



universität
wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Virtual screening for fragment-sized modulators of
Bromodomains in *Candida albicans* and *Candida auris*

verfasst von | submitted by

Erna Bajric BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Magistra pharmaciae (Mag. pharm.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on
the student record sheet:

UA 066 605

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Pharmazie

Betreut von | Supervisor:

Assoz. Prof. Mag. Dr. Johannes Kirchmair

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Johannes Kirchmair, for giving me the opportunity to work on this fascinating topic. I am immensely grateful for his guidance, insightful discussions, and constant encouragement throughout my research journey.

Furthermore, I acknowledge Assoc.-Prof. Dr. Julien Orts from the University of Vienna, for suggesting this research topic, sharing his insights on the protein targets, and his enthusiasm to experimentally evaluate the compounds identified as part of this work.

I especially want to thank Thi Ngoc Lan Vu and Vincent-Alexander Scholz for guiding me through the research, patiently helping me, and proofreading my thesis. I also extend my most sincere appreciation to the research group, COMP3D, for their unwavering support and motivation. Their collaborative spirit, willingness to help, and the warm welcome enriched this experience.

I would also like to express my sincere thanks to my former supervisor, Mag. iur. Maximilian Kunzfeld, for his caring support and valuable help throughout my time at university.

Finally, my heartfelt thanks go to my partner, Mag. iur. Pascal Steigberger, and my family for their unwavering love, encouragement, and understanding throughout this process. Their constant support was instrumental in helping me overcome challenges and persevere.

Abstract

This thesis explores the use of structure-based virtual screening techniques, specifically docking, to investigate potential modulators for the bromodomains of *Candida* species, including *Candida albicans* and *Candida auris*. The aim is to contribute to the research of innovative antifungal drugs, especially given the challenges posed by multidrug-resistant *Candida* strains. Using computational methods for subsequent experimental validation, the study systematically explores the chemical space of molecule libraries and identifies promising candidates for further investigation. The results reveal a variety of molecules identified by docking as potential ligands for the bromodomains. Similarity analyses are then used to investigate potential binding modes and structure-activity relationships, identifying key structural features that are critical for ligand recognition. Predicted interactions offer valuable insights into the mechanistic details of ligand binding, paving the way for future experimental validation, such as through nuclear magnetic resonance (NMR) spectroscopy. Overall, this study underscores the importance of docking methods in drug discovery and provides a systematic approach to identify promising candidates for new antifungal agents targeting the bromodomains of *Candida*.

Kurzfassung

In dieser Arbeit wird der Einsatz strukturbasierter virtueller Screening-Techniken, insbesondere Docking, untersucht, um potenzielle Modulatoren für die Bromodomänen von *Candida*-Arten, einschließlich *Candida albicans* und *Candida auris*, zu untersuchen. Ziel ist es, einen Beitrag zur Erforschung innovativer Antimykotika zu leisten, insbesondere angesichts der Herausforderungen, die sich durch multiresistente *Candida*-Stämme ergeben. Mit Hilfe von computergestützten Berechnungsmethoden, die anschließend experimentell validiert werden sollen, erforscht die Studie systematisch den chemischen Raum von Molekülbibliotheken und identifiziert vielversprechende Kandidaten für weitere Untersuchungen. Die Ergebnisse zeigen eine Vielzahl von Molekülen, die durch Docking als potenzielle Liganden für die Bromodomänen identifiziert wurden. Anhand von Ähnlichkeitsanalysen werden dann potenzielle Bindungsmodi und Struktur-Aktivitäts-Beziehungen untersucht und wichtige Strukturmerkmale identifiziert, die für die Ligandenerkennung entscheidend sind. Die vorhergesagten Wechselwirkungen bieten wertvolle Einblicke in die mechanistischen Details der Ligandenbindung und ebnen den Weg für eine künftige experimentelle Validierung, beispielsweise durch Kernresonanzspektroskopie (NMR). Insgesamt unterstreicht diese Studie die Bedeutung von Docking-Methoden in der Arzneimittelentdeckung und bietet einen systematischen Ansatz zur Identifizierung vielversprechender Kandidaten für neue Antimykotika, die auf die Bromodomänen von *Candida* abzielen.

Table of content

Abstract	III
Kurzfassung	V
Table of Figures	X
List of Tables	XIV
Abbreviation	XV
1 Introduction	1
2 Candida	2
2.1 <i>Candida albicans</i>	2
2.2 <i>Candida auris</i>	3
2.3 <i>Treatment of invasive fungal infections</i>	3
2.4 <i>Pathogenesis</i>	4
2.5 <i>Bromo- and extra-terminal domain proteins</i>	5
2.6 <i>Bromodomain</i>	6
2.6.1 <i>Antifungal strategy to inhibit bromodomains</i>	7
2.6.2 <i>Bromodomain protein structures</i>	7
2.6.3 <i>Protein-ligand interactions</i>	10
2.7 <i>Virtual screening</i>	12
2.7.1 <i>Structure-based virtual screening</i>	12
2.7.2 <i>Ligand-based virtual screening</i>	13
2.7.3 <i>Docking</i>	13
2.8 <i>Aim of this thesis</i>	15
3 Methods	16
3.1 <i>Protein structure and co-crystallized ligands</i>	16
3.2 <i>Databases</i>	16
3.3 <i>Schrödinger Platform</i>	17
3.4 <i>Protein preparation</i>	17
3.5 <i>Ligand preparation</i>	17

3.6	<i>Receptor grid generation</i>	18
3.7	<i>Redocking</i>	19
3.8	<i>Glide</i>	19
3.9	<i>Principal component analysis</i>	21
3.10	<i>Clustering</i>	21
3.11	<i>Hardware</i>	21
4	Results	22
4.1	<i>Database analysis</i>	22
4.2	<i>Docking results</i>	23
4.3	<i>Predicted hits</i>	23
4.3.1	Predicted hits for the target <i>Candida albicans</i> bromodomain factor 1 bromodomain 1 23	
4.3.2	Predicted hits for the target <i>Candida albicans</i> bromodomain factor 1 bromodomain 2 28	
4.3.3	Predicted hits for the target <i>Candida auris</i> bromodomain	35
4.4	<i>Chemical space analysis</i>	37
4.4.1	Chemical space analysis of MolPort compounds with molecular weight between 250 and 320 Daltons	37
4.4.2	Chemical space analysis of MolPort natura compound library.....	39
4.5	<i>Clustering</i>	42
4.6	<i>Similarity analysis</i>	43
4.6.1	Similarity analysis of the co-crystallized ligand of <i>Candida</i> bromodomain containing factor 1 bromodomain 1 and hits obtained from MolPort natural compounds.....	43
4.6.2	Similarity analysis of the co-crystallized ligand of <i>Candida</i> bromodomain containing factor 1 bromodomain 2 and hits obtained from MolPort natural compounds.....	44
4.6.3	Similarity analysis of the co-crystallized ligand of <i>Candida</i> bromodomain containing factor 1 bromodomain 1 and hits obtained from MolPort compounds with molecular weight between 250 and 320 Daltons.....	45
4.6.4	Similarity analysis of the co-crystallized ligand of <i>Candida</i> bromodomain containing factor 1 bromodomain 2 and hits obtained from MolPort compounds with molecular weight between 250 and 320 Daltons.....	46
5	Conclusions	47
6	References	49
	APPENDIX	52

Summary of all predicted hits for the target <i>Candida albicans</i> bromodomain factor 1 bromodomain 1	52
Summary of all predicted hits for the target <i>Candida albicans</i> bromodomain factor 1 bromodomain 2	54
Summary of all predicted hits for the target <i>Candida auris</i> bromodomain	57

Table of Figures

Figure 1 - PDB 5N17 showing the Bdf1 BD 1 with the co-crystallized ligand 8FK, the four helices, and the loops.	8
Figure 2 - PDB 5N18 showing the Bdf1 BD2 with the co-crystallized ligand 8HZ, the four helices, and the loops.	9
Figure 3 - Homology model structure of <i>C. auris</i> BD resolved with “Swiss model with the four helices, and the loops	10
Figure 4 - CaBdf1 BD1 (PDB 5N17) represented as cartoon showing interactions as yellow dashes with the co-crystallized ligand 8FK.	11
Figure 5 - CaBdf1 BD2 (PDB 5N18) represented as cartoon showing interactions as yellow and blue dashes with the co-crystallized ligand 8HZ.	12
Figure 6 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-015-136-309 from the subset of MolPort library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.	24
Figure 7 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-006-821-342 from the subset of MolPort library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.	25
Figure 8 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-046-848-539 from the subset of MolPort library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.	26
Figure 9 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-008-348-689 compound from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.	27
Figure 10 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-005-909-990 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.	28
Figure 11 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-019-801-991 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks.	

The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.....	29
Figure 12 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-003-997-631 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.....	30
Figure 13 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-038-415-941 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.....	31
Figure 14 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-001-026-787 compound from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.	32
Figure 15 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-000-672-474 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.....	33
Figure 16 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-001-732-545 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.....	34
Figure 17 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-001-759-343 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.....	35
Figure 18 - Protein structure C. auris represented as cartoon, with the MolPort-000-840-542 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.....	36

Figure 19 - The PCA scatter plot illustrates the chemical space covered by the subset of the MolPort database with the obtained docking hits for the target CaBdf 1 BD2, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP. 38

Figure 20 - The PCA scatter plot illustrates the chemical space covered by the subset of the MolPort database with the obtained docking hits for the target CaBdf 1 BD1, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP. 39

Figure 21 - The PCA scatter plot illustrates the chemical space covered by the MolPort database natural compounds with the obtained docking hits for the target CaBdf 1 BD2, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP. 40

Figure 22 - The PCA scatter plot illustrates the chemical space covered by the MolPort database natural compounds with the obtained docking hits for the target CaBdf 1 BD1, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP. 41

Figure 23 - The PCA scatter plot illustrates the chemical space covered by the MolPort database natural compounds with the obtained docking hits for the target C. auris, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP. 42

Figure 24 - The bar chart shows the similarity between our promising hits of the natural compound library and the co-crystallized ligand (8FK) of CaBdf1 BD1. Hits are represented as blue bars, and the average similarity as a red, dashed line... 43

Figure 25 - The bar chart shows the similarity between between our promising hits of the natural compound library and the co-crystallized ligand (8HZ) of CaBdf1 BD2. Hits are represented as blue bars, and the average similarity as red dashed line.44

Figure 26 - The bar chart shows the similarity between our promising hits of subset of MolPort library and the co-crystallized ligand (8FK) of CaBdf1 BD1. Hits are represented as blue bars, and the average similarity as red dashed line. 45

Figure 27 - The bar chart shows the similarity between our promising hits of subset of MolPort library and the co-crystallized ligand (8HZ) of CaBdf1 BD2. Hits are represented as blue bars, and the average similarity as red dashed line.....46

List of Tables

Table S1 - The most promising docking results for CaBdf 1 BD1 were identified by molecular docking with the subset of MolPort library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.	53
Table S2 - The most promising docking hits for CaBdf 1 BD1 were identified by molecular docking with the MolPort natural products library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.	54
Table S3 - The most promising docking hits for CaBdf 1 BD2 were identified by molecular docking with the subset of MolPort library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.	56
Table S4 - The most promising docking hits for CaBdf 1 BD2 were identified by molecular docking with the MolPort natural compound library. The Table contains the following information: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.	56
Table S5 - The most promising docking hits for C. auris were identified by molecular docking with the MolPort natural compound library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.	57

ChatGPT and Grammarly were used in this work to create nicer expression.

Abbreviation

2D	two-dimensional
3D	three-dimensional
BDF	bromodomain factor
BD	bromodomain
SDF	structure data file
SMILES	simplified molecular-input line-entry system
BET	bromo- and extra-terminal domain
VS	virtual screening
CVS	comma-separated values
HDAC	histone deacetylases
HAT	histone acetyltransferases
HTS	high throughput screening
SBVS	structure-based virtual screening
LBVS	ligand-based virtual screening
PCA	principal component analysis
PDB	Protein Data Bank
NMR	nuclear magnetic resonance
CaBdf	<i>Candida</i> bromodomain factor
HIV	human immunodeficiency virus

1 Introduction

Candida albicans and *Candida auris*, pose significant threats to public health due to their ability to cause a wide range of infections. Infections caused by species can manifest as oral thrush, vaginal yeast infections, or invasive candidiasis, the latter of which can lead to severe systemic infections such as candidemia and disseminated candidiasis, resulting in septic shock and organ failure. Immunocompromised individuals, including those in intensive care units, cancer patients undergoing chemotherapy, transplant recipients, and individuals with HIV (human immunodeficiency virus), are particularly vulnerable to infections. Moreover, the widespread use of broad-spectrum antibiotics, immunosuppressive therapies, and invasive medical procedures has contributed to the increasing incidence of infections, further exacerbating the healthcare burden.^{1, 2}

Candida infections are currently exhibiting a concerning rise in mortality rate, which is attributed to the growing resistance of these fungi. The emergence of multidrug-resistant *Candida* strains, including the notorious *Candida auris*, further complicates treatment options and underscores the urgent need for novel antifungal therapies.²

Researchers have intensified efforts to identify new therapeutic targets and develop innovative antifungal agents in response to the escalating threat posed by *Candida* infections. One promising target for drug development is the fungal bromodomain, an epigenetic reader module of DNA implicated in regulating gene expression and virulence in *Candida* species.^{2, 3}

Mietton et al.³ demonstrated that the fungal bromodomain is a promising target for antifungal drug discovery. Their results emphasize this by successfully targeting the fungal bromodomain rather than the human one and demonstrating a reduction in fungal growth and lethality by deletion and mutation of the bromodomains.^{2, 3} Targeting fungal bromodomain, therefore, opens a promising opportunity for drug development. However, traditional drug discovery and development approaches are very slow and require a vast financial investment, which results in a small number of potential drug candidates.⁴

Computer-aided drug design methods can help streamline the drug discovery process. They enable the virtual screening of vast compound libraries, significantly reducing the time and resources required for hit identification. Virtual screening methods save money and open the door to exploring a much larger chemical space, potentially discovering more effective and innovative drugs.⁵

In this work, we performed structure-based virtual screening of large compound libraries with Glide. The most promising virtual hits will be purchased and subjected to NMR (nuclear magnetic resonance) binding studies with the *Candida* bromodomain structures.

