IN SILICO IDENTIFICATION OF NOVEL CANCER DRUGS WITH 3D INTERACTION PROFILING

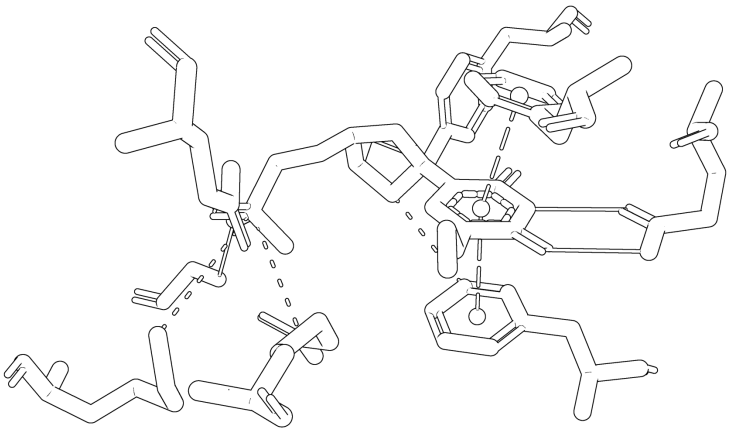
Kurzfassung der Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt an der
Technischen Universität Dresden
Fakultät Informatik

TECHNISCHE UNIVERSITÄT DRESDEN



eingereicht von

M.Sc. Sebastian Salentin
geboren am 25.11.1987 in Düren

Gutachter: Prof. Dr. Michael Schroeder (Technische Universität Dresden)
Zweitgutachter Prof. Dr. Andrew Torda (Universität Hamburg)

Tag der Verteidigung: 2. Juni 2017

Dresden, den 7. März 2017

INTRODUCTION

NEW APPROACHES FOR CANCER Cancer is one of the leading causes of death worldwide with tens of millions new annual incidents [3]. However, costly [24] and failure-prone studies [14] are major barriers to develop new drugs. Drug repositioning provides a shortcut for development by trying to find new uses for old drugs, reducing costs and shortening development time [24]. Today, repositioning already accounts for up to 30 % of new drug approvals [11] and can look back on a history of many successful repositioning cases in cancer [20].

BREAKING RESISTANCE: HSP27 AS A TARGET Repositioning works because drugs can bind to multiple unrelated proteins. Often, the sites where the drugs bind in these proteins show striking similarities [7]. One remarkable example (Figure 1) are a kinase from *Herpes* virus (green) and a heat shock protein found in cancer cells (blue).

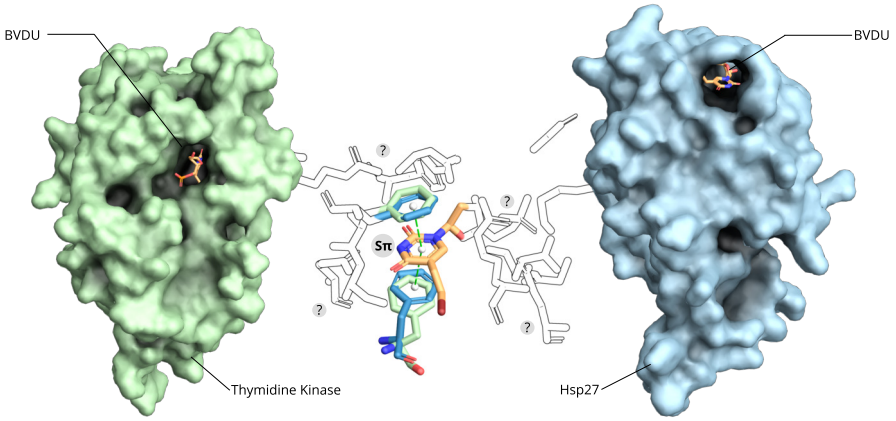


Figure 1: The Herpes drug BVDU (orange) can bind to both a kinase in *Herpes* (green) and the heat shock protein Hsp27 relevant for cancer (blue). Both proteins bind BVDU via a special motif (Sπ).

While the two proteins are vastly different, the Herpes drug BVDU (Figure 1, orange) can bind both of them and thus also works in cancer [9]. This is possible since both proteins can recognize BVDU with a special binding motif, called the sandwich π-stacking (Figure 1, Sπ colored in the center). Understanding BVDU binding in more detail is key to find more proteins with drug binding sites similar to Hsp27 and in this way also new drugs that could target Hsp27, a key player for chemoresistance in cancer [10]. Thus, the aim of the thesis is:

Aim: Finding new cancer drugs.

