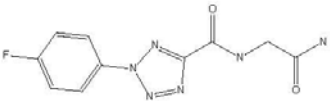
Gold.ASP.Fitness: 35.0684



Scaffold: 11
F2072-0195

65

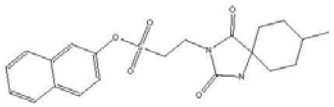
Gold.ASP.Fitness: 35.0627



Scaffold: 20
F6237-1865

66

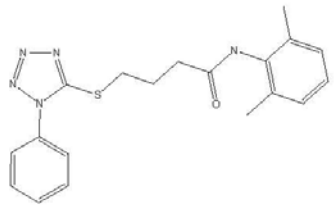
Gold.ASP.Fitness: 34.9859



Scaffold: 21
F1678-0313

67

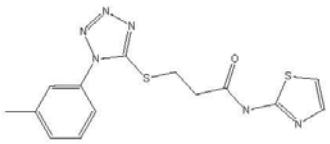
Gold.ASP.Fitness: 34.9468



Scaffold: 10
F1142-4862

68

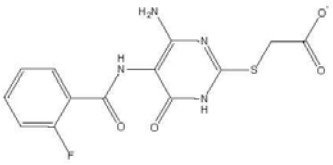
Gold.ASP.Fitness: 34.9423



Scaffold: 10
F3375-2055

69

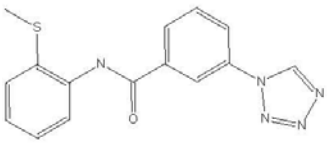
Gold.ASP.Fitness: 34.8324



Scaffold: 7
F1880-0428

70

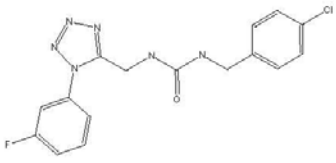
Gold.ASP.Fitness: 34.8259



Scaffold: 13
F3296-0338

71

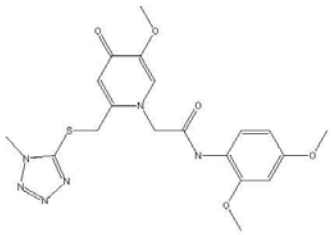
Gold.ASP.Fitness: 34.7885



Scaffold: 11
F2072-0207

72

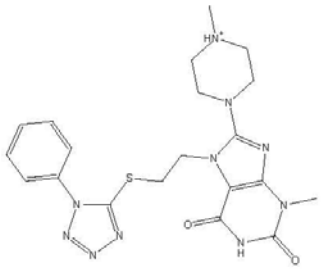
Gold.ASP.Fitness: 34.7881



Scaffold: 3
F2848-0024

73

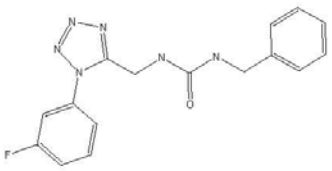
Gold.ASP.Fitness: 34.7664



Scaffold: 9
F0570-0088

74

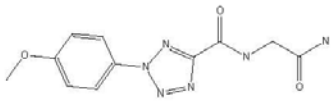
Gold.ASP.Fitness: 34.7430



Scaffold: 11
F2072-0201

75

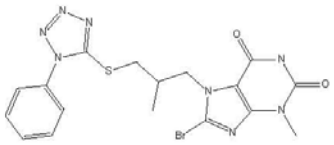
Gold.ASP.Fitness: 34.7159



Scaffold: 20
F6237-1087

76

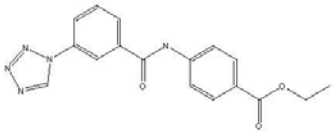
Gold.ASP.Fitness: 34.7006



Scaffold: 9
F2582-0096

77

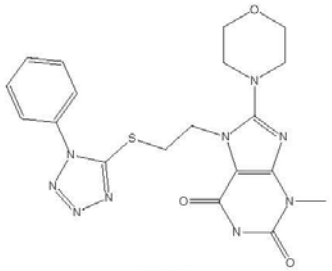
Gold.ASP.Fitness: 34.6899



Scaffold: 13
F3296-0263

78

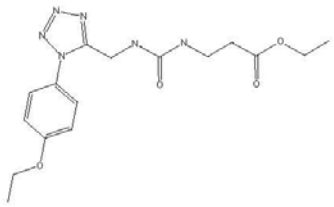
Gold.ASP.Fitness: 34.6763



Scaffold: 9
F0570-0074

79

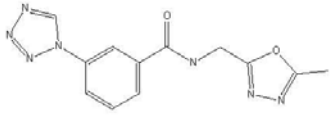
Gold.ASP.Fitness: 34.6182



Scaffold: 11
F2072-0091

80

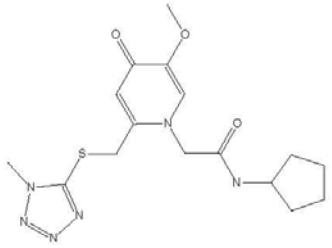
Gold.ASP.Fitness: 34.6025



Scaffold: 13
F5791-3351

81

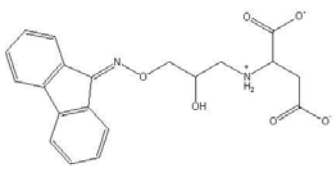
Gold.ASP.Fitness: 34.5668



Scaffold: 3
F2848-0052

82

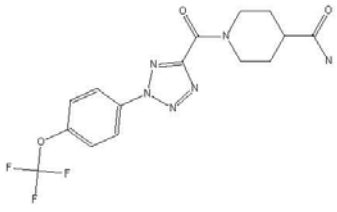
Gold.ASP.Fitness: 34.5291



Scaffold: 22
F1746-0008

83

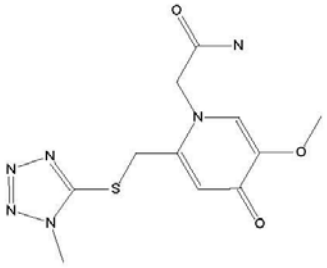
Gold.ASP.Fitness: 34.5090



Scaffold: 1
F6237-3555

84

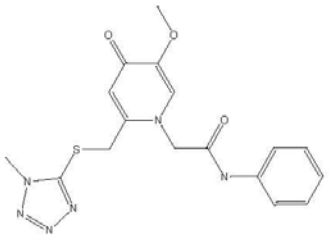
Gold.ASP.Fitness: 34.4868



Scaffold: 17
F2848-0045

85

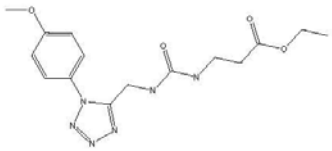
Gold.ASP.Fitness: 34.3733



Scaffold: 17
F2848-0006

86

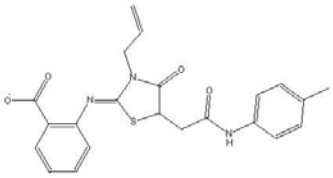