2 **Candida**

2.1 **Candida albicans**

Candida albicans is the leading cause of fungal infections in humans globally.^{6, 7} It is a mostly harmless pathogen that has adapted to live in a commensal way in the human body, which means that it benefits from the relationship without harming the host. The human body regulates the spread and pathogenicity of the pathogen through the immune system and its microbiome of the intestine and mucosal surfaces.^{1, 8, 9}

Candida albicans (*C. albicans*), a commensal member of the human microbiota, colonizes the gastrointestinal tract, oral cavity, and reproductive tract in healthy individuals without causing diseases.⁹ Commensal areas occupied by *C. albicans* are the oral mucosa, esophageal mucosa, gastrointestinal tract, vaginal mucosa, and nail beds. *Candida's* commensal properties, which can turn them into opportunistic pathogens, should be pointed out. Due to its adaptability, *C. albicans* is an excellent pathogen that thrives under different conditions. During an infection, *C. albicans* colonizes different host niches that differ in pH value, oxygen and CO₂ content, and nutrient availability.⁸ Moreover, this fungus can switch from yeast to hyphae, convert to white opaque, form biofilms, adhere to cells, and remodel the cell wall.¹⁰

2.2 *Candida auris*

Candida auris, a member of the *Candida* species, is a multidrug-resistant fungal pathogen and was first isolated in 2009 from the external ear canal of a patient in Japan.¹¹ It is considered one of the leading causes of nosocomial infections, precipitating candidemia, and various invasive conditions, including pericarditis, respiratory tract infections, and urinary tract infections.^{12, 13, 14}

Like *C. albicans*, *C. auris* is a member of the *Candida* species and the CTG clade. Fellows of the CTG clade translate the CTG Codon as serin rather than leucine.¹⁵ In contrast to *C. albicans*, *C. auris* is not a commensal yeast mostly found on the mucosal surface or gastrointestinal tract; it predominantly colonizes the skin.¹² Due to its affinity for human skin and abiotic surfaces, *C. auris* can persist for a prolonged period and inhabit healthcare environments and equipment despite the use of disinfectants. Transmission occurs from skin to skin or surface to skin.¹² This fungus form is mainly associated with high mortality rates and bloodstream infections, especially in immunocompromised patients.¹⁶

The risk of infection with *C. auris* is similar to that of *C. albicans*. Patients at risk primarily include older diabetics, patients who have recently undergone surgery, and patients with implanted or permanently implanted medical devices. Other risk factors for infection include broad-spectrum antibiotic therapy, antifungal therapy, and chronic kidney disease. Most *C. auris* infections are associated with diarrhea and broad-spectrum antibiotics.¹⁷

One unique feature of *C. auris* is that it is thermostable. High temperatures above 40°C do not affect its growth. Another characteristic is its ability to tolerate high salt concentrations.¹³ *C. auris* grows *in vitro* in different forms, such as oval-shaped, ellipsoidal, or elongated cells. It produces white or gray, smooth, and gleaming colonies that appear as single cells or aggregates clumped together.¹⁸

2.3 Treatment of invasive fungal infections

The standard approach to treating invasive *Candida* infections typically involves four classes of antimycotic drugs: polyenes (e.g., amphotericin B), pyrimidine analogs, echinocandins, and azoles.¹⁹ Azoles are the most frequently prescribed antimycotic

drugs. Their mechanism of action centers on the selective inhibition of lanosterol-14 α -demethylase, the critical enzyme in ergosterol synthesis. By targeting this enzyme, azoles interrupt the conversion of lanosterol to ergosterol, which leads to changes in membrane permeability and the activity of membrane-bound proteins, ultimately impeding fungal growth.²⁰

Polyenes target ergosterol in fungal cell walls. Ergosterol is only present in fungi and is essential for maintaining membrane fluidity and integrity and transmitting cell signals. By binding to ergosterols, polyenes inhibit the physiological function of the fungal membrane.²¹ Echinocandins inhibit the catalytic subunit of 1,3- β -D-glucan-synthase. This group of antifungal drugs can only be administered intravenously, making long-term treatment and therapy at home impossible.² The pyrimidine analog flucytosine is converted to 5-fluorouracil, which inhibits RNA and DNA synthesis. The antifungal activity depends on cytosine permease, cytosine deaminase, and uracil phosphoribosyl transferase.²⁰

Both *C. albicans* and *C. auris* have developed resistance to various mentioned agents through different mechanisms. The main resistance mechanisms can be categorized as reduction of therapeutic drug concentration, alteration of drug targets, and metabolic changes.¹⁹

Reduction of therapeutic drug concentration is meant, e.g., increasing drug efflux or the number of targets, as well as the ability to sequester drugs within extracellular and intracellular compartments and pro-drug conversion.²⁰ Mutation within the fungus, which decreases the affinity for drugs, is primarily responsible for altering drug targets. Metabolic changes lead to the loss of specific functions or reduced function in the metabolic process.¹⁹

2.4 Pathogenesis

The type and severity of the infection depend on the immune status and the extent to which the pathogen damages the host. Therefore, the disease caused by *Candida* is a complex outcome of these causes.⁹

On the one hand, *Candida's* pathogenesis depends on the expression of virulence factors such as germ tube formation, adhesion, phenotype switching between specific cell types, biofilm formation, and the production of hydrolytic enzymes.²²

Disruption of the immune system, the body's barrier functions, or the local microenvironment can alter the growth, gene, and protein expression, metabolism, and morphology of *C. albicans*, causing damage to the host. The severity of this damage determines whether the disease is acute or recurrent, which may be related to virulence, viability, and antifungal resistance in the environment and host.^{9, 6} Transmission of *C. albicans* occurs shortly after or at birth from mother to child, and at this point, the fungus can remain commensal or cause neonatal infections.⁹

We distinguish between two types of candidiasis for medically necessary infections caused by *C. albicans*: systemic and mucosal. Mucosal candidiasis is limited to one area, while systemic candidiasis can affect more regions and spread through the bloodstream.⁶ As a result, infections with *Candida* species are considered a severe problem in public health and represent one of the most prominent nosocomial pathogens.⁹

2.5 Bromo- and extra-terminal domain proteins

The majority of eukaryotic DNA is stored in the cell nucleus. The carriers of nuclear DNA are the chromosomes, each of which consists of a single DNA molecule with an envelope of packaging proteins. These packaging proteins are histone and non-histone proteins. The histones have many alkalic amino acids: arginine and lysine.¹ Chromatin is essential in regulating DNA accessibility for gene transcription, DNA repair, and replication.¹ It is organized into subunits, the nucleosomes, which consist of 147 pairs of nucleotides coiled in a left turn around a histone octamer. These nucleosomes each contain two copies of histone types H2A, H2B, H3, and H4, forming an octamer.²³

Three fundamental mechanisms control the regulation of specific gene expression programs: post-translational histone modifications, the incorporation of histone variants, and the precise positioning of nucleosomes.²⁴ Histone lysine acetylation, an essential post-translational modification mediated by histone acetyltransferases

(HATs) and counterbalanced by histone deacetylases (HDACs), represents a key regulatory mechanism in transcription and chromatin signaling pathways.²⁵

HAT neutralizes the positive charges of lysine; acetylating weakens the interaction between histones and DNA, loosens the densely packed chromatin, and enables gene expression.^{1, 26} In contrast, HDAC enzymes do precisely the opposite and are associated with transcription repression. Acetylated lysine residues serve as recognition sites for protein modules known as bromodomains.²⁶

2.6 Bromodomain

As crucial epigenetic reader modules, bromodomains are integral to the protein family of bromo- and extra-terminal (BET) domains. They function as chromatin-associated proteins, playing a pivotal role in epigenetic regulation. Found in diverse organisms, from humans to fungi, bromodomains, and BET proteins are critical players in the intricate processes of chromatin remodeling and gene expression regulation.²⁷

Four BET family members have been identified in humans: Brd2, Brd3, Brd4, and Brdt. Each member contains two bromodomains (BDs) responsible for binding chromatin through recognizing acetylated histones and lysine residues. Canonical BDs bind monoacetylated peptides, while BET BDs recognize diacetylated peptides due to their larger binding pocket, enabling them to accommodate them. Unlike humans or *Saccharomyces cerevisiae*, *C. albicans* only express one bromodomain.²

Research by Mietton et al.³ explored the potential of targeting the bromodomain within the BET protein family as an antifungal strategy in *Candida albicans*. Their findings suggest that bromodomains are crucial for fungal survival and virulence, as they are involved in chromatin remodeling during transcription. Furthermore, bromodomains recruit proteins involved in transcriptional regulation, including transcription factors, co-activators, co-repressors, and chromatin remodeling complexes, thereby linking histone acetylation to transcriptional control. This study demonstrated the crucial role of the bromodomain *in vitro* and *in vivo* mouse studies.³

BET proteins are found in the nucleolus of the Fungi. They specifically recognize acetylated histone tails, contributing to chromatin remodeling during transcription. Studies suggest that bromodomains are essential for fungal gene expression and their ability to cause disease.^{1, 3}

Histone acetylation, regulated by opposing enzymes known as histone acetyltransferases (HATs) and histone deacetylases (HDACs), play a crucial role in chromatin remodeling by loosening chromatin structure and facilitating access by DNA polymerase and transcription factors.²⁸

While human BET proteins share sequence homology with fungal BET proteins and target chromatin through their bromodomains, they differ sufficiently in sequence to develop drugs that potentially target fungal BET proteins.³

2.6.1 Antifungal strategy to inhibit bromodomains

C. albicans is a diploid organism. When one of the two BDs is deleted, the fungus's growth is significantly affected. Based on these experiments, the bromodomain is a promising antifungal therapeutic strategy. Mietton et al.³ investigated that deletion of Bdf1 in *C. albicans* leads to fungal death and that a mutation in the protein sequence of BDs abolishes BD-mediated ligand-binding activity, severely impairing growth. This confirms inhibition of BET BD as a potential antifungal strategy in *C. albicans*.³ They emphasized that combined inactivation of Bdf1 BD1 and BD2 and their deletion leads to death, and the single inactivation of one BD shows a reduction in growth. Notably, the inhibition of BD2 has a more significant effect on growth reduction.³

2.6.2 Bromodomain protein structures

Human bromodomains have improved as a target strategy in treating cancer and other non-infectious diseases but not in antifungal strategies.¹ Both protein regions BD1 and BD2 form four helices with the same name but differ in their primary sequence. Two of the four helices form two loops that define the protein's binding pocket. We used the primary protein sequences to search the Protein Data Bank (PDB) for matching structures. The PDB is a publicly available database that contains data on biological macromolecules. (<https://www.rcsb.org>)³⁹ In the PDB, five *C. albicans* bromodomain structures are available: 5N13, 5N15, 5N16, 5N17, and 5N18. For CaBdf1 BD1, three 3D structures can be found in the PDB as the

PDB codes 5N15, 5N16, and 5N17. Further, two 3D structures for CaBdf1 BD2 can be found in the PDB as the PDB codes 5N13 and 5N18.

We used crystallized protein structures better to understand the protein structure and its possible interactions. Fungal bromodomains have four alpha right-handed helices (Z, A, B, C), which are shown in Figures 1, 2, and 3. The well-conserved binding pocket is defined between the loops ZA and BC. The binding pocket shows hydrophobic properties and contains water molecules inside.³

Figure 1 shows CaBdf 1 BD 1 from PDB 5N17 as cartoon with the four alpha right-handed helices and two loops forming the binding pocket bound to the co-crystallized ligand 8FK. The protein binding site is located between the ZA- and BC-loop where the ligand 8FK binds to the protein due to different interactions.

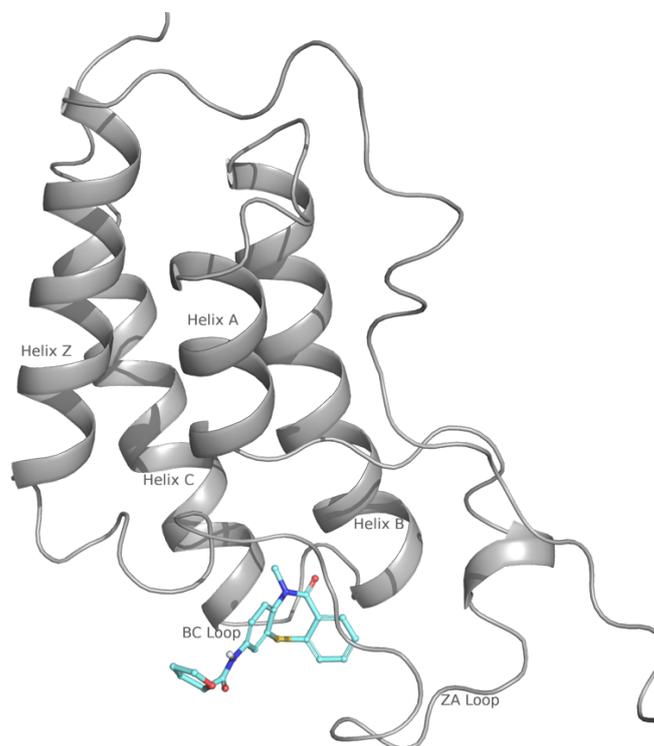


Figure 1 - PDB 5N17 showing the Bdf1 BD 1 with the co-crystallized ligand 8FK, the four helices, and the loops.

Figure 2 shows CaBdf 1 BD 2 from PDB 5N18 as cartoon with the four alpha right-handed helices and two loops forming the binding pocket bound to the co-

crystallized ligand 8HZ. The protein binding site is located between the ZA- and BC-loop where the ligand 8HZ binds to the protein due to different interactions.

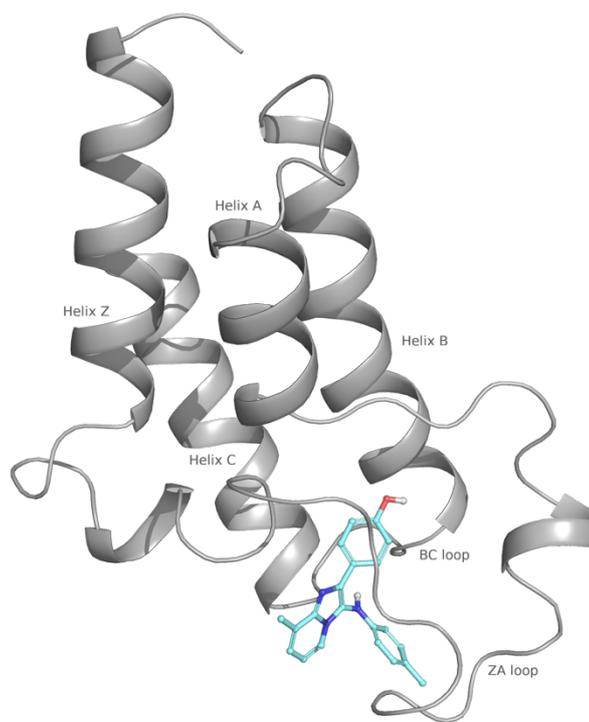


Figure 2 - PDB 5N18 showing the Bdf1 BD2 with the co-crystallized ligand 8HZ, the four helices, and the loops.

The third protein structure investigated is *C. auris* BD. Figure 3 shows the homology model of *C. auris* BD homology model by using the “Swiss model”, as no resolved 3D was available at the time of our study. Since protein folding is conserved in bromodomains, the four alpha right-handed helices (Z, A, B, C) with the two loops (ZA and BC) that define the binding pocket are also found in this case.