MORE INSIGHTS INTO DRUG BINDING Proteins can recognize and bind their ligands via specific arrangements of reversible, weak contacts [16]. Next to the so-called π-stacking shown in Figure 1, at least seven more types of contacts play an important role, next to water molecules and cofactors [17]. However, none of the currently available tools to derive interaction data from 3D structures considers all relevant aspects of binding to full extent and is suitable for high-throughput characterization of interactions. To get more insights into the binding of drugs (ligands) to proteins, an algorithm is required to translate mechanisms of binding into simple rule sets for computational detection of contacts (Open Problem I).

> **Open Problem I**
> How can protein-ligand interactions be characterized?

SYSTEMATIC SCREENING FOR CANCER DRUGS With a solution to the Open Problem I, data on interactions between drugs and proteins on a large scale opens up the possibility to compare the complex binding motifs. Finally, drugs can be found which bind in a similar fashion as BVDU and they would be candidates for targeting the cancer protein Hsp27 (see Figure 1). Existing methods largely paid attention to the protein binding site or the drugs themselves, whereas a focus on the contacts between them has only recently come into focus [16]. This special viewpoint enables to find similarities in binding characteristics between unrelated chemical classes (scaffold hopping) or unrelated proteins (target hopping), as in the case of the Herpes kinase and the Hsp27 (see Figure 1).

To this end, the complex arrangement of three-dimensional contacts between drug and protein has to be reduced to a simple representation without loss of information. Such interaction *profiles* or *fingerprints* can be stored as feature vectors. The Open Problem II addressed by the thesis is to provide such an approach which can capture key binding features and can reliably discriminate similar from dissimilar binding profiles.

> **Open Problem II**
> How can similar binding behaviour reliably detected?

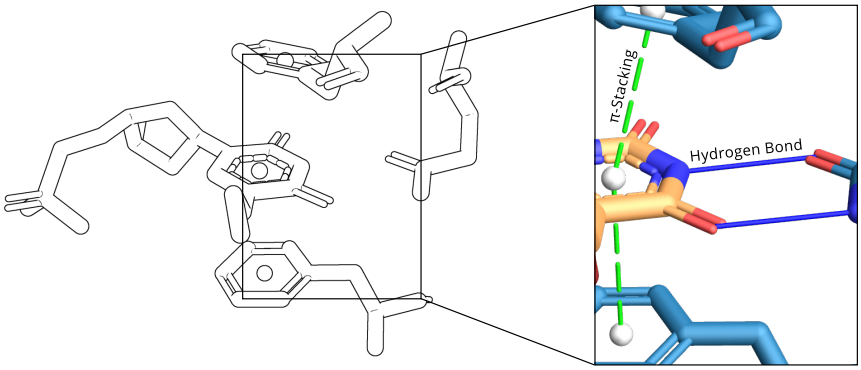
1

PLIP: THE PROTEIN-LIGAND INTERACTION PROFILER

This chapter is based on the publication `PLIP: fully automated protein-ligand interaction profiler` [17]. S. Salentin is the main contributor. He conceived and wrote the paper as well as designed and implemented the presented software PLIP.

COMPREHENSIVE CHARACTERIZATION OF INTERACTIONS The Protein-Ligand Interaction Profiler (PLIP) is a specialized algorithm to detect possible interactions between proteins and ligands, e.g. drugs, from structural data (see Figure 2). It integrates eight different interaction types and considers water molecules as well as cofactors [17].

The detection algorithm is based on rule sets for each interaction type: after automatic extraction of bound ligands and their binding environments, functional groups in both ligand and protein are characterized. Next, geometric criteria are applied to detect different contact types. Several prioritization routines reduce the number of contacts to report only the most important to the user. With this fully automatic profiling, PLIP is first in class to provide not only a comprehensive characterization of contacts without structure preparation, but also visualization, and atomic-level output files.



A Crystal structure data **B** Derived interaction patterns

Figure 2: Interactions between proteins and ligands can be derived from structural data. PLIP as a specialized tool detects interactions such as π-stacking (green) or H-Bonds (blue) with high reliability.

PLIP IS A VALIDATED TOOL PLIP was validated on a manually curated and diverse set of structural protein-ligand complexes [16]. With this set, thresholds from literature were fine-tuned for reliable detection of relevant contacts in both low- and high-quality structures. For the detection of contacts in comparison to annotations by chemical experts, PLIP reaches a recall of 97 % at a precision of 83 %.