Gold.ASP.Fitness: 34.3727



Scaffold: 11
F2072-0056

87

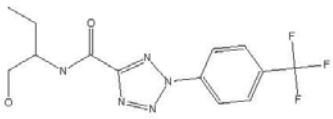
Gold.ASP.Fitness: 34.3272



Scaffold: 16
F1293-0045

88

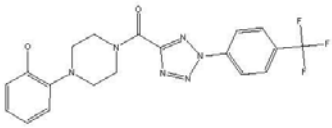
Gold.ASP.Fitness: 34.1854



Scaffold: 1
F6237-2818

89

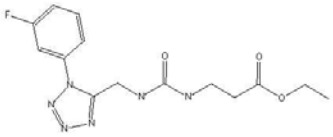
Gold.ASP.Fitness: 34.1519



Scaffold: 1
F6237-3061

90

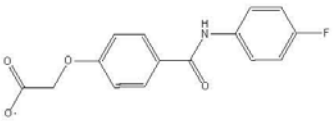
Gold.ASP.Fitness: 34.1402



Scaffold: 11
F2072-0196

91

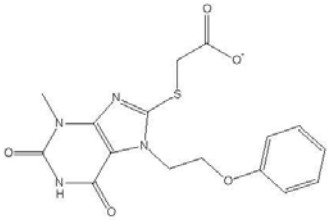
Gold.ASP.Fitness: 34.1032



Scaffold: 23
F3099-7179

92

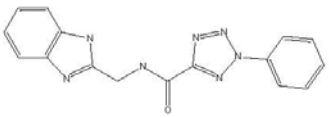
Gold.ASP.Fitness: 34.0915



Scaffold: 9
F3260-0944

93

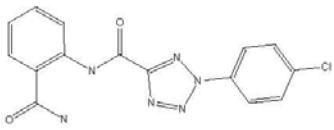
Gold.ASP.Fitness: 34.0851



Scaffold: 20
F6237-0009

94

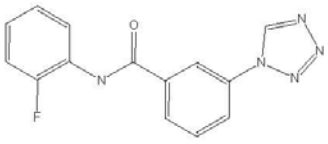
Gold.ASP.Fitness: 34.0613



Scaffold: 20
F6237-2149

95

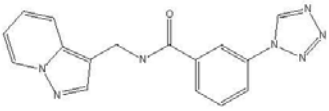
Gold.ASP.Fitness: 33.9356



Scaffold: 2
F3296-0413

96

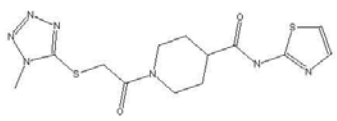
Gold.ASP.Fitness: 33.7861



Scaffold: 2
F6230-0002

97

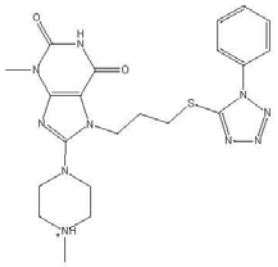
Gold.ASP.Fitness: 33.7643



Scaffold: 19
F5871-1722

98

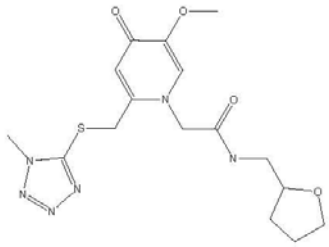
Gold.ASP.Fitness: 33.7196



Scaffold: 9
F1889-0416

99

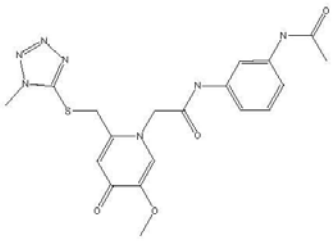
Gold.ASP.Fitness: 33.6541



Scaffold: 17
F2848-0029

100

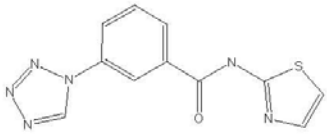
Gold.ASP.Fitness: 33.5930



Scaffold: 3
F2848-0076

101

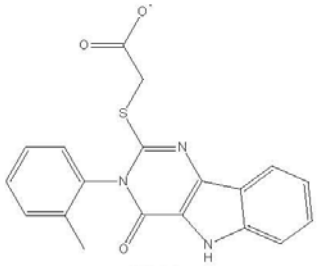
Gold.ASP.Fitness: 33.5717



Scaffold: 13
F3296-0271

102

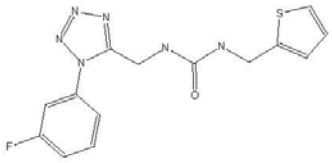
Gold.ASP.Fitness: 33.5357



Scaffold: 24
F1882-0126

103

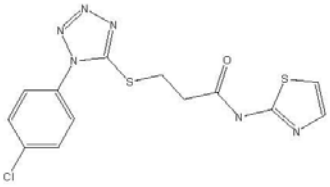
Gold.ASP.Fitness: 33.5333



Scaffold: 11
F2072-0206

104

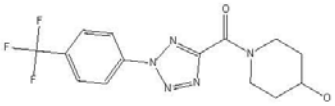
Gold.ASP.Fitness: 33.4962



Scaffold: 10
F3375-2057

105

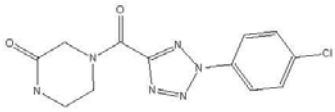
Gold.ASP.Fitness: 33.4578



Scaffold: 1
F6237-2773

106

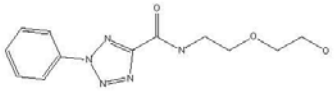
Gold.ASP.Fitness: 33.4428



Scaffold: 20
F6237-2276

107

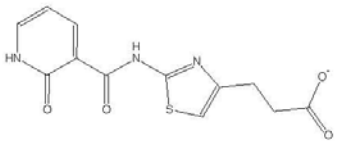
Gold.ASP.Fitness: 33.4389



Scaffold: 20
F6237-0098

108

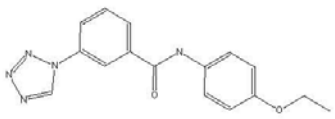