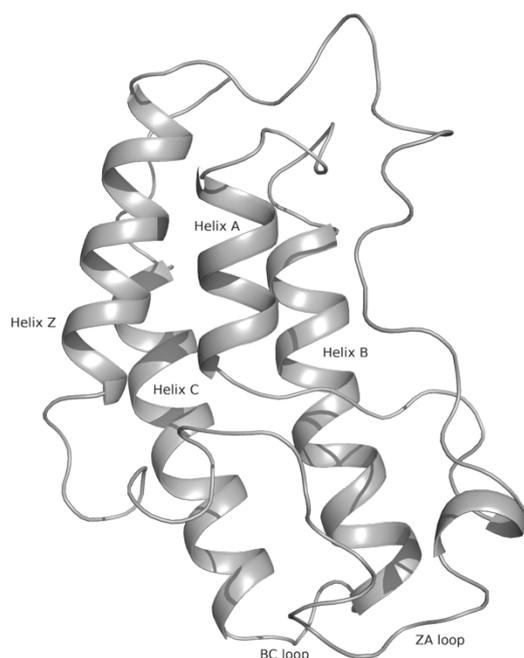


Figure 3 - Homology model structure of *C. auris* BD resolved with "Swiss model with the four helices, and the loops

2.6.3 Protein-ligand interactions

Zhou et al.²⁹ performed MD analyses to investigate the binding pocket of CaBdf1 BDs. Like mammalian BDs, the *C. albicans* BDs contain a highly conserved binding pocket within five water molecules in unbound conformation. These water molecules form hydrogen bonds with the protein main and side chains.²⁹ The authors postulated that the water molecules reduce the volume of the binding pocket, which can have a crucial role in protein-ligand identification since they can interact via hydrogen bonds with the ligand and establish the interaction and connection between the protein and its ligand. Furthermore, they showed that the stability of those water molecules varies between both domains and single molecules.²⁹

The primary interaction in CaBdf 1 BD1 of the co-crystallized ligand is the hydrogen bond with ASN in position 291 shown in Figure 4.

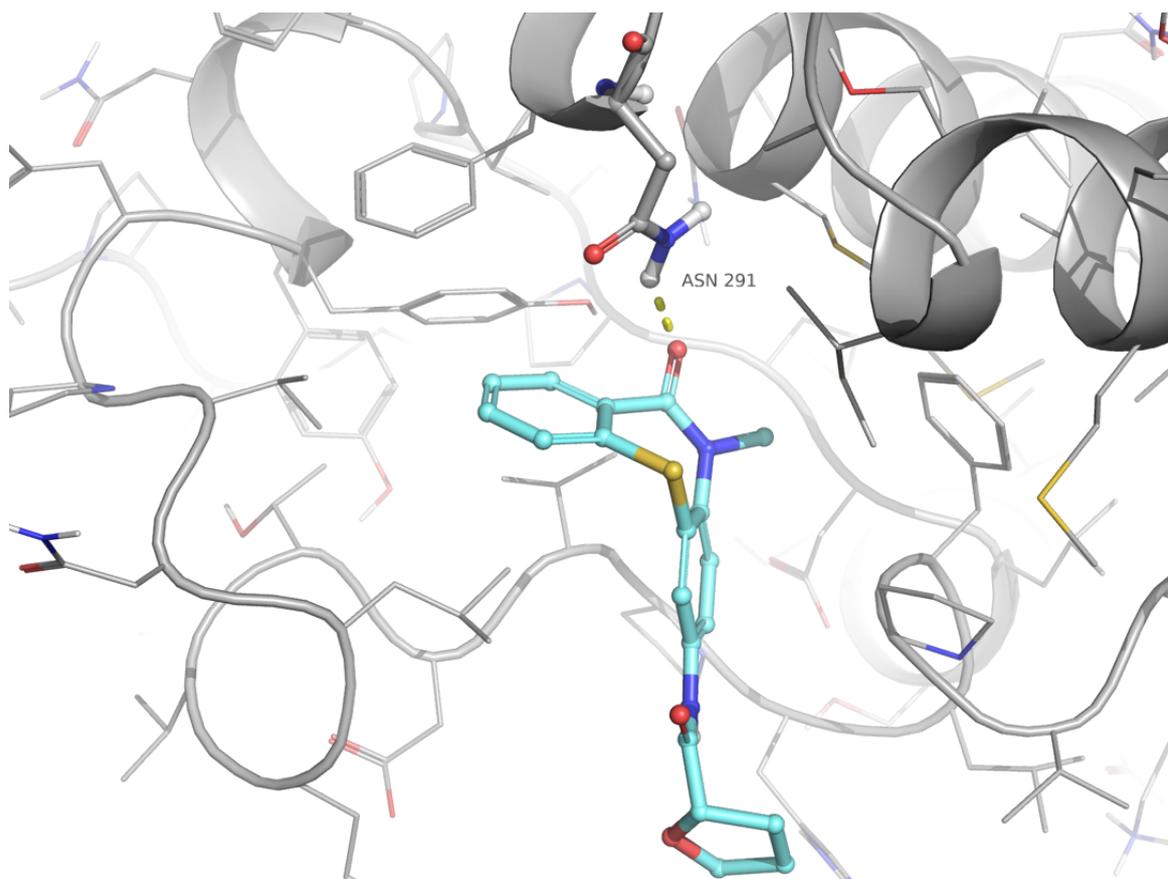


Figure 4 - CaBdf1 BD1 (PDB 5N17) represented as cartoon showing interactions as yellow dashes with the co-crystallized ligand 8FK.

The interactions within the 5N18 binding pocket, shown in Figure 5, are a hydrogen bond formed between the ligand and ASP in position 468 and a π - π -interaction between the ligand and phenyl in position 467. Those interactions are also found in the docking validation.

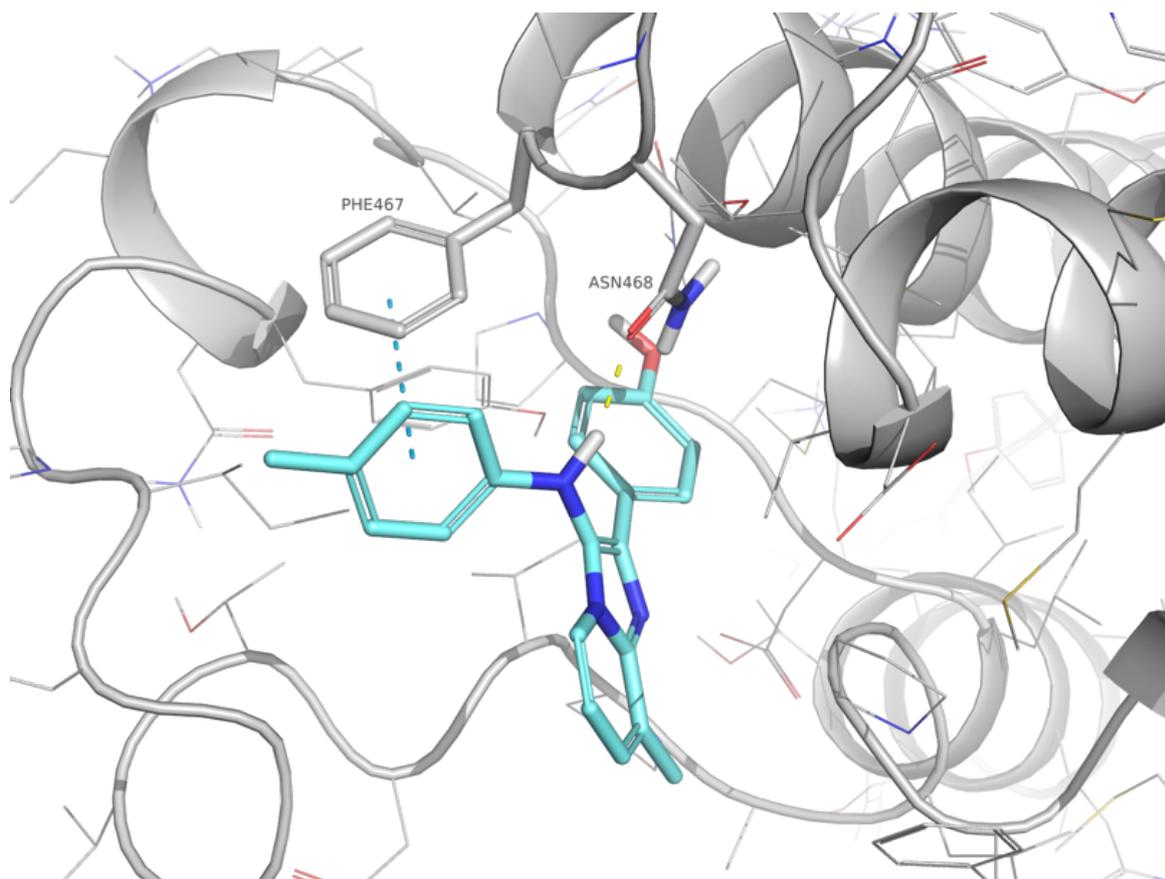


Figure 5 - CaBdf1 BD2 (PDB 5N18) represented as cartoon showing interactions as yellow and blue dashes with the co-crystallized ligand 8HZ.

2.7 Virtual screening

Virtual screening (VS) is a computational method for investigating the potential interaction between small molecules and a target molecule. Due to its cost-efficiency, virtual screening is becoming increasingly essential compared to traditional lead discovery tools.³⁰

VS is prominent for lead identification, lead optimization, and scaffold hopping.⁵

This method offers two primary approaches: ligand-based and structure-based. These methods can be further distributed into different techniques.

2.7.1 Structure-based virtual screening

In general, structure-based virtual screening (SBVS) docks compound databases into the target of interest. Overall, structure-based screening results in predictions of the ligand-target complex and a ranking from the docking.³¹ This approach requires knowledge of the target (e.g., receptor, channel, protein). In the first step,

target identification and validation should be made, and modeling the 3D protein structure of the biological component of interest is mandatory.³²

The 3D structure is mainly determined through experimental methods like X-ray crystallography and NMR or computational methods like homology modeling or alpha fold. Computational algorithms dock the target protein with huge libraries and databases of small molecules.³²

2.7.1.1 Homology model

If no 3D structure is available to work with in docking, this issue can be resolved by creating computer-aided models of the 3D structures. The sequence of the target must be known for modeling. Those modeling tools use templates with sequence similarity, and predicting the structures is based on extrapolation experimental information from related protein structures.^{33, 34}

2.7.2 Ligand-based virtual screening

Another method in virtual screening is the ligand-based approach. This method uses information from several active ligands, not information from the target protein's structure. Based on the ligand information, this is the method of choice if the 3D target structure is not available or unknown, e.g., a G-protein-coupled receptor or protein structures resolved in the apo form. By establishing this method, an assumption was made that similar structures show similar biological activities. The active compounds were selected for this process to match the screening candidates.^{5, 35}

2.7.3 Docking

Molecular docking generally aims to predict the ligand-receptor complex structure using computation methods.³⁶

Furthermore, molecular docking is a powerful tool for high-throughput virtual screening of large compound libraries. In the field of drug discovery, structure-based drug design (SBDD) utilizes molecular docking simulations to predict the interaction between a small molecule (ligand) and a biological target (usually a protein). This simulation mimics the natural binding process, allowing researchers to estimate the binding mode (ligand orientation) and affinity (strength of the interaction) between the two molecules. Due to its effectiveness in predicting these crucial aspects,

molecular docking has become widely used in drug design research. Glide is software that leverages this approach, employing semi-empirical calculations to estimate the free energy of the ligand-target complex, which serves as an indicator of binding affinity.³⁷

2.7.3.1 Scoring function

The scoring function, a crucial component of the docking process, assigns values to the binding affinity between ligands and the target using suitable valuing functions. Glide was designed with a focus on efficiency, enabling it to comprehensively search positions, orientations, and conformations for the ligand in the Binding pocket at a consistent calculation pace for screening, providing reassurance about its effectiveness.³⁸

The scoring function in Glide is instrumental in assessing the suitability of a ligand within a protein's binding site, quantifying its binding affinity. It aims to rank ligand poses based on their potential to form stable complexes with the protein. The terms in the scoring function collectively compute a score for each ligand pose, with lower scores indicating stronger binding affinity. Glide subsequently ranks ligand poses based on these scores, facilitating prioritization for further investigation or optimization.^{38, 39, 40}

When calculating the scoring function, van der Waals interactions, Coulomb potential, hydrogen bonds, entropy, torsional stress, ligand, and protein desolvation penalty are taken into account.³⁹

2.8 Aim of this thesis

This master thesis aims to apply structure-based virtual screening techniques using Glide software to the protein structures of bromodomains of the *Candida* species *C. albicans* and *C. auris*. The main aim is to generate results through virtual screening using docking that will subsequently inform and guide NMR (Nuclear Magnetic Resonance) studies focusing on the bromodomain structures of *Candida* species. This approach will contribute to identifying potential modulators for the research of innovative antifungal agents and potent targets to combat *Candida* infections, especially in view of the increasing challenges posed by multidrug-resistant strains such as *Candida auris*. Furthermore, using computational methods in drug discovery, this study aims to streamline the process and reduce the time and resources required while expanding the chemical space to discover new and effective drug candidates.

3 Methods

3.1 Protein structure and co-crystallized ligands

As part of our initial information-gathering on the CaBdf 1, we conducted a literature review based on the primary sequence. This helped us gain a better understanding of the targets general knowledge and interactions, among other things. We also used the primary protein sequences to search the PDB for matching structures. The PDB is a publicly available database that contains data on biological macromolecules. The structure can be searched for in the PDB using the PDB code, information about the macromolecule, or the primary sequence itself. (<https://www.rcsb.org>)⁴¹

Our research led us to opt for structure-based virtual screening using a docking approach since we were not aware of any active small molecules. For our docking process, we utilized two structures of CaBdf 1 from the PDB and one structure for *Candida auris*. Unfortunately, at the time of our study, no resolved 3D structure was available for *Candida auris*. However, we were able to determine it through homology modeling using the "Swiss model". In the PDB, five *C. albicans* bromodomain structures are available: 5N13, 5N15, 5N16, 5N17, and 5N18.

We extracted the 3D structures of fungal bromodomains from the PDB. The structures were created using X-ray diffraction and have a resolution of 1.6 Å for CaBdf1 BD1 (5N17) and 1.45 Å for Ca Bdf1 BD2 (5N18).

3.2 Databases

To facilitate docking, we utilized various compounds from different libraries. Our first round of VS focused on natural compounds available from MolPort. (<https://www.molport.com/shop/index>). For the second round, we used the NMR department's in-house library. Lastly, we docked molecules with molecular weights ranging from 250 to 320 Da from MolPort to find hits suitable for NMR studies. These databases generally store compounds in line representations, such as SMILES, SMARTS, and InChI, which require careful attention to ensure accurate assignment of stereochemistry, partial charges, and ionization states during conversion.

3.3 Schrödinger Platform

For this thesis, the Schrödinger Platform for small-molecule drug discovery was used to perform the preparation and docking steps. The docking algorithm employed was Schrödinger's Glide algorithm.⁴²

3.4 Protein preparation

Protein preparation was carried out using Maestro's Protein Preparation wizard. It was utilized to prepare the proteins for docking. The first step involved importing the structures from the target. The protein structures were downloaded from the PDB (PDB entries 5N17 and 5N18).

For preparation, default settings were applied, except for missing side chains, for which the "Filling missing side chains using Prime" setting was used. Next, all unnecessary atoms and molecules, such as solvents and water, with a distance greater than 3 Å were removed from the workspace.

In the subsequent step, the H-Bond assignment was performed using default settings, and redundant waters were deleted, except for conserved waters in the binding pocket. These binding-site water molecules may play a crucial role in ligand binding affinity. They can, for example, increase the binding strength and build hydrogen bonds between the ligand and the target.⁴³

3.5 Ligand preparation

Ligands for docking were obtained from MolPort in either SDF or CSV format and from Assoc.-Prof. Julien Orts. The CSV format was converted to SDF. Conversion was accomplished using Maestro, DataWarrior, and Python3 directly from the command line. The primary objective of ligand preparation is to convert the 2D structures into suitable 3D structures, considering tautomerism and stereochemistry. Consequently, the potential ligands should be prepared in a manner that allows for subsequent use in docking.