PLIP IS FIRMLY ESTABLISHED WITHIN THE COMMUNITY Over the last years, PLIP has been made available as a web service, open source command line tool [16], and plugins for PyMOL and Chimera [21]. It is well established within the community, demonstrated not only by more than 12,000 users of the web service since launch in December 2015, but also its extensive usage in more than 20 published studies (see e.g. Biswas et al. [1], Borrel [2] and Chen, Qin and Zeng [4]) for binding characterization, most of them within a medical context.

EXPLORING AN UNTAPPED DATA SOURCE With large-scale analysis of protein-ligand contacts from the Protein Data Bank (PDB), PLIP opens up largely untapped interaction data for more insights on drug binding. For the first time, the binding preferences of small molecules were analyzed in such detail, with data on almost 400,000 contacts. One important result is that almost 50 % of all hydrogen bonds are formed to the backbone, questioning the long-standing focus on amino acid properties in current methods to measure similarity of binding environments.

HIGH PERFORMANCE FOR LARGE-SCALE ANALYSES PLIP can perform fast analysis of structures with low memory overhead and enables to process the entirety of publicly available structures ($>$ 120,000) overnight on a small cluster.

CONCLUSION PLIP as a comprehensive detection tool for large-scale protein-ligand interactions offers a solution to the Open Problem I. It enables to provide detailed insights into the binding of drugs to their targets and is therefore the prerequisite for all following analyses, including the screening for new compounds binding the cancer target Hsp27 (see chapter 2).

2

NOVEL CANCER DRUGS WITH 3D INTERACTION PROFILING

This chapter is based on `Honey as cancer treatment?` [18] (revised manuscript under review in PLOS Computational Biology). Here, S. Salentin conceived the *in silico* pipeline, wrote the paper and analyzed the results.

NEW HSP27 INHIBITORS As a main application in the thesis, interaction data from PLIP was used to screen for novel inhibitors of the heat shock protein Hsp27. To this end, 3D interaction profiling as a virtual screening workflow is presented to search for compounds with interaction preferences similar to a known binder.

CHARACTERIZING BVDU BINDING A manual evaluation of PLIP interaction data for 14 complexes with BVDU (Figure 3A) revealed that the drug can use a total of 10 small binding motifs or *subpatterns*. The key patterns are an aromatic interaction with the central ring structure of BVDU and a parallel hydrogen bond motif. Following the characterization, a *metapattern* (Figure 3B) was constructed, which describes the binding preferences of the Herpes drug.

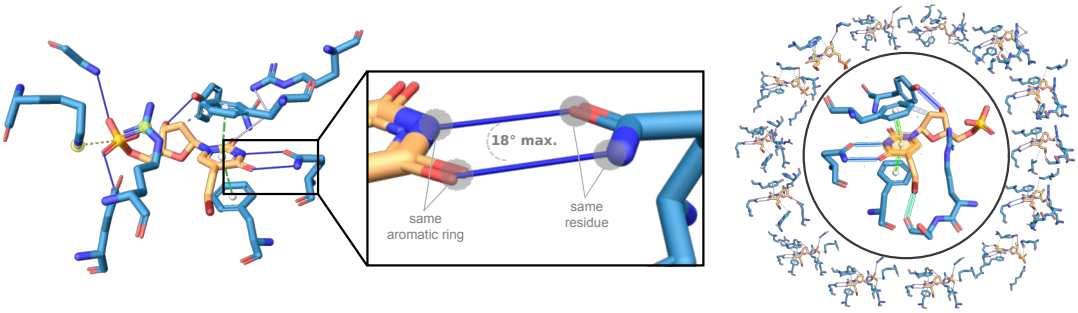


Figure 3: Characterization of BVDU binding preferences. **A** Single motifs or *subpatterns* are characterized in detail. **B** The superset of all observed subpatterns in BVDU structures yields the *metapattern*.

SCREENING WITH 3D INTERACTION PROFILING Interaction patterns in over 170,000 structural complexes from the PDB was used to search for drugs which use the same subpatterns as BVDU for binding to proteins. Figure 4 shows the distribution of the number of BVDU subpatterns in screening complexes: 58 drugs with very si-

milar interaction patterns in comparison to BVDU were found. The screening yielded a diverse set comprising > 5 chemical classes.