Gold.ASP.Fitness: 33.4385



Scaffold: 25
F2158-0145

109

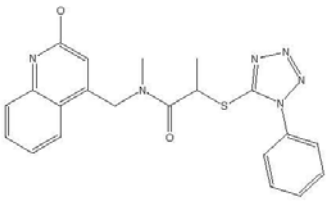
Gold.ASP.Fitness: 33.3535



Scaffold: 13
F3296-0341

110

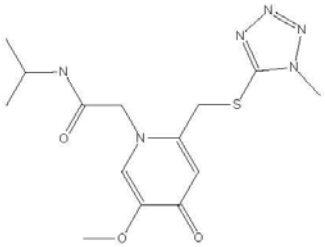
Gold.ASP.Fitness: 33.2453



Scaffold: 10
F3375-0529

111

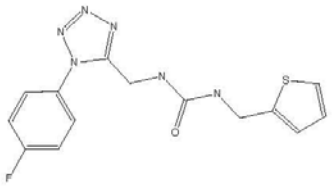
Gold.ASP.Fitness: 33.2403



Scaffold: 17
F2848-0069

112

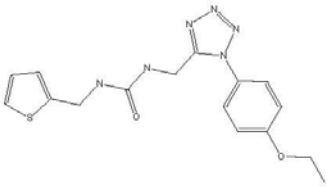
Gold.ASP.Fitness: 33.1301



Scaffold: 11
F2072-0381

113

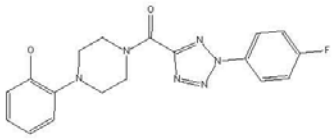
Gold.ASP.Fitness: 33.0220



Scaffold: 11
F2072-0101

114

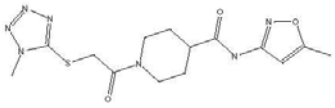
Gold.ASP.Fitness: 32.9523



Scaffold: 20
F6237-1894

115

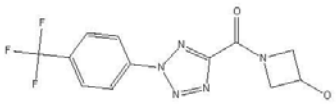
Gold.ASP.Fitness: 32.8436



Scaffold: 19
F5871-1612

116

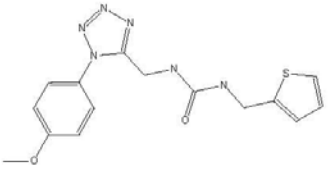
Gold.ASP.Fitness: 32.8065



Scaffold: 1
F6237-3108

117

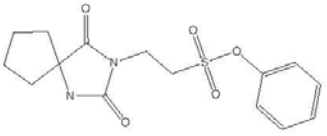
Gold.ASP.Fitness: 32.7746



Scaffold: 11
F2072-0066

118

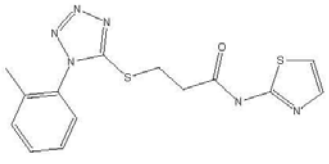
Gold.ASP.Fitness: 32.7583



Scaffold: 21
F0307-0689

119

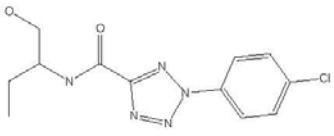
Gold.ASP.Fitness: 32.6636



Scaffold: 10
F3375-2049

120

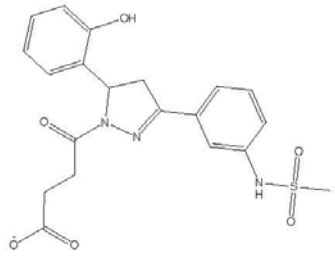
Gold.ASP.Fitness: 32.6076



Scaffold: 20
F6237-2040

121

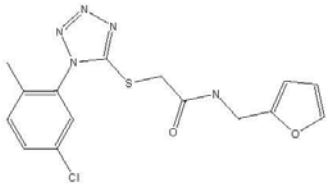
Gold.ASP.Fitness: 32.5776



Scaffold: 26
F3258-0176

122

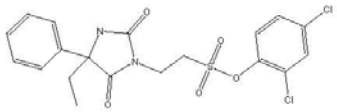
Gold.ASP.Fitness: 32.5395



Scaffold: 10
F3040-4065

123

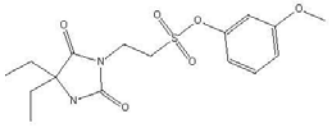
Gold.ASP.Fitness: 32.5250



Scaffold: 21
F1678-0303

124

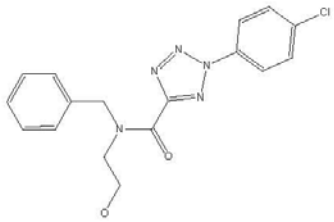
Gold.ASP.Fitness: 32.4904



Scaffold: 21
F1678-0312

125

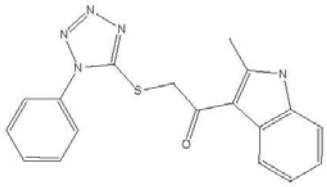
Gold.ASP.Fitness: 32.4894



Scaffold: 20
F6237-1982

126

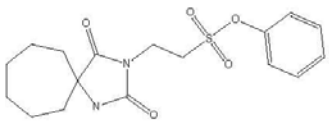
Gold.ASP.Fitness: 32.3847



Scaffold: 10
F0307-0337

127

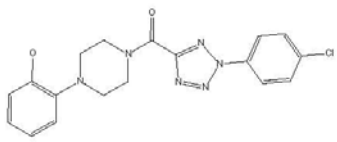
Gold.ASP.Fitness: 32.3403



Scaffold: 21
F1408-0302

128

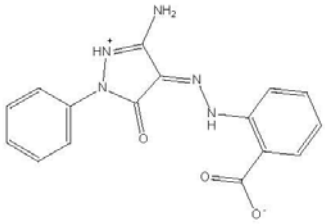
Gold.ASP.Fitness: 32.2381



Scaffold: 20
F6237-2283

129

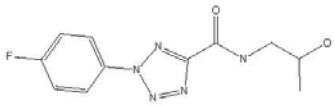
Gold.ASP.Fitness: 32.2137



Scaffold: 27
F1143-3807

130

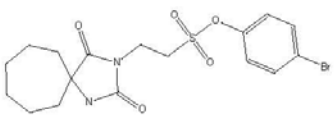