Regarding ligand preparation, there are two approaches: the first involves performing the preparation directly within Maestro using the LigPrep tool. Alternatively, this task can be carried out from the command line. This work employed both methods.⁴⁴

LigPrep from the command line was done with this script:

```
import os
import multiprocessing as mp
from datetime import datetime

def run(input_file):
    input_dir = '/data/local/ebajric/Fragment_database/Split_sdf/Splitted_files/'
    output_dir = '/data/local/ebajric/Fragment_database/Split_sdf/ligprep_files_split/'
    os.system('/data/shared/software/schrodinger2021-1/ligprep -isd '+input_dir+'/'+input_file+'
-WAIT -ma 90 -bff 16 -pht 0.0 -epik -s 1 -osd '+output_dir+'/'+input_file)

if __name__ == "__main__":
    start_time = datetime.now()
    input_dir = '/data/local/ebajric/Fragment_database/Split_sdf/Splitted_files/'

    #get jobList for multiprocessing
    jobList = []

    for file in os.listdir(input_dir):
        if file.endswith('.sdf'):
            jobList.append(file)

    pool = mp.Pool(processes=30)
    pool.map(run, jobList)
    pool.close()
    pool.join()
    print('Finished in:')
    print(datetime.now()-start_time)
```

3.6 Receptor grid generation

After completing the protein and ligand preparation, the next step involves generating the receptor grid, which is crucial for docking. This step entails examining the binding pocket with the co-crystallized ligand and selecting the ligand. Upon selecting the ligand, a purple box appears on the screen, confirming the correct grid selection and that no further settings were made. In this process, the binding site is calculated and represented as a grid. The grid can be set differently to refine the scoring function during docking. This step was performed only once for each target. Before running Glide, we ensured that this process was successful. Redocking was performed using Ligand-docking with Glide and the co-crystallized ligand. The primary purpose of the Receptor Grid Generation tool is to define the active-site properties.⁴⁵

3.7 Redocking

A redocking was carried out to validate the grid. During redocking, the co-crystallized ligand was prepared and docked to the protein, and then the redocking poses were evaluated to determine if the fit was the same.

3.8 Glide

All docking steps were performed with Glide. Docking was performed in two different ways. First, docking with the natural products from the MolPort library was performed directly within the Maestro tool "Glide." The previously prepared ligands and the respective grid-receptor were selected as inputs, and three poses were generated. All other settings were kept on default.

The experiments were performed from the command line to scan small fragments. One more step was required: first, a Glide input had to be created, and then Glide could be performed.

Performing Glide input:

```
import os
import sys

# usage
# python run_glidefile.py

list =
os.listdir("/data/local/ebajric/Fragment_database/Split_sdf/ligprep_files_split/") #
directory of input ligands
grid = "/data/local/ebajric/Fragment_database/glide-
grid_1_small_frag_5N18.zip"
target_name = '5N18' # depends on how you named your grid...

for in_file in list:
    #if (in_file[-4:] == ".mae" and in_file[:36] == "iissc-002-500-000--002-999-
999_prep_8") :
        outfile_name =
        "/data/local/ebajric/Fragment_database/Split_sdf/glide_input_files/" +
        target_name + in_file[:-4] + "_glide.in"
        outfile = open(outfile_name, 'w')
        print ("Preparing glide file for file: " + in_file)
        file_content = "FORCEFIELD OPLS_2005\nGRIDFILE " + grid \
        +
        "\nLIGANDFILE /data/local/ebajric/Fragment_database/Split_sdf/ligprep_files_s
plit/" + in_file \
        + "\nPOSES_PER_LIG 3\nPRECISION SP"
        outfile.write(file_content)
        outfile.close()
```

Utilizing the terminal for docking to run glide with this script:

```
import os
from datetime import datetime

if __name__ == "__main__":
    start_time = datetime.now()
    input_dir = '/data/local/ebajric/Fragment_database/Split_sdf_5N17_1-300/Glide_Input_files_1-300_no_water/'

    for file in os.listdir(input_dir):
        if file.endswith('.in'):
            input_file=input_dir+'/'+file
            os.system('/data/shared/software/schrodinger2021-1/glide -WAIT -OVERWRITE -adjust '+input_file+' -HOST localhost:32')

    print('Finished in:')
    print(datetime.now()-start_time)
```

In the end, we received 30 files for the docking, which should be merged to obtain the docking results in one file.

For this, we run from the command line:

```
outfile_name = '/data/local/ebajric/Fragment_database/Split_sdf/merge_input.lst'

outfile = open(outfile_name, 'w')
for i in range(1,30):
    file_content =
        "/data/local/ebajric/Fragment_database/Split_sdf/glide_results/5N18Selection_of_2023-12-19_molport_molecularWeight_600-restPart"+str(i)+"_glide_pv.maegz\n"
    outfile.write(file_content)

outfile.close()
```

To combine the results, we executed:

```
/data/shared/software/schrodinger2021-1/utilities/glide_merge -o all_pv.maegz -r all.rept -f merge_input.lst &
```

Next, we ranked the best docking results:

```
/data/shared/software/schrodinger2021-1/utilities/glide_sort -o top10k_pv.maegz -r top10k.rept -n 10000 all_pv.maegz &
```

3.9 Principal component analysis

Principal component analysis is one of the oldest and most widely used statistical techniques for reducing dimensionality in data analysis and pattern recognition.^{46, 47, 48}

PCA was performed with six descriptors to present the chemical space of the structures used for docking. The descriptors were total molar weight, partition coefficient for n-octanol to water (cLogP), number of hydrogen bond donors and acceptors, polar surface area, and rotatable bonds.

The PCAs were determined with DataWarrior 6.1.0.⁴⁹ The two main components, which effectively capture the largest parts of the variance within the dataset, were plotted.

3.10 Clustering

Clustering was performed to subdivide the results of similar compounds and to check if there were similar clusters of docking results. We first clustered all results by ring fragments to see if the resulting hits were similar and if and how they differed from the co-crystallized ligands.

3.11 Hardware

The computational methods used hardware from an in-house setup consisting of a Linux computer with CentoOS Stream 8 as the operating system, 1 GPU (NVIDIA GeForce RTX 3090 with 24 GB of memory), and 32 CPU cores (AMD Ryzen 9 7950X 16-Core Processor). The 32 CPU cores were used for docking.

4 Results

We performed a virtual screening of libraries with more than one million molecules by using a docking approach. In total, we docked four different libraries: two databases of Assoc. Prof. Dr. Julien Orts, the MolPort collection of natural compounds and the MolPort library with molecular weight between 250 and 320 Dalton to the targets CaBdf 1 BD1, BD2 and the homology model of *C. auris* BD.

By manually inspecting the hit lists, we identified more than 60 promising potential modulators for the *C. albicans* BDs and the *C. auris* BD. These modulators were selected based on their docking score, interactions with key residues, and fit with the binding pocket, taking molecular surface complementarity into account. Here, we will present the docking poses of some selected molecules with the most promising properties, including high docking scores and interactions with key residues in the bromodomain binding pocket.

In addition, we analyzed the chemical space of the virtual libraries by using principal component analysis with calculated physicochemical descriptors to highlight the differences and similarities in the properties of our hit list. Moreover, we performed a similarity analysis between the hits and the co-crystallized ligands of the bromodomains of *C. albicans* and *C. auris*, with the results also discussed here.

These top candidates will be further analyzed *in vitro* by NMR to validate their binding to the *Candida* bromodomain and to evaluate their potential as antifungal agents. All potential hits are listed in the Tables S1, S2, S3, S4 and S5 in the appendix according to the target and their respective physicochemical properties.

4.1 Database analysis

For the virtual screening, four different databases were utilized. The MolPort library of natural compounds was recommended by Assoc.-Prof. Dr. Julien Orts due to its high chemical diversity, encompassing numerous molecules that are either natural compounds or derivatives thereof with established pharmaceutical applications. The molecular weight of substances in the natural products library is between 45,04 g/mol and 8099,25 g/mol, and it contains more than 113.000 chemical structures. This database comprises both natural products and natural-like products.

Given the considerable variance in size and molecular structure among natural compounds, a subset of the whole MolPort database with more than 886.000 chemical structures with a molecular weight between 250 and 320 Da was employed during the docking study. This specific molecular weight range was selected to identify molecules suitable for subsequent NMR studies, ensuring their solubility for further experimentation.

The molecule list provided by Assoc. Prof. Dr. Julien Orts includes his existing compounds in the laboratory, which were screened to determine the molecules already available for in vitro studies on bromodomains. The molecular weight in this database was between 109.13 g/mol and 404.40 g/mol for a total of 896 compounds. In addition, the FDA database which contains drugs, was curated by Assoc. Prof. Dr. Julien Orts was consulted to assess the docking affinity of 686 substances already on the market with the target proteins under investigation.

4.2 Docking results

The analysis revealed 65 molecules suitable for in vitro NMR studies. A hit is defined as a molecule that, on one hand, exhibits the same interactions as the co-crystallized ligand or more interactions than the ligand, and, on the other hand, fits into the binding pocket, considering molecular surface complementarity. Hits are primarily ranked based on a scoring function; the lower the score, the better the binding in the binding pocket described by the docking algorithm. While interactions between the ligand and target are of primary importance, it is crucial to assess complementarity by analyzing surfaces and to consider this when determining whether the pose is well accommodated in the binding pocket.

4.3 Predicted hits

4.3.1 Predicted hits for the target *Candida albicans* bromodomain factor 1 bromodomain 1

For the CaBdf1 BD1 target to which the subset of the MolPort library was docked, we identified 24 molecules based on 18 different scaffolds. The ones we identified as the most promising hits are based on two different scaffold types. All scaffolds are shown in Table S1 in the appendix. The docking poses in Figures 6, 7, and 8

illustrate the precise interactions between the protein and the most potential scaffolds.

First, they were selected based on the docking score results. All results showed a docking score between -9,7 and -8. Second, there were very similar interactions with the protein, identical to the co-crystallized ligand but with even more interactions. Furthermore, our docked molecules adopted positions deeper in the binding pocket and showed a nicer molecular surface complementarity than the co-crystallized ligand.

MolPort-014-136-309 docked on the target CaBdf1 BD1 in Figure 6, which contains water inside the binding pocket, shows hydrogen bonds with the residues PRO 233, ASN 291, and TYR 248. The hydrogen bond with PRO 233 is formed between the oxygen carbonyl of PRO233 and the secondary amino group of the ligand. Together with its secondary amino group, ASN 291 forms a hydrogen bond with the carbonyl oxygen of the thiazepine ring. The next interaction in Figure 1 is a hydrogen bond mediated via water in position 514 between TYR 248 and thiazepine's carbonyl oxygen.

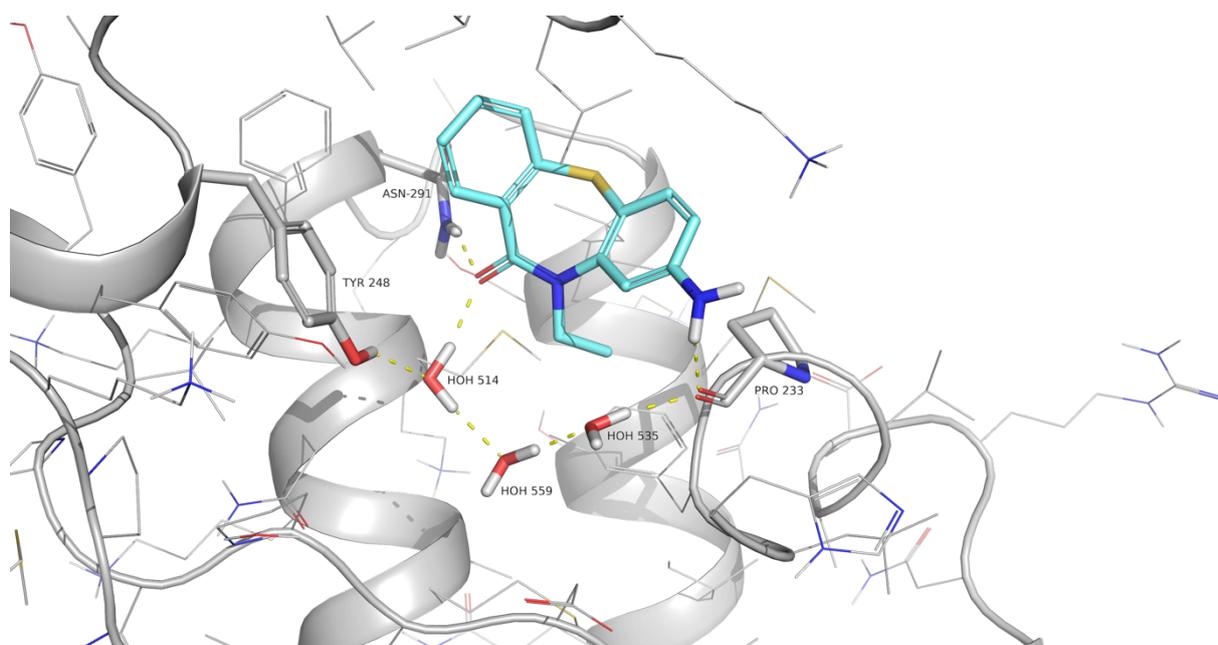


Figure 6 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-015-136-309 from the subset of MolPort library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

Our second hit for CaBdf1 BD1, MolPort-006-821-342, shown in Figure 7, is predicted to form four key interactions with the target. The ligand forms two hydrogen bonds with the residue ASN 291. One interaction is predicted to establish a hydrogen bond between the secondary amino group of the pyrimidine ring and the carbonyl oxygen of ASN 291. Another interaction with ASN 291 involves the amino group of the amino acid and a nitrogen atom of the pyrimidine. Additionally, a hydrogen bond is formed between the nitrogen of the imidazole ring and PRO 233. Another hydrogen bond is observed between the second nitrogen of the imidazole, mediated through a water molecule in position 514, and TYR 248.

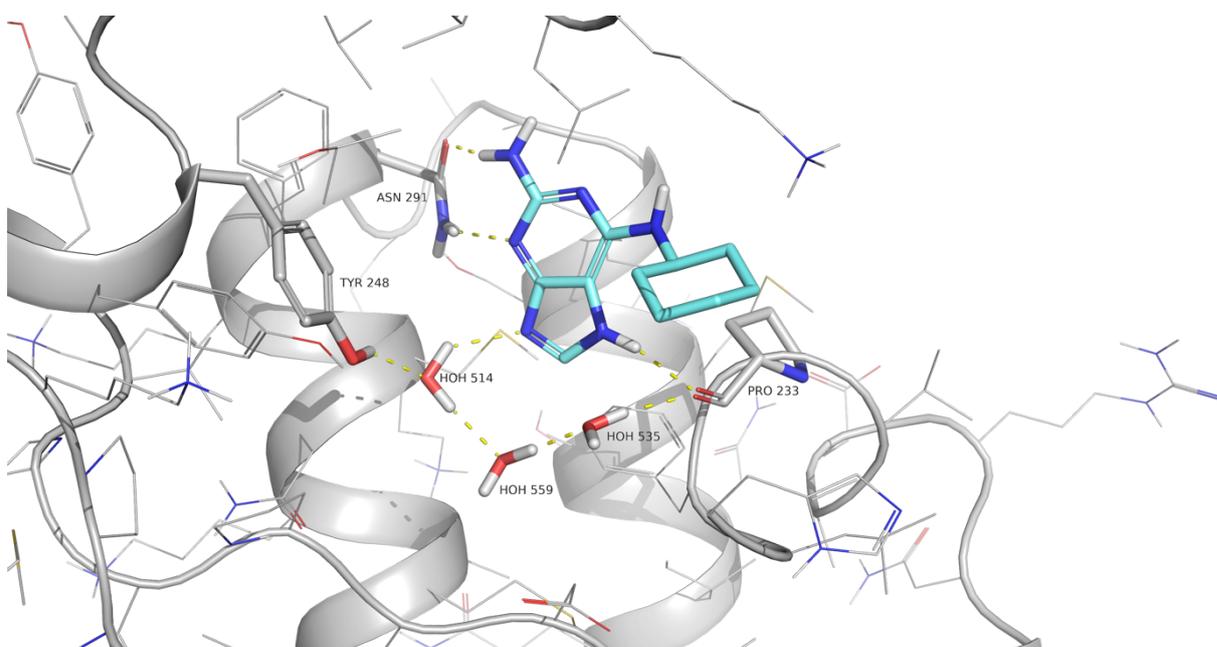


Figure 7 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-006-821-342 from the subset of MolPort library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

The third ligand, MolPort-046-848-539, at the target CaBdf1 BD1, is predicted to establish the same interactions as the previously described ligand MolPort-006-821-342 in Figure 8.