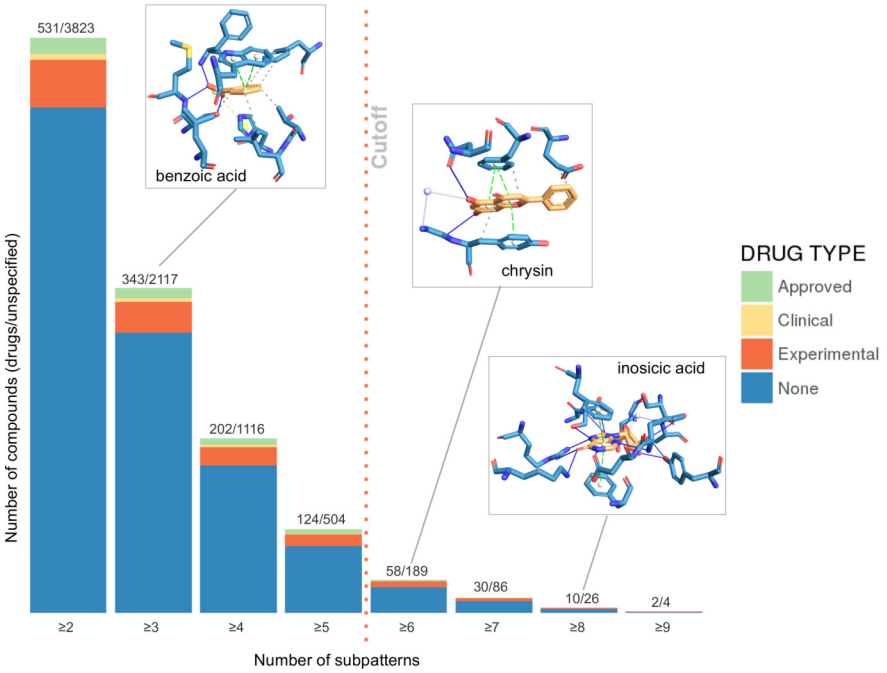


Figure 4: Distribution of the number of BVDU subpatterns in PDB complexes. Only 247 compounds in the screening set, among them 58 drugs show 6 or more of the BVDU interaction subpatterns. Boxed images show examples of screening compounds with low similarity (benzoic acid), high similarity (chrysin) and almost equal (inosicic acid) binding patterns in comparison to BVDU.

HONEY AGAINST CANCER? After docking of selected compounds against the known BVDU targets viral thymidine kinase and Hsp27, the natural flavone and honey component chrysin emerged as the top candidate, with higher predicted energies than the known binder in both targets. In literature, benefical effects of chrysin have been already shown against viral infections (see e.g. Lyu, Rhim and Park [15] and Schnitzler et al. [19]) as well as in several cancer cell lines (see e.g. Kasala et al. [12] and Zhang et al. [25]). In experiments after the screening, we could indeed demonstrate that chrysin inhibits the cancer target Hsp27 and reduces chemoresistance in multiple myeloma cells with higher potency than the known binder BVDU.

CONCLUSION 3D interaction profiling with large-scale data on protein-ligand interactions was used to predict the natural product chrysin as an inhibitor of the cancer target Hsp27. The experimental validation of this prediction demonstrates the power of the presented workflow for identifying drugs with similar binding behaviour and the potential of the method for future repositioning efforts.

3

BEYOND CANCER: MORE APPLICATIONS

This chapter contains material from `Computational Drug Repositioning by Target Hopping` [8], where S. Salentin analyzed interactions in hit targets and helped to write the paper.

REPOSITIONING FOR CHAGAS DISEASE Chagas disease is a widespread tropical disease [23] with limited treatment options [5, 22]. To identify new drug candidates, the PDB was screened with ProBiS [13] and LigandRMSD [7] for targets with binding sites similar to this of the known binder alendronate in the drug target farnesyl pyrophosphate synthase (FPPS) from the disease-causing parasite *Trypanosoma cruzi*. An interaction analysis on the top hit complex with phosphonoacetate hydrolase (PhnAH) and the antiviral drug foscarnet (Figure 5) shows that the drug binds in a similar fashion as alendronate in FPPS. Hence, the antiviral foscarnet may modulate one of the main drug targets for Chagas disease and is a top candidate for repositioning.