Gold.ASP.Fitness: 32.1653



Scaffold: 20
F6237-1650

131

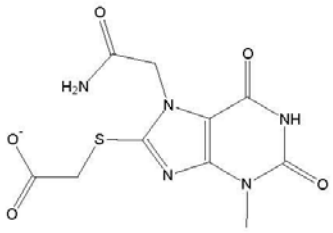
Gold.ASP.Fitness: 32.1568



Scaffold: 21
F1408-0277

132

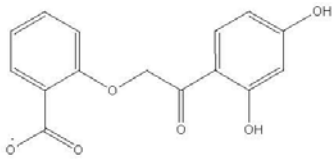
Gold.ASP.Fitness: 31.8650



Scaffold: 28
F0017-0905

133

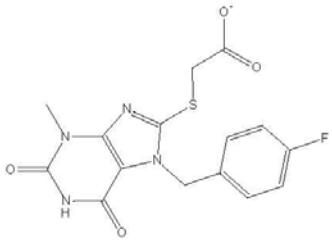
Gold.ASP.Fitness: 31.7599



Scaffold: 29
F3139-0699

134

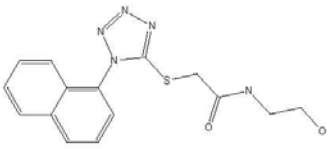
Gold.ASP.Fitness: 31.7387



Scaffold: 28
F3260-0598

135

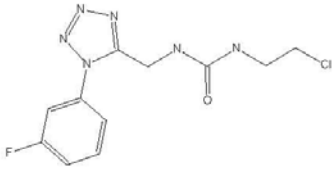
Gold.ASP.Fitness: 31.7030



Scaffold: 30
F3040-4612

136

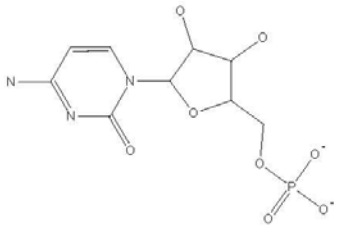
Gold.ASP.Fitness: 31.6912



Scaffold: 11
F2072-0197

137

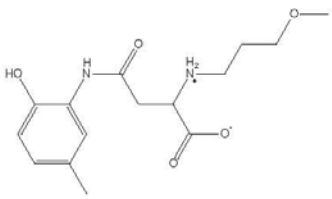
Gold.ASP.Fitness: 31.4142



Scaffold: 31
F1443-3170

138

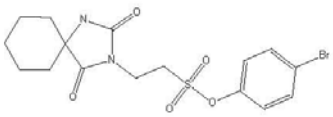
Gold.ASP.Fitness: 31.3755



Scaffold: 32
F2721-0721

139

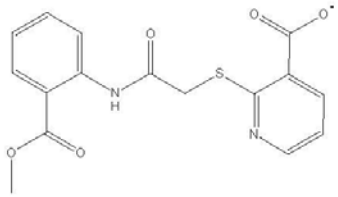
Gold.ASP.Fitness: 31.3522



Scaffold: 21
F0013-0613

140

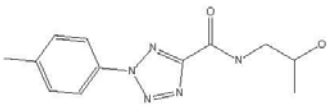
Gold.ASP.Fitness: 31.3498



Scaffold: 29
F3183-0002

141

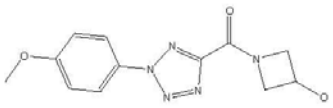
Gold.ASP.Fitness: 31.3022



Scaffold: 20
F6237-0483

142

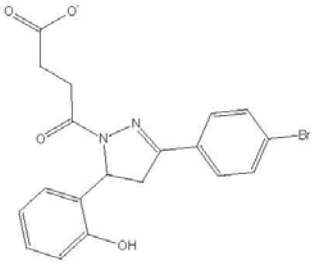
Gold.ASP.Fitness: 31.2493



Scaffold: 20
F6237-1163

143

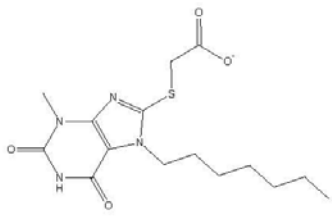
Gold.ASP.Fitness: 31.2268



Scaffold: 33
F3074-0056

144

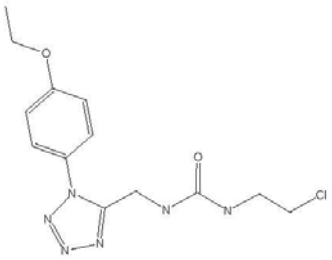
Gold.ASP.Fitness: 31.2192



Scaffold: 28
F0373-0482

145

Gold.ASP.Fitness: 31.1855

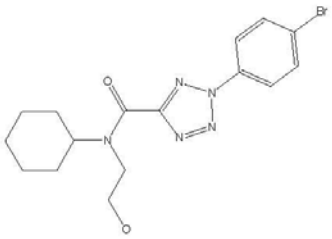


Scaffold: 11

F2072-0092

146

Gold.ASP.Fitness: 31.1206

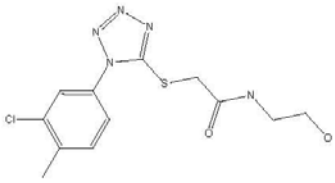


Scaffold: 20

F6237-3967

147

Gold.ASP.Fitness: 30.9649

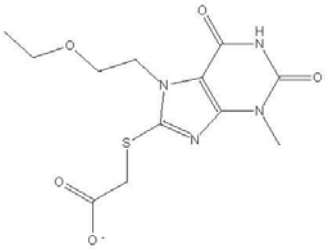


Scaffold: 34

F3040-4614

148

Gold.ASP.Fitness: 30.8907

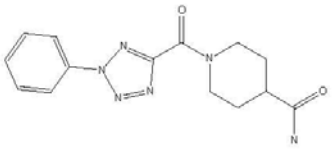


Scaffold: 28

F0372-0663

149

Gold.ASP.Fitness: 30.7370

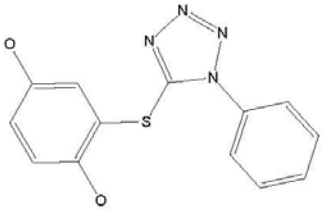


Scaffold: 20

F6237-0054

150

Gold.ASP.Fitness: 30.7051

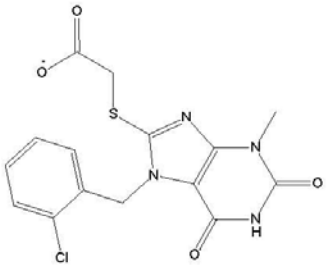


Scaffold: 34

F0405-0206

151

Gold.ASP.Fitness: 30.4439

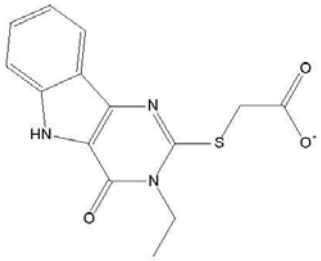


Scaffold: 28

F0372-0256

152

Gold.ASP.Fitness: 30.4433