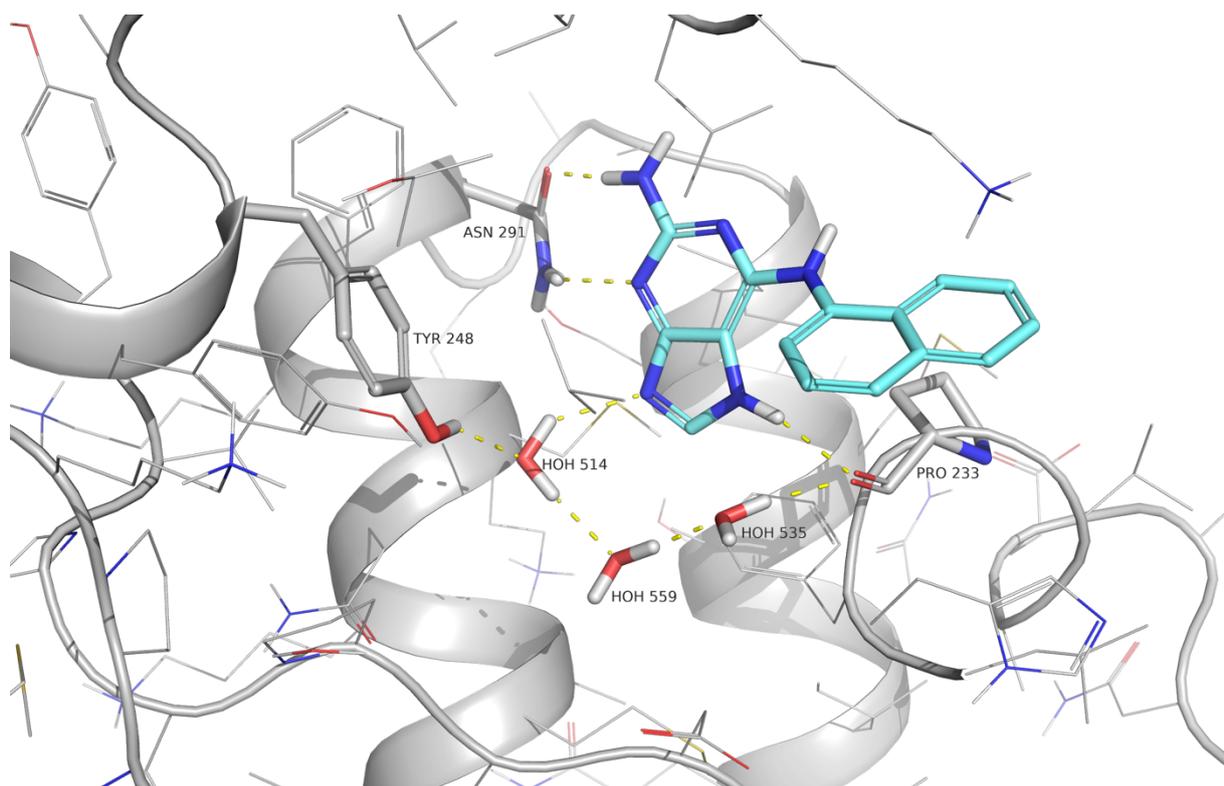


Figure 8 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-046-848-539 from the subset of MolPort library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

In the results of the docking run for CaBdf1 BD1 with the natural compound database, we identified six promising compounds based on four different scaffold types. The scaffolds are shown in Table S2 in the appendix. The detailed interactions between the protein and the most potential scaffolds are shown as docking poses in Figure 9 and in Figure 10.

The hits were analyzed according to the docking score, which was between -9,2 and -8 in this case. Finally, they were evaluated as hits based on their interactions and interactions compared to those of the protein's co-crystallized ligand and their fit in the binding pocket. Again, all the molecules we selected showed good complementarity of the molecular surface with the protein's surface.

Our first docking pose for the MolPort natural compound database in Figure 9 illustrates the interaction between the hit compound MolPort-008-348-689 from the MolPort natural compound database and CaBdf1 BD1. This docking pose highlights several hydrogen bonds between the ligand and the protein. The first hydrogen bond at the bottom of Figure 9 is formed between the hydroxyl group of the phenol moiety

in the ligand and the carbonyl oxygen in the backbone of MET 283. Additionally, a water molecule within the binding pocket mediates a hydrogen bond between TYR 248 and the ligand's phenol hydroxyl group. One more hydrogen bond is formed with the ASN 291 residue.

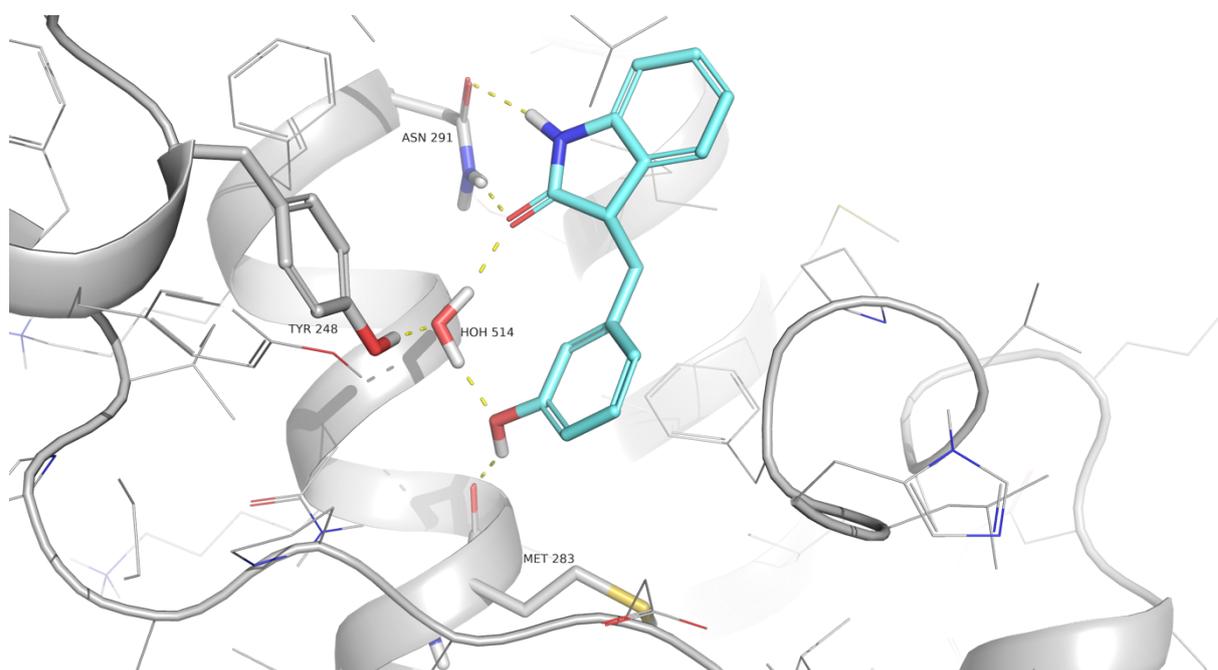


Figure 9 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-008-348-689 compound from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

Figure 10 shows the second compound MolPort-005-909-990, which was evaluated as a hit docked on the target CaBdf1 BD1. This compound shares almost the same interactions with the same amino acid residues as the MolPort-008-348-689 compound. However, MolPort-005-909-990 forms a hydrogen bond with the other MET in position 256, where the nitrogen from the piperazine moiety interacts with the amino residue.

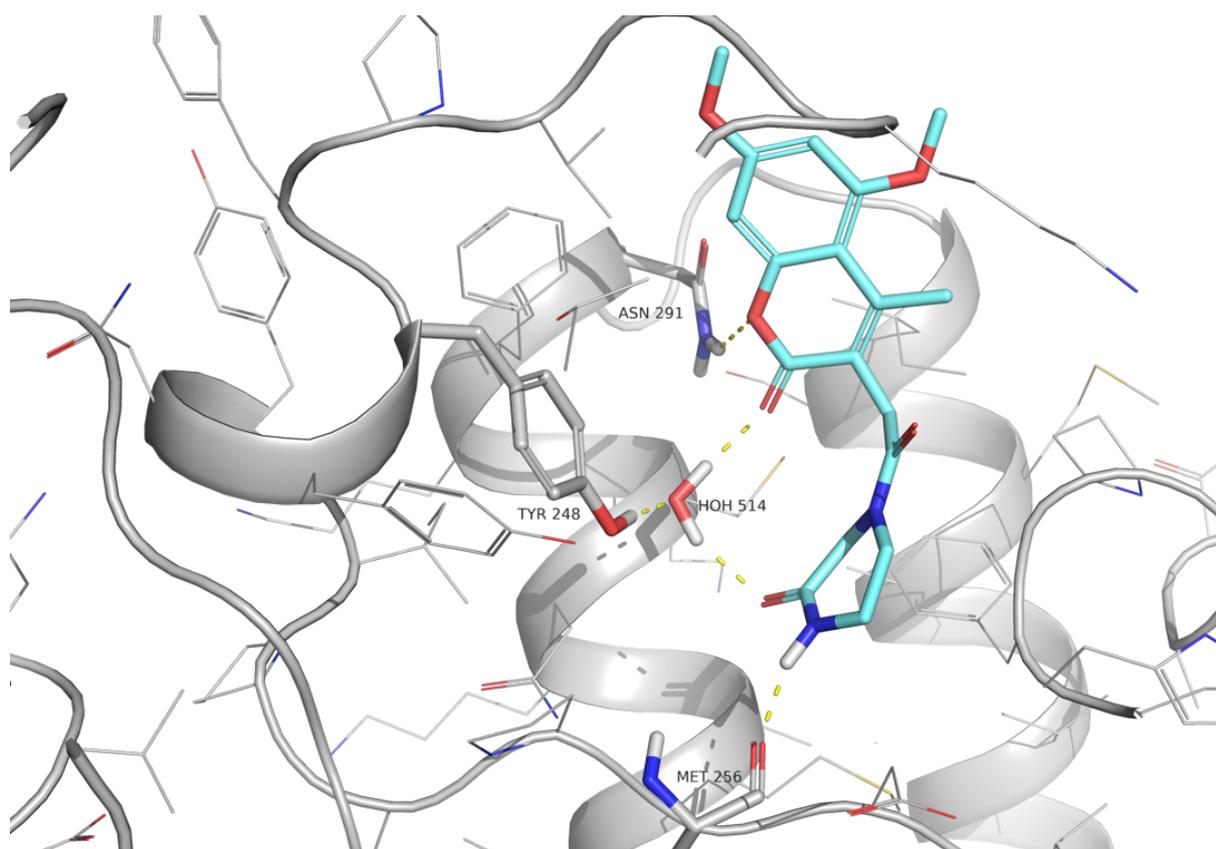


Figure 10 - Protein structure CaBdf1 BD1 represented as cartoon, with the MolPort-005-909-990 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

4.3.2 Predicted hits for the target *Candida albicans* bromodomain factor 1 bromodomain 2

For the target CaBdf1 BD2, we identified 24 molecules as promising in the docking results with the subset of the MolPort library based on 16 different scaffold types. The scaffolds are shown in Table S3 in the appendix. The detailed interactions of the most potential scaffolds with the protein are shown as docking poses in Figure 11, 12, 13 and 14.

First, the results were selected based on their docking score. All results showed a docking score between -9.5 and -8. Secondly, they interact very similarly with the protein as its co-crystallized ligand but show even more interactions. Our docked molecules also showed good surface complementarity in the binding pose and are adopting deeper positions in the binding pocket.

Figure 11 shows MolPort-019-801-991. In this docking pose, water molecules are present in the binding site. All hydrogen bonds in Figure 11 are mediated by water molecules. The ligand's pyrimidine structure forms a water-mediated hydrogen bond in position 640 with PRO 410. The other hydrogen bond is formed between the phenol's hydroxyl group and VAL 460 via the water molecule in position 633.

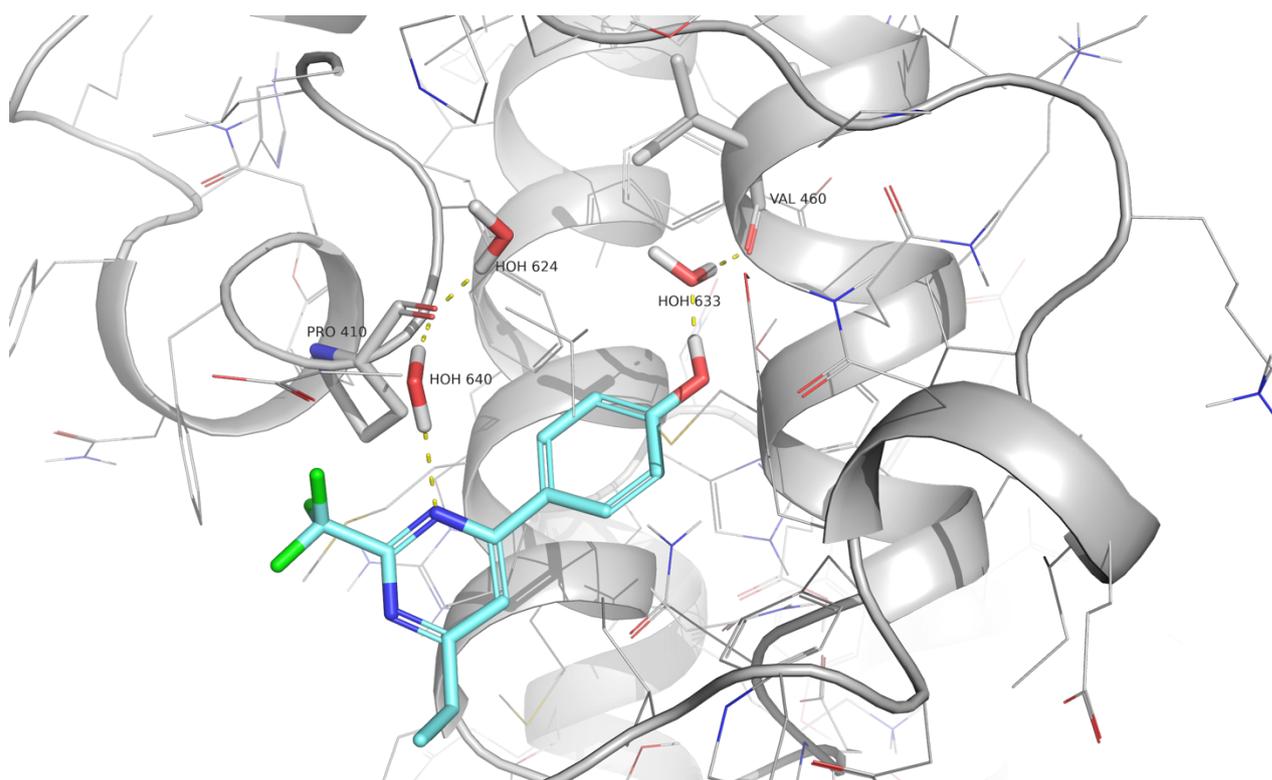


Figure 11 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-019-801-991 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.