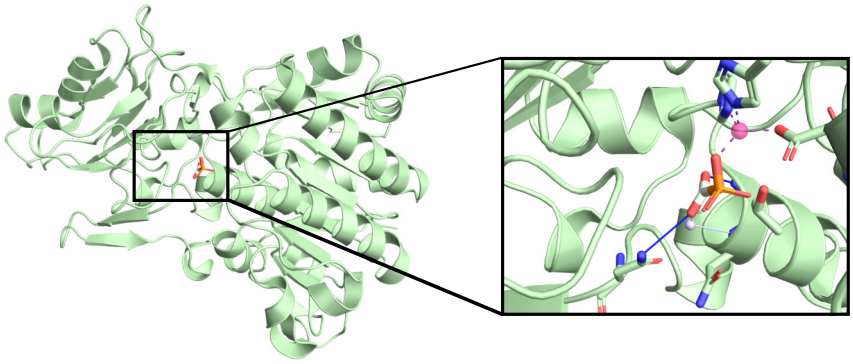


Figure 5: Characterization of interactions between PhnAH (green) and foscarnet (orange). The ligand interacts via hydrogen bonds (blue), a water bridge (light blue) and coordination of a zinc ion (pink).

EXPLOITING INTERACTION DATA Further proof-of-concept analyses demonstrate that data generated with PLIP can improve detection of false positives in docking, rapid identification of key binding features for lead optimization, and discrimination of binding modes.

CONCLUSION PLIP is a flexible tool which can be readily integrated with other similarity scoring tools and improves existing workflows by providing key insights into ligand binding.

4

TOWARDS AUTOMATED SCREENING OF THE PROTEOME

A NOVEL FINGERPRINT With the PLIP Fusion Fingerprint, an interaction fingerprint was developed, which enables to automatically encode and rapidly compare the complex three-dimensional interaction patterns observed in biological complexes (see Figure 6). It is invariant to the structures of the partaking protein or ligand and improves on previous comparable approaches especially by considering all relevant interaction types as well as water-mediated interactions.

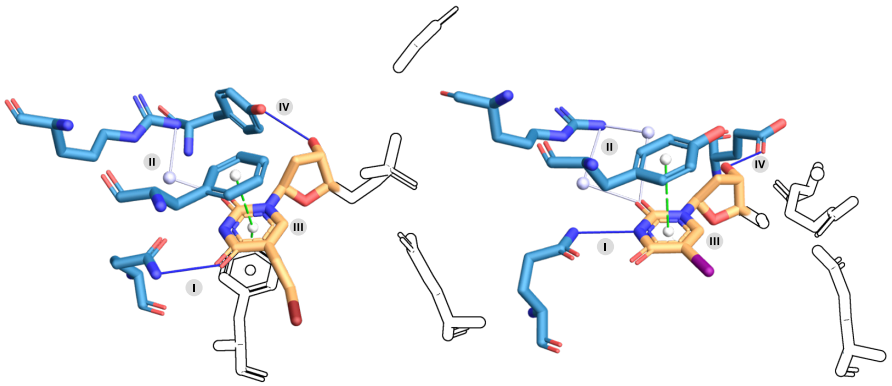


Figure 6: Binding characteristics can be shared between two-protein ligand complexes: for an automated similarity comparison, similar three-dimensional patterns (I-IV) need to be identified and encoded.

HIGH-THROUGHPUT SIMILARITY SCREENING The new fingerprint design outperformed the current state-of-the art approach TIFP [6] in discrimination of similar and dissimilar binding patterns: in a direct comparison on a set of 1800 complexes, the new fingerprint reached a F1-score of 0.877 in comparison of 0.830 with the TIFP fingerprint.

PROOF-OF-CONCEPT SCREENING: BVDU In a screening with the BVDU interaction profile, 20 % of the compounds from the 3D interaction profiling screen were recovered in the top 0.3 % ranks. Additionally, some previously undiscovered similarities were identified.

CONCLUSION The PLIP Fusion Fingerprint is a consequent development towards a fully automated screening for similar interaction patterns. The superior performance of the design was demonstrated in a benchmark and a first screening showed the potential as a standalone method or complementary to 3D interaction profiling.

BIBLIOGRAPHY

[1] A. Biswas et al. 'Role of sequence evolution and conformational dynamics in the substrate specificity and oligomerization mode of thymidylate kinases'. In: *Journal of Biomolecular Structure and Dynamics* 1102 (2016), pp. 1–19.

[2] A. Borrel. 'Development of Computational Methods to Predict Protein Pocket Druggability and Profile Ligands using Structural Data'. PhD thesis. 2016.

[3] Cancer Research UK. *Worldwide cancer statistics*. 2016.

[4] M. Chen, X. Qin and G. Zeng. 'Single-walled carbon nanotube release affects the microbial enzyme-catalyzed oxidation processes of organic pollutants and lignin model compounds in nature'. In: *Chemosphere* 163 (2016), pp. 217–226.