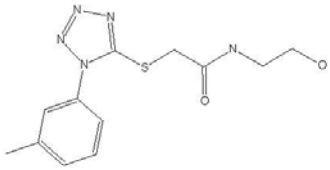


Scaffold: 35

F1882-1334

153

Gold.ASP.Fitness: 30.4147

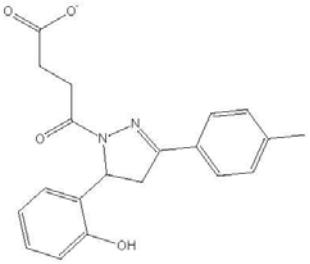


Scaffold: 10

F3040-4598

154

Gold.ASP.Fitness: 30.4083

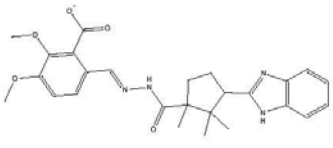


Scaffold: 36

F3074-0054

155

Gold.ASP.Fitness: 30.3200

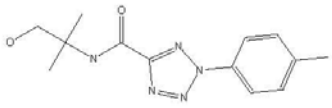


Scaffold: 37

F1284-0731

156

Gold.ASP.Fitness: 30.2892

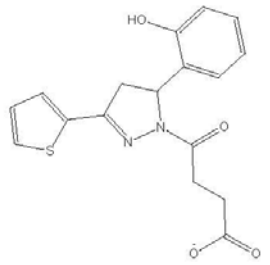


Scaffold: 20

F6237-0486

157

Gold.ASP.Fitness: 30.2280

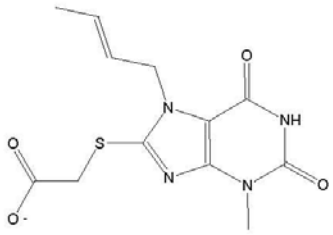


Scaffold: 36

F3074-0045

158

Gold.ASP.Fitness: 30.1692

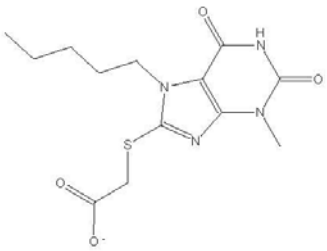


Scaffold: 28

F0372-0583

159

Gold.ASP.Fitness: 29.9391

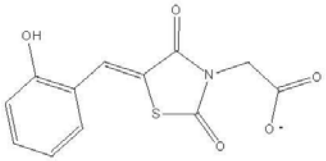


Scaffold: 28

F0017-0898

160

Gold.ASP.Fitness: 29.8324

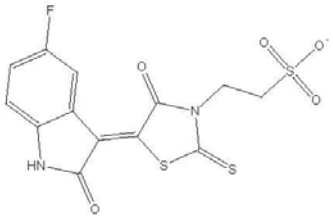


Scaffold: 16

F0804-0013

161

Gold.ASP.Fitness: 29.8241

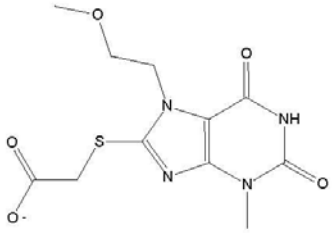


Scaffold: 16

F1691-2810

162

Gold.ASP.Fitness: 29.8020

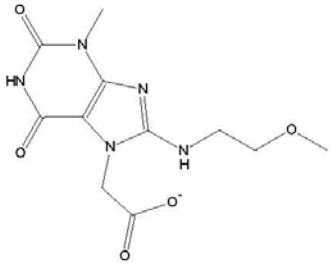


Scaffold: 28

F1191-0141

163

Gold.ASP.Fitness: 29.7872

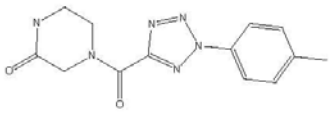


Scaffold: 28

F5038-0090

164

Gold.ASP.Fitness: 29.7537

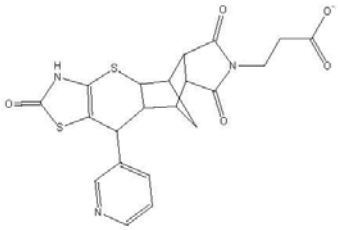


Scaffold: 20

F6237-0720

165

Gold.ASP.Fitness: 29.6680

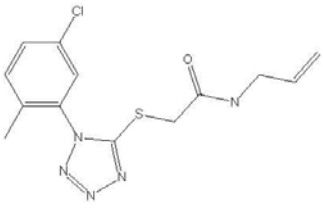


Scaffold: 38

F1734-0046

166

Gold.ASP.Fitness: 29.6262

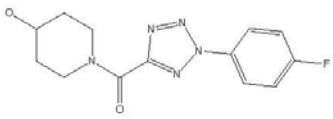


Scaffold: 34

F3040-4505

167

Gold.ASP.Fitness: 29.5403

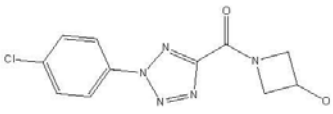


Scaffold: 20

F6237-1606

168

Gold.ASP.Fitness: 29.5026

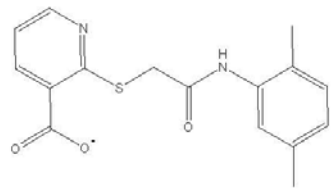


Scaffold: 20

F6237-2330

169

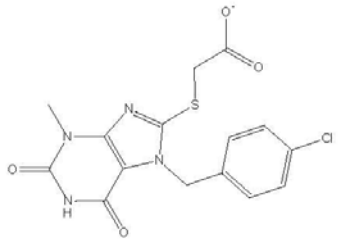
Gold.ASP.Fitness: 29.4525



Scaffold: 29
F3328-0015

170

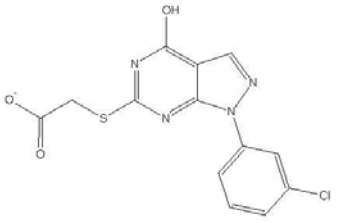
Gold.ASP.Fitness: 29.4329



Scaffold: 28
F0373-0880

171

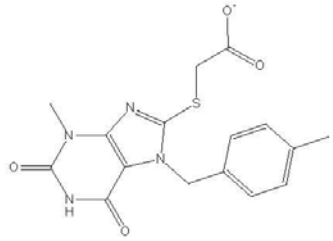
Gold.ASP.Fitness: 29.1328



Scaffold: 39
F2330-0261

172

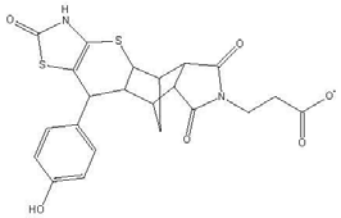
Gold.ASP.Fitness: 29.0638



Scaffold: 28
F0373-1252

173

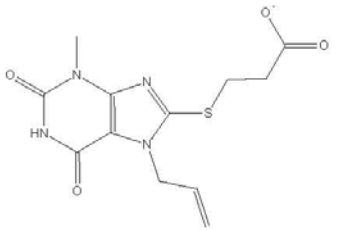
Gold.ASP.Fitness: 28.8861



Scaffold: 38
F1734-0036

174

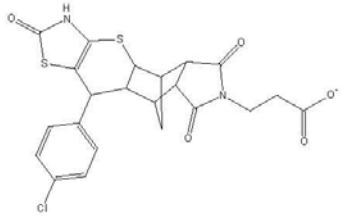