MolPort-003-997-631 in Figure 12 displays the same interactions as MolPort-019-801-991 with PRO 410 and VAL 460.

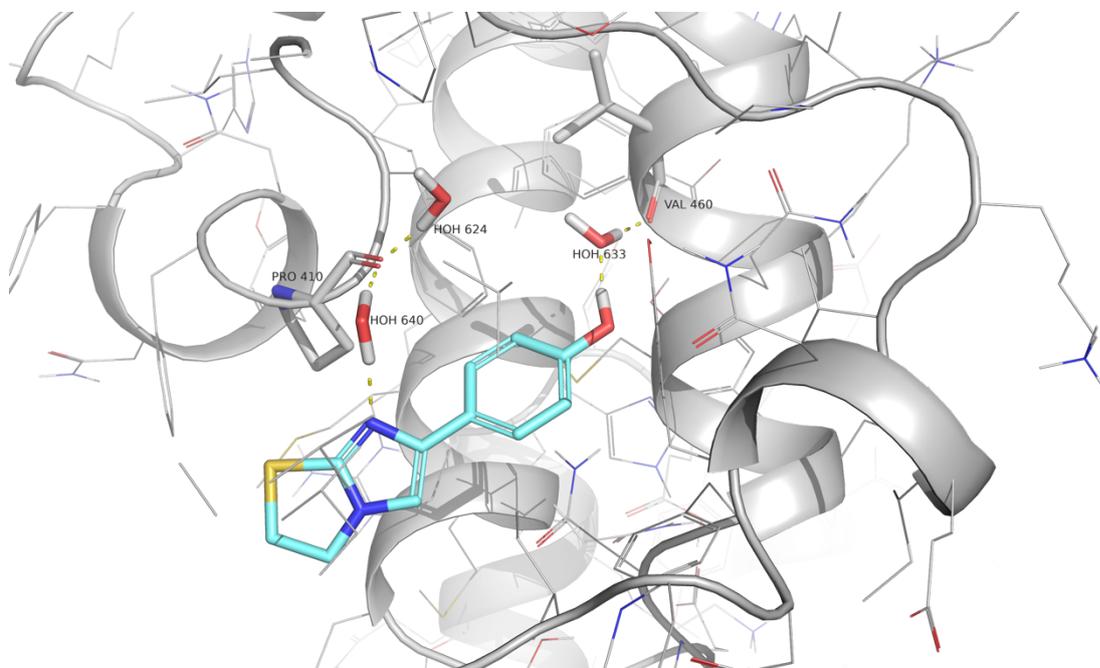


Figure 12 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-003-997-631 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.

In Figure 13, MolPort-038-415-941 is docked into CaBdf1 BD2 without water molecules in the binding pocket. The molecule is docked slightly deeper in the binding pocket and shows interactions with PRO 410 and ASN 468.

ASN 468 forms two hydrogen bonds with the ligand. One bond is located between the secondary amino group of the ligand and the carbonyl oxygen of ASN 468. The next bond is between the pyrimidine of the ligand and the amino group of ASN 468. The hydrogen bond with PRO 410 is predicted to establish an interaction between prolines carbonyl oxygen and the ligand's secondary amino group.

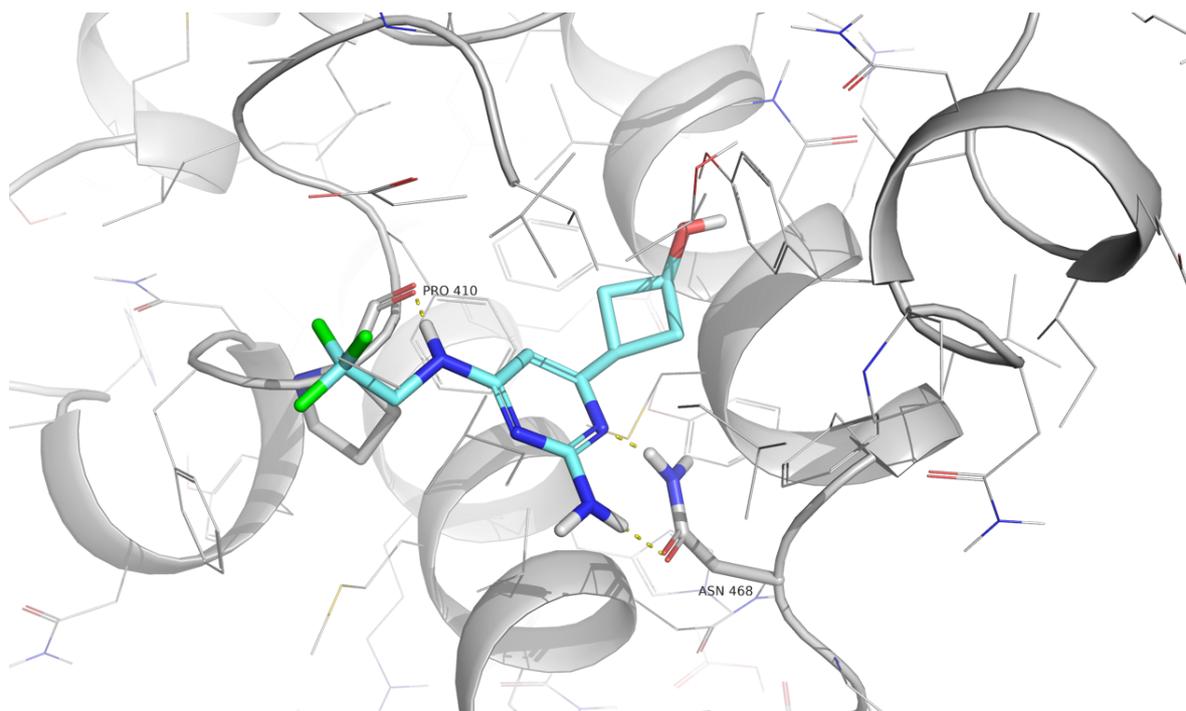


Figure 13 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-038-415-941 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.

Figure 14 illustrates the docked molecular structure of MolPort-001-026-787 and its interactions with the target CaBdf1 BD2. The ligand forms a hydrogen bond with ALA 413 via its secondary amino group to the carbonyl oxygen of ALA 413. Additionally, three hydrogen bonds are formed with the carbonyl oxygen of PRO 410. The third hydrogen bond is predicted to establish an interaction between the carbonyl oxygen of the ligand and the amino group of ASN in position 416. MolPort-001-026-787 forms a π - π interaction between TYR 425 and the benzene ring of the ligand in this docking pose.

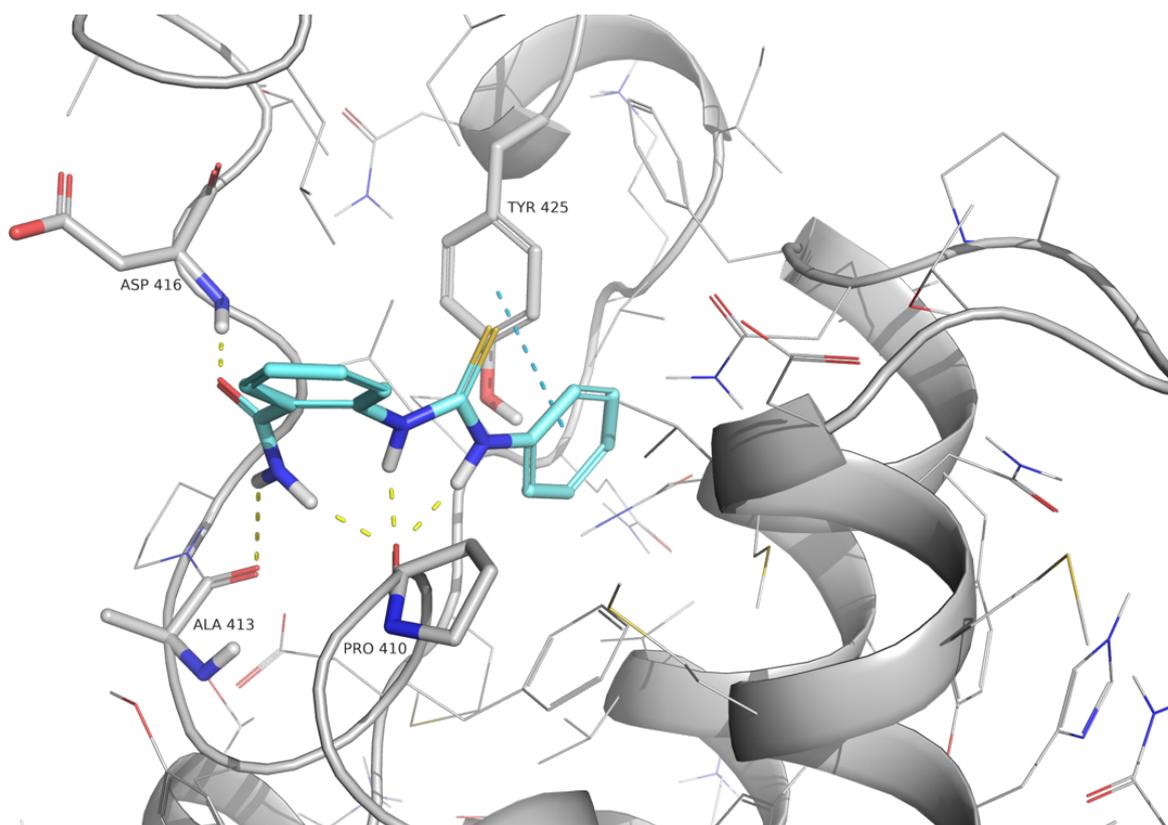


Figure 14 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-001-026-787 compound from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.

In the results of the docking run with the natural compound database, we identified seven compounds of interest based on six different scaffold types. All scaffolds are shown in Table S4 in the appendix. The detailed interactions of the most promising scaffolds with the protein are shown as docking poses in the Figure 15 and Figure 16.

The compounds were analyzed based on their docking scores ranging between -10 and -8,6. Finally, they were evaluated as hits based on their interactions with the protein and their fit within the binding pocket. All our selected molecules exhibited good complementarity between the molecular and protein surfaces.

One of our hits from the natural product database for the target CaBdf1 BD2, with MolPort-000-672-474, shown in Figure 15, exhibits completely different interactions compared to the co-crystallized ligand. The compound forms two hydrogen bonds: one with MET 433 and another with PRO 410. The hydrogen bond with MET 433 is formed via the hydroxyl group on the phenol moiety to the carbonyl oxygen of

methionine in position 433. The other hydrogen bond is formed via a nitrogen atom in the pyrimidine ring to the carbonyl oxygen of PRO 410.

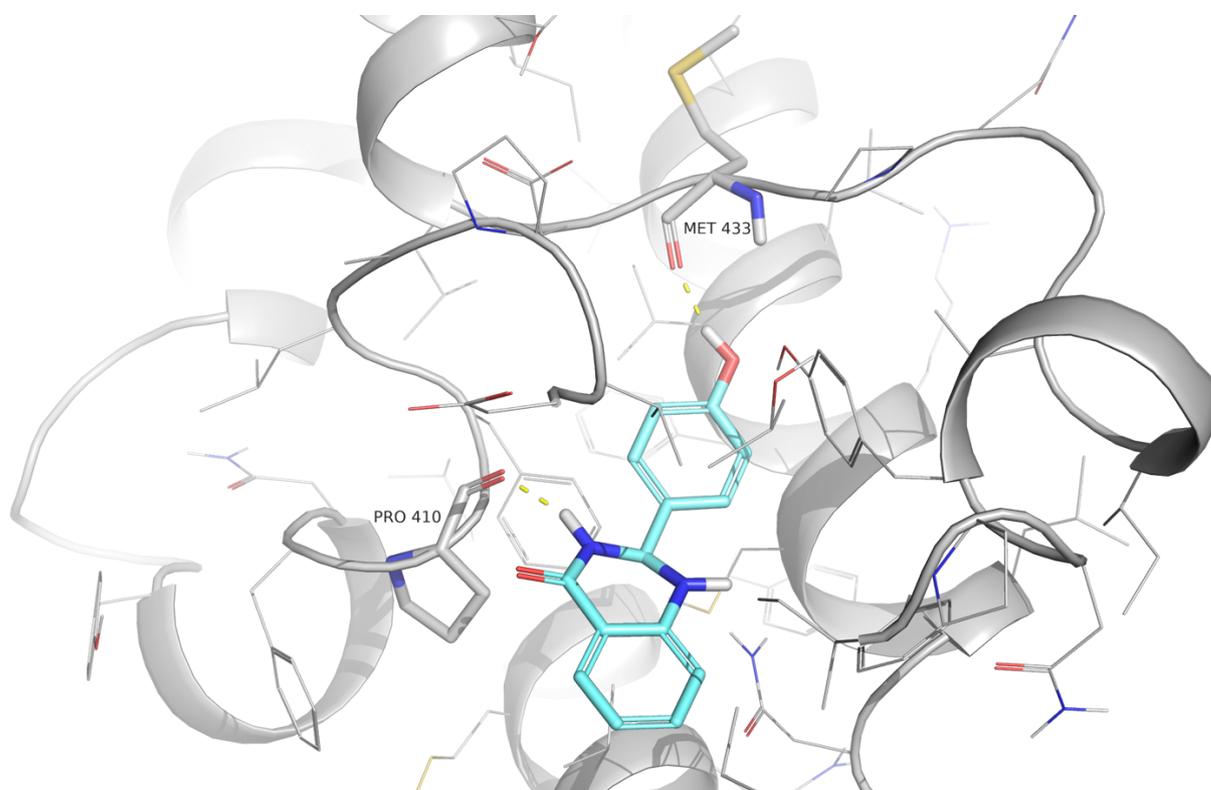


Figure 15 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-000-672-474 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

Figure 16 shows the second molecule, MolPort 001-732-545, from the natural product database, which we classified as a promising hit concerning the target CaBdf1 BD2. This molecule also forms the two hydrogen bonds already mentioned in Figure 15 with the amino acids MET 433 and PRO 410. Here, the phenol hydroxyl group forms a hydrogen bond with the carbonyl oxygen of MET 433. Additionally, the ligand forms a hydrogen bond between a nitrogen atom of the imidazole and the carbonyl oxygen of PRO 410. Furthermore, the hydrogen bond already observed in the co-crystallized ligand with ASN 468 is present; here, the second nitrogen of the imidazole ring forms a hydrogen bond with the primary amino group of ASN 468.