[5] J. R. Coura and S. L. De Castro. 'A critical review on chagas disease chemotherapy'. In: *Memorias do Instituto Oswaldo Cruz* 97.1 (2002), pp. 3–24.

[6] J. Desaphy et al. 'Encoding protein-ligand interaction patterns in fingerprints and graphs.' In: *Journal of Chemical Information and Modeling* 53.3 (2013), pp. 623–37.

[7] V. J. Haupt, S. Daminelli and M. Schroeder. 'Drug Promiscuity in PDB: Protein Binding Site Similarity Is Key.' In: *PLOS ONE* 8.6 (2013), e65894.

[8] V. J. Haupt et al. 'Computational Drug Repositioning by Target Hopping : A Use Case in Chagas Disease Drug Repositioning by Target Hopping'. In: *Current Pharmaceutical Design* 22.21 (2016), pp. 3124–3134.

[9] J. C. Heinrich et al. 'RP101 (brivudine) binds to heat shock protein HSP27 (HSPB1) and enhances survival in animals and pancreatic cancer patients.' In: *Journal of Cancer Research and Clinical Oncology* 137.9 (2011), pp. 1349–61.

[10] J. C. Heinrich et al. 'New HSP27 inhibitors efficiently downregulate resistance development in cancer cells'. In: *Oncotarget* 7.42 (2016), pp. 68156–68169.

[11] G. Jin and S. T. C. Wong. 'Toward better drug repositioning: Prioritizing and integrating existing methods into efficient pipelines'. In: *Drug Discovery Today* 19.5 (2014), pp. 637–644.

[12] E. R. Kasala et al. *Chemopreventive and therapeutic potential of chrysin in cancer: Mechanistic perspectives*. Vol. 233. 2. Elsevier Ireland Ltd, 2015, pp. 214–225.

[13] J. Konc and D. Janezic. 'ProBiS algorithm for detection of structurally similar protein binding sites by local structural alignment.' In: *Bioinformatics* 26.9 (2010), pp. 1160–8.

[14] J. Li et al. 'A survey of current trends in computational drug repositioning.' In: *Briefings in Bioinformatics* 17.1 (2015), pp. 1–11.

[15] S.-Y. Lyu, J.-Y. Rhim and W.-B. Park. 'Antiherpetic activities of flavonoids against herpes simplex virus type 1 (HSV-1) and type 2 (HSV-2) in vitro.' In: *Archives of Pharmacal Research* 28.11 (2005), pp. 1293–1301.

[16] S. Salentin et al. 'Polypharmacology rescored: Protein-ligand interaction profiles for remote binding site similarity assessment.' In: *Progress in biophysics and molecular biology* 116.2 (2014), pp. 174–186.

[17] S. Salentin et al. 'PLIP: fully automated protein-ligand interaction profiler.' In: *Nucleic acids research* 43.W1 (2015), W443–W447.

[18] S. Salentin et al. 'Honey as cancer treatment? Interaction patterns identify chrysin as suppressor of cancer chemoresistance'. In: *PLoS Computational Biology* (2017), revised manuscript under review.

[19] P. Schnitzler et al. 'Antiviral Activity and Mode of Action of Propolis Extracts and Selected Compounds'. In: *Phytotherapy Research* 24 (2010), S20–S28.

[20] J. S. Shim and J. O. Liu. 'Recent advances in drug repositioning for the discovery of new anticancer drugs'. In: *International Journal of Biological Sciences* 10.7 (2014), pp. 654–663.

[21] T. Stular et al. 'Discovery of Mycobacterium tuberculosis InhA Inhibitors by Binding Sites Comparison and Ligands Prediction'. In: *Journal of Medicinal Chemistry* 59.24 (2016), pp. 11069–11078.

[22] J. a. Urbina and R. Docampo. 'Specific chemotherapy of Chagas disease: controversies and advances'. In: *Trends in Parasitology* 19.11 (2003), pp. 495–501.

[23] WHO. *Chagas disease (American trypanosomiasis)*. 2015.

[24] E. J. Yang et al. 'Revisiting Non-Cancer Drugs for Cancer Therapy'. In: *Current Topics in Medicinal Chemistry* 16.19 (2016), pp. 2144–2155.

[25] T. Zhang et al. 'Chrysin and its phosphate ester inhibit cell proliferation and induce apoptosis in Hela cells.' In: *Bioorganic & Medicinal Chemistry* 12.23 (2004), pp. 6097–105.