Gold.ASP.Fitness: 28.5287



Scaffold: 28
F0381-3656

175

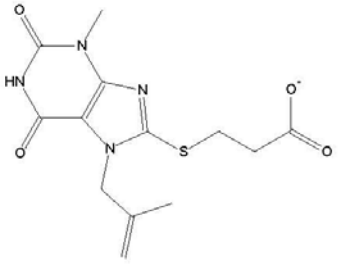
Gold.ASP.Fitness: 28.4823



Scaffold: 38
F1734-0065

176

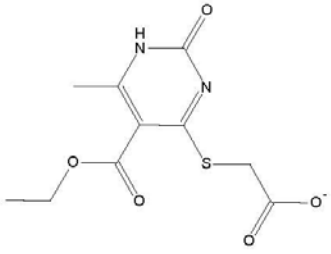
Gold.ASP.Fitness: 27.6953



Scaffold: 28
F0404-0034

177

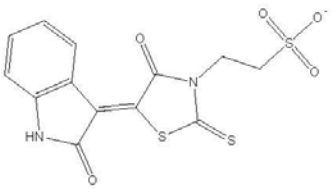
Gold.ASP.Fitness: 27.5820



Scaffold: 39
F2800-0009

178

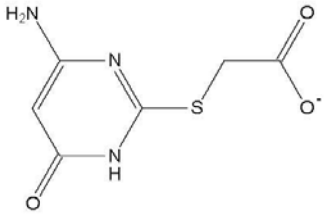
Gold.ASP.Fitness: 27.4834



Scaffold: 16
F1691-2672

179

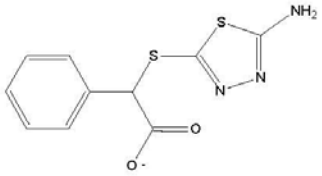
Gold.ASP.Fitness: 27.1636



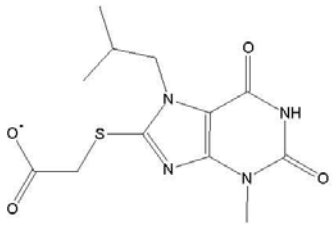
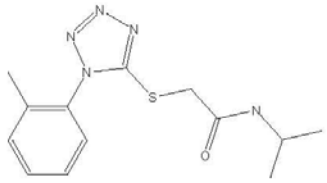
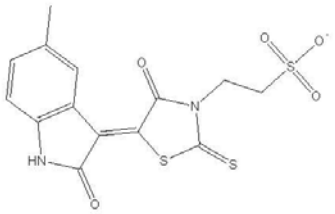
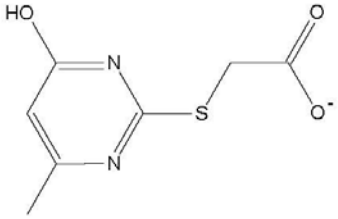
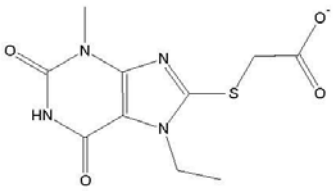
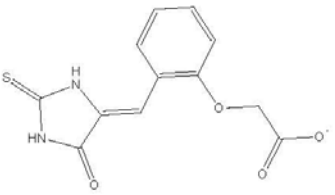
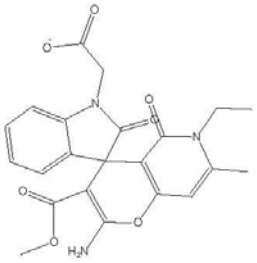
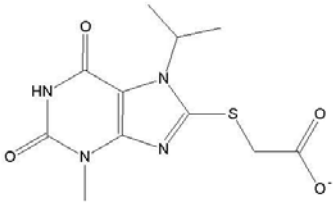
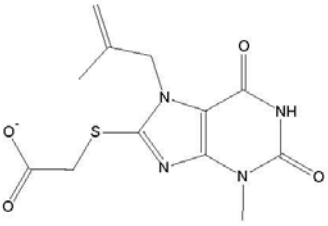
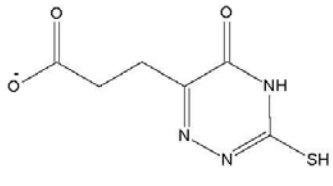
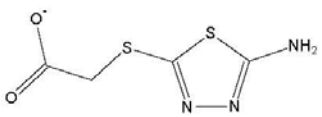
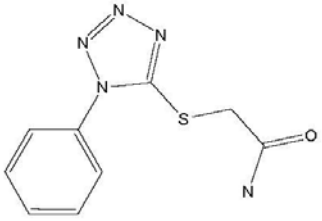
Scaffold: 39
F0918-1010

180

Gold.ASP.Fitness: 26.7757

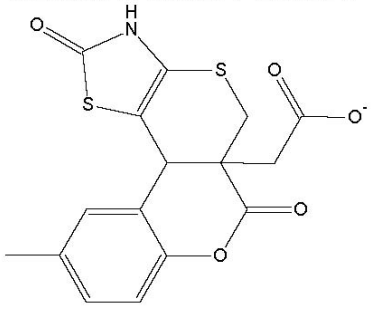


Scaffold: 39
F3328-0053

181 **Gold.ASP.Fitness:** 26.6047  Scaffold: 28 F0017-0897	182 **Gold.ASP.Fitness:** 26.5265  Scaffold: 34 F3040-4069	183 **Gold.ASP.Fitness:** 26.4794  Scaffold: 16 F3003-1043
184 **Gold.ASP.Fitness:** 26.4104  Scaffold: 40 F1019-0015	185 **Gold.ASP.Fitness:** 26.2357  Scaffold: 28 F0373-0090	186 **Gold.ASP.Fitness:** 26.2317  Scaffold: 16 F1204-0078
187 **Gold.ASP.Fitness:** 26.1206  Scaffold: 41 F2677-0087	188 **Gold.ASP.Fitness:** 26.0563  Scaffold: 28 F0373-0094	189 **Gold.ASP.Fitness:** 25.9584  Scaffold: 28 F0404-0033
190 **Gold.ASP.Fitness:** 25.6271  Scaffold: 40 F2111-0012	191 **Gold.ASP.Fitness:** 25.1257  Scaffold: 42 F2124-0410	192 **Gold.ASP.Fitness:** 24.2585  Scaffold: 34 F0212-0072

193

Gold.ASP.Fitness: 21.5130



Scaffold: 41

F1726-0167

B. Abstract