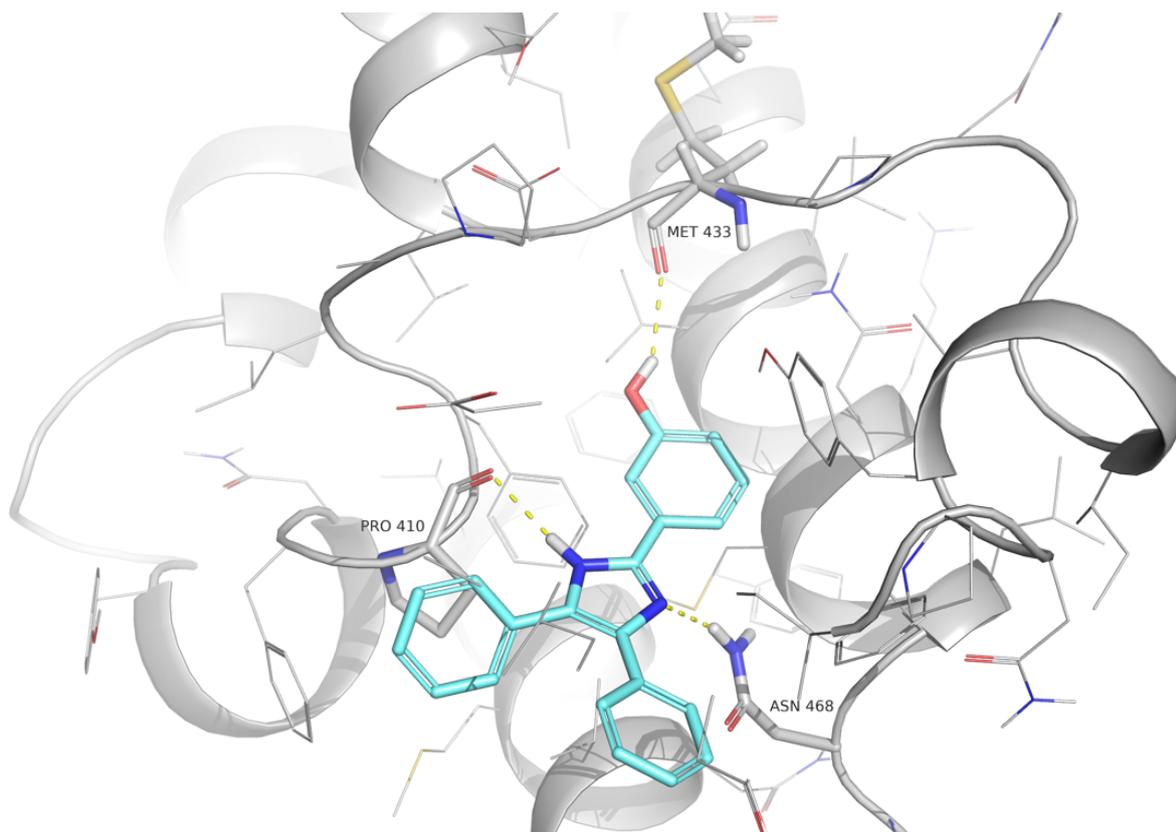


Figure 16 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-001-732-545 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

Figure 17 shows the last hit for the target CaBdf1 BD2 with the predicted interactions from the MolPort natural products Database with the MolPort-001-759-343. This molecule shows the already-known interaction with MET 433. In position 233, you can see a hydrogen bond between the ligand's primary amine group and MET 233 carbonyl oxygen.

The next hydrogen bond is between the carbonyl oxygen of the ligand and the hydroxy group of TYR 425. TYR 425 shows another interaction, namely a π - π -interaction between the tyrosine phenol and the ligand's pyrrole ring.

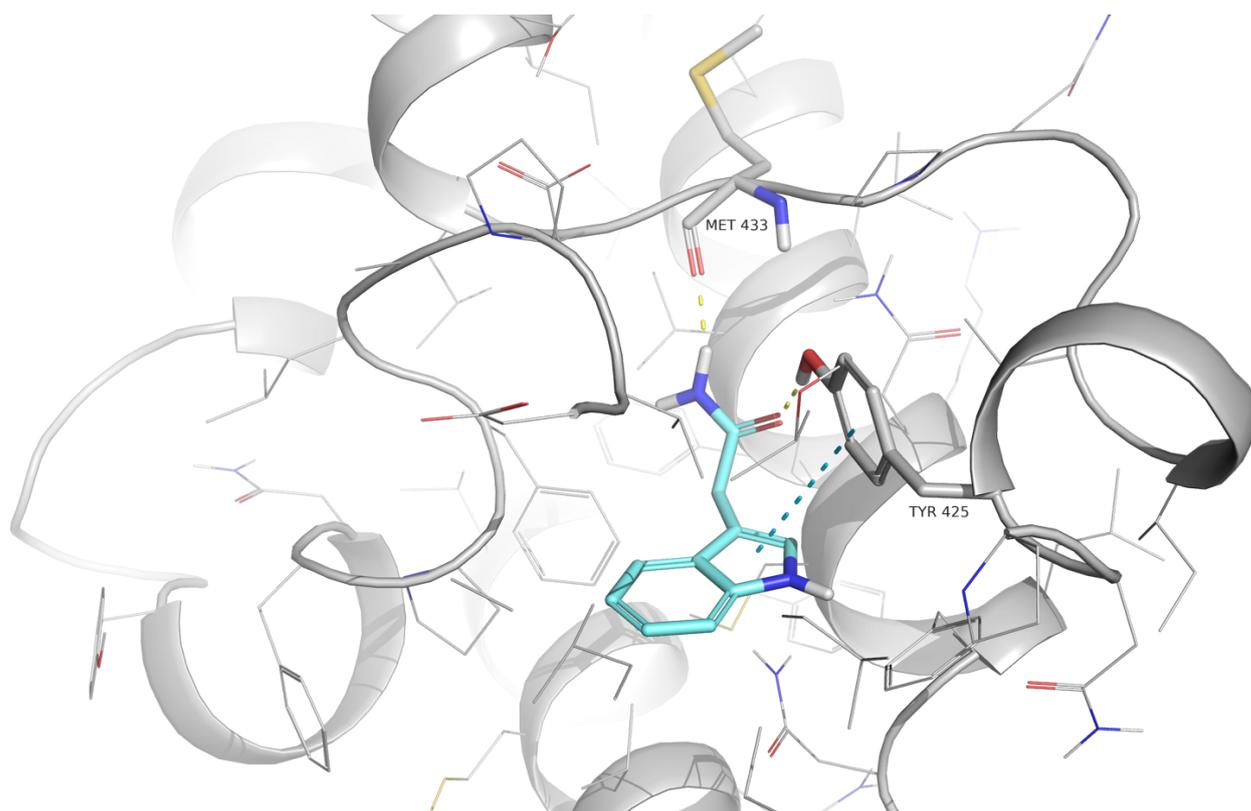


Figure 17 - Protein structure CaBdf1 BD2 represented as cartoon, with the MolPort-001-759-343 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes for hydrogen bonds and as blue dashes for π - π -interactions.

4.3.3 Predicted hits for the target *Candida auris* bromodomain

For the target *C. auris*, we identified four compounds of interest based on two different scaffold types for the docking run with the MolPort natural compound database. All scaffolds are shown in Table S5 in the appendix. The detailed interactions of the most promising scaffolds with the protein are shown as docking poses in Figure 18.

The compounds were also analyzed based on their docking scores, which ranged between -9,1 and -8,5. Finally, the compounds were evaluated as hits based on their interactions with the protein and their fit within the binding pocket. All our selected molecules exhibited good complementarity between the molecular and protein surfaces.

The docking analysis of ligand MolPort-000-840-542 with the bromodomain of *C. auris* reveals several key interactions. The amino group of ASN 101 forms a hydrogen bond with the carbonyl oxygen of the quinoline moiety, while the nitrogen

of the quinoline forms a hydrogen bond with the backbone's carbonyl oxygen of PRO 43. Additionally, the amino group of the quinoline interacts with the carbonyl oxygen in the backbones of MET 93 and MET 66.

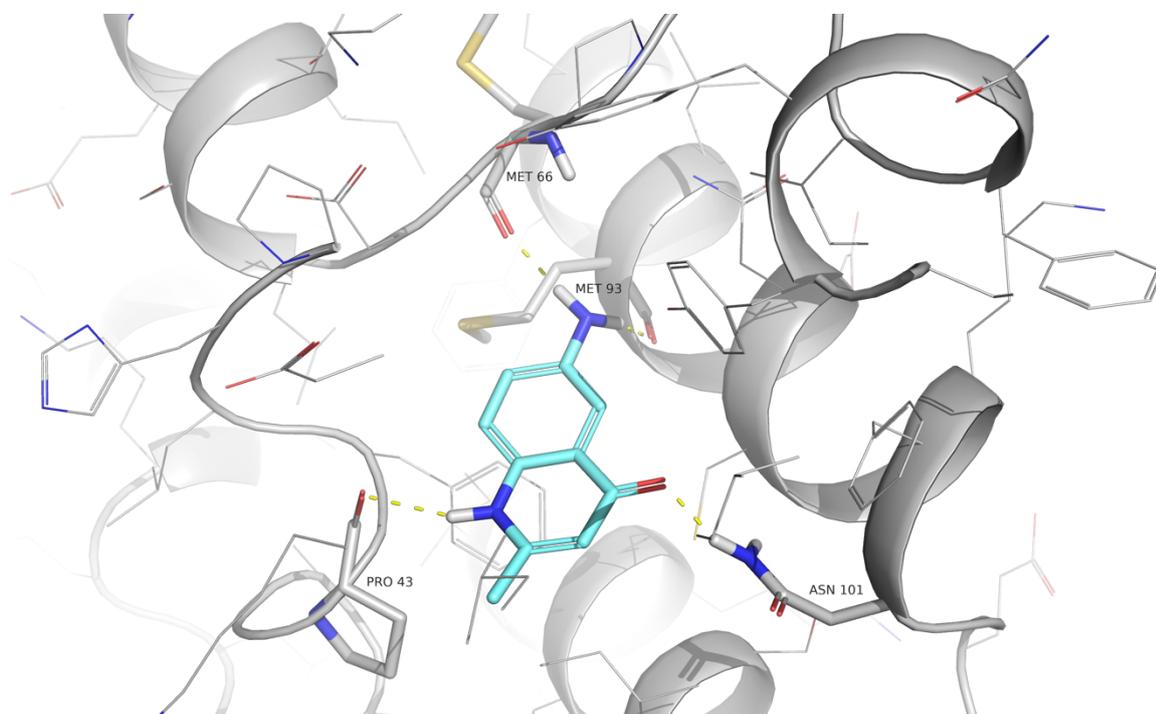


Figure 18 - Protein structure C. auris represented as cartoon, with the MolPort-000-840-542 from the MolPort natural compound library, shown as aquamarine sticks and the amino acid residues forming interactions represented as gray sticks. The predicted interactions are shown as yellow dashes.

4.4 Chemical space analysis

The PCA results complement the docking analysis by providing an overview of the molecular landscape investigated in the study. While docking analysis provides detailed information about the interactions between a compound and the target protein, PCA offers a comprehensive perspective on the overall distribution of our hits in the different libraries we used for docking.

4.4.1 Chemical space analysis of MolPort compounds with molecular weight between 250 and 320 Daltons

4.4.1.1 *Chemical space analysis of MolPort compounds with molecular weight between 250 and 320 Daltons for the target Candida bromodomain factor 1 bromodomain 2*

A principal component analysis (PCA) shown in Figure 19 was performed to visualize the distribution of a subset of the MolPort library with a molecular weight range of 250-320 daltons and the chemical space of the more than 800.000 molecules contained in this database.

Figure 19 shows the MolPort docked against the target CaBdf1 BD2 based on their physicochemical properties (H-bond-acceptors, H-bond-donors, total molecular weight, topological surface area, rotatable Bonds, cLogP). The first two principal components (PC1 and PC2) are shown, explaining 35,47% and 29,34% of the variance in the data, respectively. The orange circles represent the investigated hits identified during the docking process, while the blue circles represent the remaining compounds in the MolPort library that are not considered hits. The distribution of data points across the PCA space indicates various physicochemical properties within the screened library according to the chosen descriptors. Notably, the investigated virtual hits (orange circles) are not segregated within a specific region of the PCA plot, they are rather central but relatively far apart in general. Occasionally, however, a few hits are closer together than others.

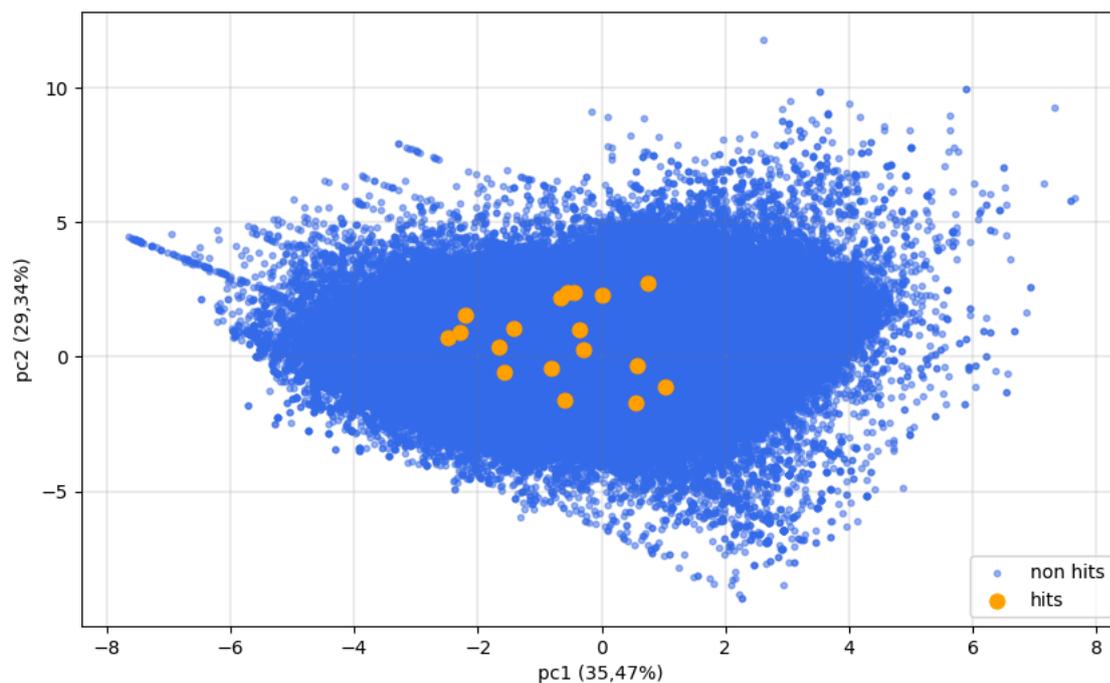


Figure 19 - The PCA scatter plot illustrates the chemical space covered by the subset of the MolPort database with the obtained docking hits for the target CaBdf 1 BD2, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP.