Obesity is a major health problem in many countries of the world. Weight gain is associated with several diseases. It is linked to cardiovascular disease, diabetes mellitus, osteoarthritis, some forms of cancer and sleep-breathing disorders. The common treatment option nowadays is lifestyle change including diet with calorie restriction and physical exercise. However, the treatment often remains without success. Bariatric surgery and some pharmacological treatments also exist. They lead to moderate weight loss but also possess a lot of side effect and compliance problems. Obese individuals have an increased adipocyte size as well as adipocyte number. However, the treatment options listed above just interact with adipocyte size not number.

According to recent studies, Ribonuclease A, an enzyme of the ribonuclease family, is associated to obesity. The enzyme is expressed in adipose tissue and findings show that inhibition of RNase A leads to decreased adipocyte number. Because of this, Ribonuclease A is an interesting novel target for drug development for treatment of obesity.

The aim of this study is to identify new ligands that can act as antagonists for Ribonuclease A by structure-based in silico screening. First, a structure-based pharmacophore model has been created and used for screening of the LifeChemicals database. The obtained hit compounds were further analyzed using the ChemGPS-NP web service. For final priorisation of the hit list, also docking studies were performed. As a result, 51 compounds were identified, which are currently tested for Ribonuclease A inhibition activity.

C. Zusammenfassung

Adipositas oder Fettsucht wird mit zahlreichen Krankheiten assoziiert, wie zum Beispiel Herz-Kreislauferkrankungen, Diabetes Mellitus Typ 2, Arthritis, einigen Krebsarten und Bluthochdruck. Die häufigste Behandlungsmethode ist heutzutage die Änderung der Lebensgewohnheiten, durch eine kalorienreduzierte Diät und sportliche Betätigung. Jedoch ist der Erfolg solch einer Therapie oft unzureichend. Magenverkleinerung und einige pharmakologischen Behandlungen führen zu einem gewissen Gewichtsverlust, sind jedoch auch mit vielen Nebenwirkungen verbunden. Übergewichtige Patienten weisen meistens sowohl eine Vergrößerung der Adipozyten auf, als auch eine Erhöhung der Anzahl der Zellen. Die bereits erwähnten Behandlungsmethoden bewirken jedoch nur eine Reduktion der Adipozytengroesse, aber nicht deren Anzahl.

Ribonuclease A, ein Enzym das zu der Ribonuclease A Überfamilie gehört, wird mit Adipositas assoziiert. Das Enzym wird im Fettgewebe expremiert. Neuen Studien zufolge führt eine Hemmung der RNase A zu einer Reduktion der Adipozytenanzahl. Aus diesem Grund bietet Ribonuclease A einen neuen Angriffspunkt zur der Entwicklung von Arzneistoffen zur Behandlung von Adipositas.

Das Ziel dieser Arbeit war es, mit Hilfe von strukturbasiertem in silico Screening mögliche Inhibitoren der Ribonuclease A zu identifizieren. Nach der Erstellung eines strukturbasierten Pharmakophormodells wurde die LifeChemicals Datenbank gescreent. Die dadurch gefundenen Verbindungen wurden mit ChemGPS-NP analysiert. Zur finalen Priorisierung der Hitliste wurden Docking Methoden angewandt. Als Ergebnis wurden 51 Verbindungen identifiziert, die zur Zeit auf ihre Ribonuclease Inhibierungsaktivität getestet werden.

D. List of Abbreviations

BMI	Body mass index
CADD	Computer Aided Drug Design
ChEMBLdb	Chemical European Molecular Biology Laboratory database
EAR	Eosinophil-associated ribonclease
EBI	European Bioinformatics Institute
IL-6	Interleukin-6
IUPAC	Principal component analysis
PCA	
PDB	Protein Data Base
R&D	Research and development
RNAPubmed	Ribonucleic acid
RNase	Ribonuclease
SMILES	Simplified Molecular Input Line Entry Specification
T3P	Thymidine-3-monophosphate
TNF	Tumor necrosis factor
WHO	World Health Oragnisation

E. List of Figures

Figure 1. 3D Structure of RNase A in solution by nuclear magnetic resonance, reproduced from Santoro et al. 1993

Figure 2. Sequence diagram of bovine pancreatic ribonuclease A, reproduced from Marashall et al. 2007

Figure 3. Phylogenetic tree of human RNase A ribonucleases, reproduced from Dyer, Rosenberg 2006

Figure 4. Leptin influence on food intake, reproduced from Friedmann 2009

Figure 5. Obesity is characterized by larger adipocyte size as well as by larger adipocyte number, reproduced from Arner, Spalding 2010

Figure 6. Metabolic syndrome criteria, reproduced from Grundy et al. 2004

Figure 7. Mechanisms leading to cardiovascular diseases, reproduced from Van Gaal et al. 2006

Figure 8. Disability-adjusted life-years (DALYs) lost as a result of obesity, reproduced from Haslan, James 2005

Figure 9. Drugs with indication for obesity, approved in the US, reproduced from Yanovsi et al. 2013.

Figure 10. Drug development process: from drug discovery to the market, reproduced from Gao et al. 2010.

Figure 11. Framework of pharmacophore modeling and applications, reproduced from Yang, 2010.

Figure 12. Workflow for pharmacophore based virtual screening, reproduced from Gao et al 2010

Figure 13. ChemGPS-NP descriptors, reproduced from Larsson et al. 2007

Figure 14. Interactions between the different elements of ChemGPS-NPWeb, reproduced from Rosen et al. 2009

Figure 15. Schema of complex formation

Figure 26. ClustalW2 web input form, taken from www.ebi.ac.uk/Tools/msa/clustalw2/

Figure 17. Ligand flexibility settings

Figure 18. BLAST in Uniprot results

Figure 19. Sequence alignment results with ClustalW2

Figure 20. Overview of the complex of Ribonuclease A and Thymidine-3-monophosphate.

Figure 21. Isolated binding site of Ribonuclease A in complex with Thymidine-3-monophosphate. The surface of the binding site is coloured as following: green represents the lipophilic features, pink the hydrophilic.

Figure 22. Ligand interactions shown in 2D plot, T3P in binding site of RNase A

Figure 23. Illustration of the pharmacophore model.

Figure 24. ChemGPS-NP mapping (three first dimensions).

Figure 25. Redocking results, scored with Goldscore. Original binding pose is colored blue.

Figure 26. Redocking results, scored with Chemscore. Original binding pose is colored blue.

Figure 27. Redocking results, scored with ChemPLP. Original binding pose is colored blue.

Figure 28. Redocking results, scored with ASP. Original binding pose is colored blue.

Table 1: Inhibitors and their SMILES

Table 2: ChemGPS Results: Coordinates of PCA1 – PCA3

Ich habe mich bemüht, sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

F. Curriculum Vitae

Katsiaryna Bulyha

Ausbildung

2008-2013	**Diplomstudium Pharmazie, Universität Wien** Abschluss Mag. Pharm.
06/2008	Matura mit ausgezeichnetem Erfolg → Schwerpunktmodul „Kommunikation, Präsentation, Rhetorik“
2005-2008	Wasagasse Gymnasium (BG IX), Wien, Österreich
2003-2005	Gymnasium Gonsenheim, Mainz, Deutschland
2001-2003	Gymnasium Zitadelle, Jülich, Deutschland

Arbeitserfahrung

Seit 12/2013	**Clinical Research Associate (CRA)** Austrian Breast and Colorectal Study Group (ABCSG), Österreich
07/2012 – 09/2012	**Praktikum bei European Molecular Biology Laboratory (EMBL)** Structural and Computational Biology Unit, Heidelberg, Deutschland
08/2011- 09/2011	**Praktikum bei Austrian Institute of Technology (AIT)** Health and Environment Department, Tulln, Österreich
11/2008 - 05/2011	**wissenschaftliche Mitarbeiterin der Universität Wien** Institut für Erdwissenschaften, geringfügig beschäftigt

Sprachkenntnisse

Russisch	Muttersprache
Deutsch	Muttersprache
Englisch	Fließend
Französisch	Gute Kenntnisse
Italienisch	Grundkenntnisse

Aktivitäten und Interessen

- Tutor im Praktikum „Arzneimittelanalytik und Wirkstoffentwicklung“
- Teilnahme an „Summer School on Drug Design“ 09/13, Wien
- Peer Mentor (Leitung einer Gruppe Studenten des 1. Semesters)
- Laufen, Lesen, Ölmalerei

Sonstige Kenntnisse

MS Office (Word, PowerPoint, Excel)

Führerschein Klasse B (seit 2009)