4.4.1.2 Chemical space analysis of MolPort compounds with molecular weight between 250 and 320 Daltons for the target Candida bromodomain factor 1 bromodomain 1

For the same subset MolPort database again a PCA calculation was performed but this time for the target CaBdf1 BD1. Figure 20 shows the MolPort database distribution as blue circles and our hits in orange circles. The blue circles cover more than 800.000 molecular fragments with the obtained hits for the target CaBdf1 BD1.

As in Figure 19 the same six molecular descriptors were used for calculating the PCA. The first two principal components (PC1 and PC2) are shown and explain 35.47% and 29.34% of the variance in the data, respectively. For this target, the PCA shows no clusters or favorable hit-regions. They show a large distribution over the central area of the scatter plot.

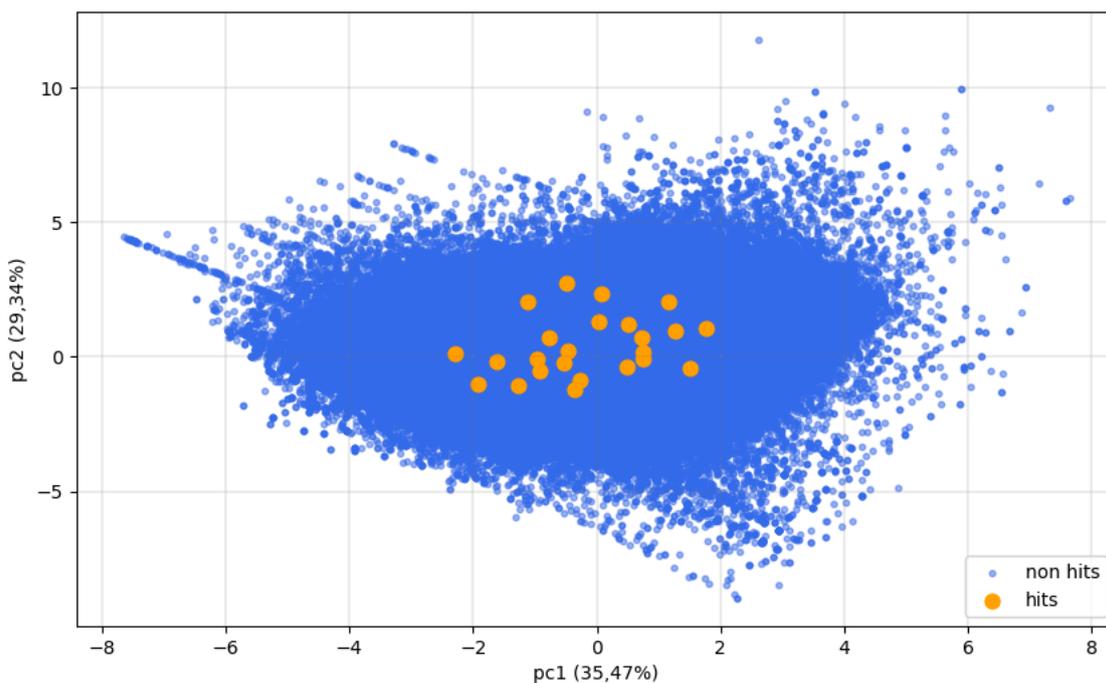


Figure 20 - The PCA scatter plot illustrates the chemical space covered by the subset of the MolPort database with the obtained docking hits for the target CaBdf 1 BD1, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP.

4.4.2 Chemical space analysis of MolPort natura compound library

4.4.2.1 Chemical space analysis of MolPort natural compounds with docking hits

Candida albicans bromodomain factor 1 bromodomain 2

Figure 21 shows a principal component analysis for the MolPort natural compound database in blue colored circles and the obtained hits for the target CaBdf1 BD2 as orange circles. The PCA was performed to visualize the chemical space and distribution of the natural compound library from MolPort used for docking the target CaBdf1 BD2 based on their physicochemical properties. The first two principal components (PC1 and PC2) are shown, explaining 60.8% and 22.59% of the variance in the data, respectively. As already mentioned, orange circles represent our investigated hits, identified during the docking process, while the blue circles represent the remaining compounds in the library that were not considered as hits. The chemical space is covered by more than 113.000 chemical structures from the natural product database of MolPort.

The distribution of data points across PCA space attests to the diversity of physicochemical properties within the library, as indicated by the descriptors chosen

(H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP). The hits are generally concentrated in the upper right quadrant of the scatter plot and are relatively close to each other, apart from a single hit.

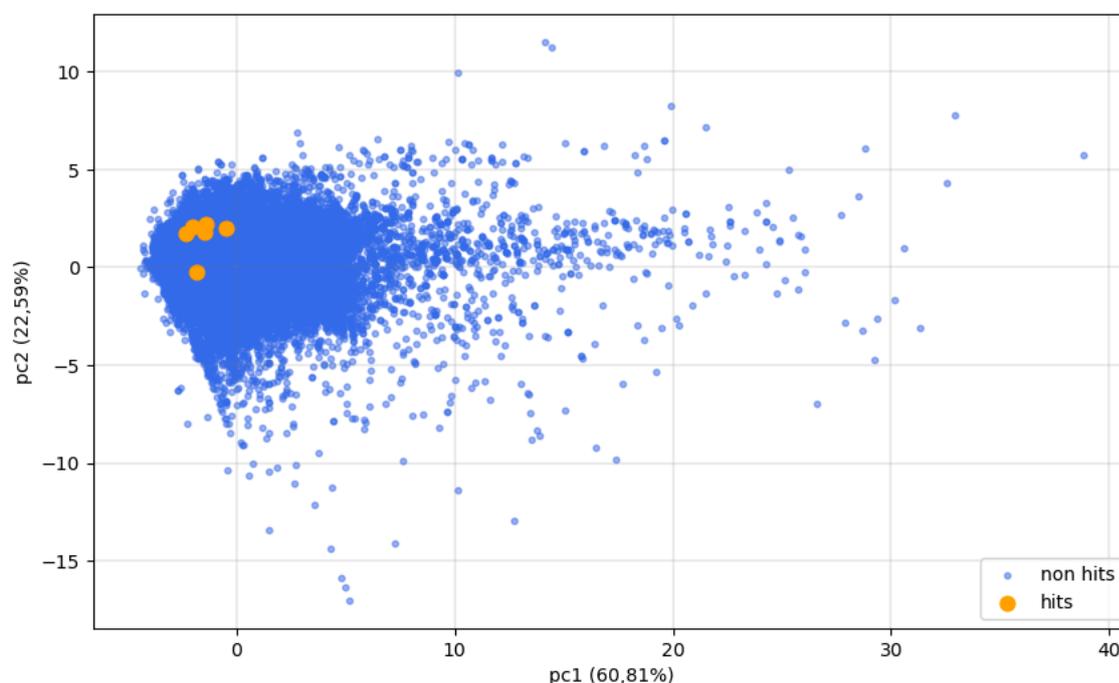


Figure 21 - The PCA scatter plot illustrates the chemical space covered by the MolPort database natural compounds with the obtained docking hits for the target CaBdf 1 BD2, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP.

4.4.2.2 Chemical space analysis of MolPort natural compounds with docking hits *Candida albicans* bromodomain factor 1 bromodomain 1

The subsequent scatter plot, visualized in Figure 22, was also performed with DataWarrior. Figure 22 shows the PCA of the MolPort natural products database with the obtained hits of the target CaBdf1 BD1. Here, the PCA was also performed to visualize the chemical space and the distribution of the natural compound library from MolPort used for docking to the target CaBdf1 BD1 based on their physicochemical properties. The chemical space is covered by the chemical structures of the natural compound library and the six identified promising hits of CaBdf1 BD1.

The PCA was calculated using these six key descriptors: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, and cLogP.

The first two principal components, PC1 and PC2, are instrumental in our analysis. PC1 accounts for a significant portion of the variance in the data, explaining 60.8% of it. PC2, while less impactful, still contributes significantly, explaining 22.59% of the variance.

In Figure 22, the chemical space of the database is shown as blue circles, and our promising hits are orange circles. The hits for the target CaBdf1 BD1 are localized in the upper part of the database and divided into three clusters. The outer two clusters represent only one hit each. The remaining hits are very close to each other and between the two outer hits.

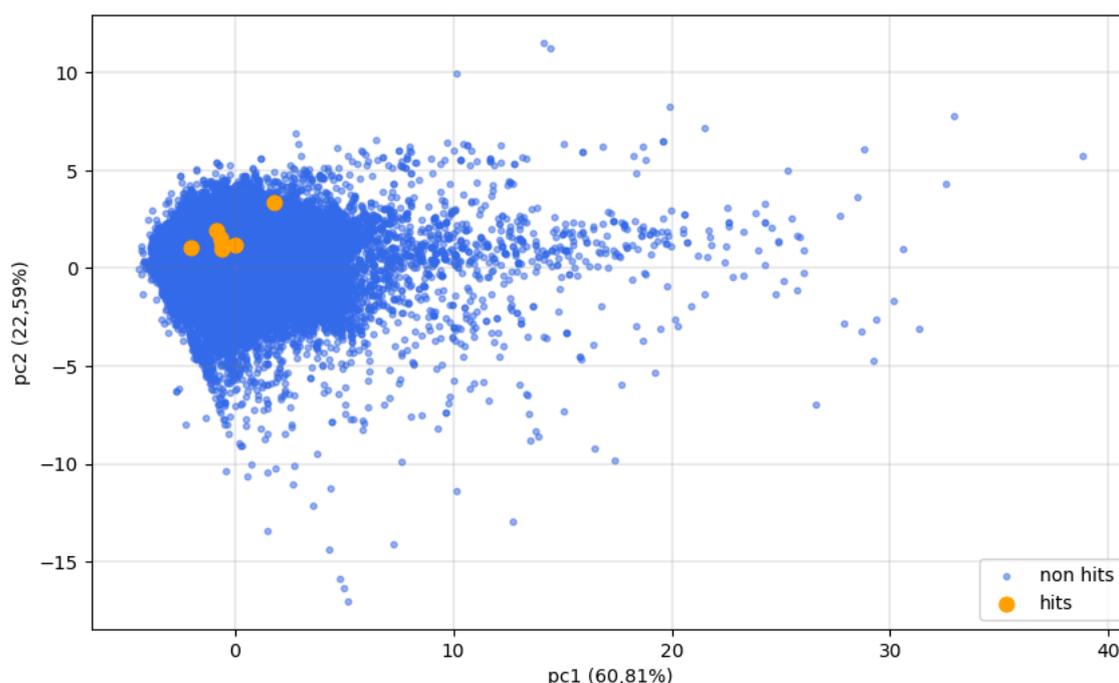


Figure 22 - The PCA scatter plot illustrates the chemical space covered by the MolPort database natural compounds with the obtained docking hits for the target CaBdf 1 BD1, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP.

4.4.2.3 Chemical space analysis of MolPort natural compounds with docking hits *C. auris*

The last scatter plot in Figure 23 again represents the chemical space and the distribution, from the MolPort natural product database, but this time with the results of docking to the *C. auris* bromodomain. In this PCA, the same chemical properties

(H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds and cLogP) were used to calculate the PCA as in Figures 21 and 22. The more than 116,000 natural products and natural product-like molecules are shown as blue circles and the potential ligands for *C. auris* as orange circles. In Figure 23, the hits are located next to each other in the upper left corner, but no clear distribution pattern was detected for them.

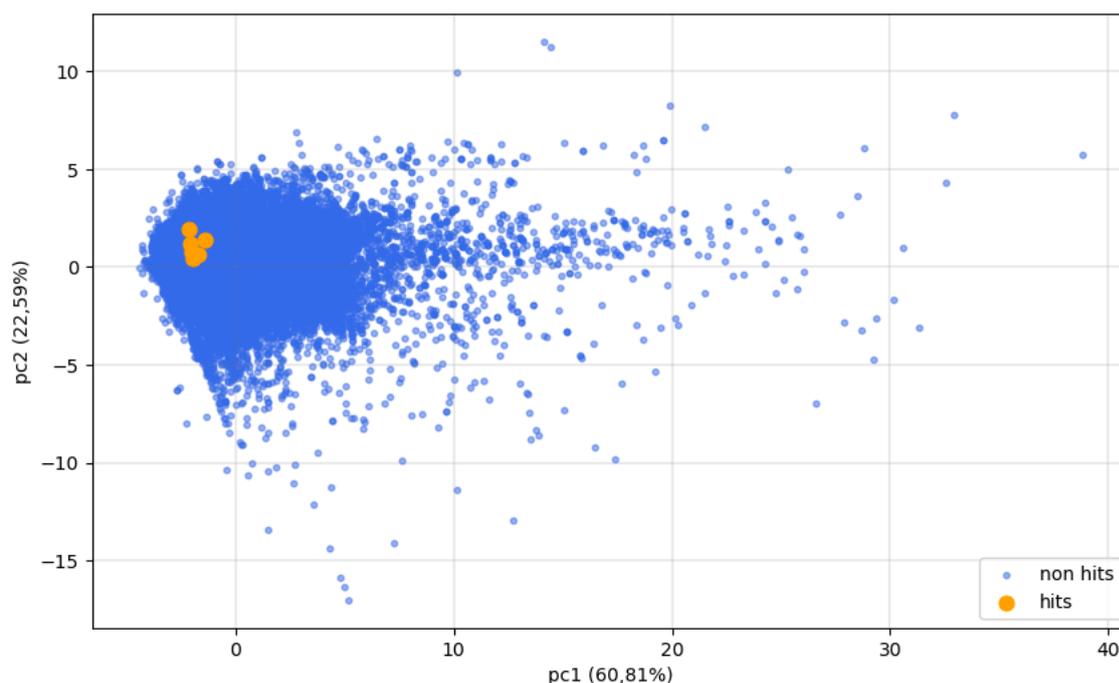


Figure 23 - The PCA scatter plot illustrates the chemical space covered by the MolPort database natural compounds with the obtained docking hits for the target *C. auris*, calculated by six physicochemical properties: H-bond acceptors, H-bond donors, total molecular weight, topological surface area, rotatable bonds, cLogP.

4.5 Clustering

To further analyze the docked hits and identify potential scaffolds showing promising activity, clustering by ring fragments was performed using the DataWarrior. This analysis was applied to hits obtained from docking to the three targets: CaBdf1 BD1, CaBdf1 BD2 and the homology model of *C. auris*. Ring analysis with DataWarrior extracted the ring fragments from our hit lists, for showing promising scaffolds which are interacting with the proteins. The aim of this analysis was to identify potential structural clusters among the hits for each target. The clustering results can be found in the Appendix in the Table S1, S2, S3, S4 and S5 in column Ring fragments.

4.6 Similarity analysis

4.6.1 Similarity analysis of the co-crystallized ligand of *Candida* bromodomain containing factor 1 bromodomain 1 and hits obtained from MolPort natural compounds

Figure 24 shows the calculation of the similarity between our promising hits of the natural compound library and the co-crystallized ligand (8FK) of CaBdf1 BD1. The calculation was performed with RDKit based on Tanimoto similarity. Figure 24 represents a bar chart where each bar represents a molecule from our hit-list in blue color. The similarity score is located on the y-axis and indicates the similarity in percent with respect to the co-crystallized ligand.

We have also calculated an average similarity, which is shown as a red dashed line in the chart. The average similarity score across all docking hits was approximately 15%, indicating a moderate structural resemblance to the reference ligand. Our docking predictions suggest that the structure of the co-crystallized ligand has not been fully optimized. This could be relevant for further ligand optimization and identifying more suitable scaffolds.

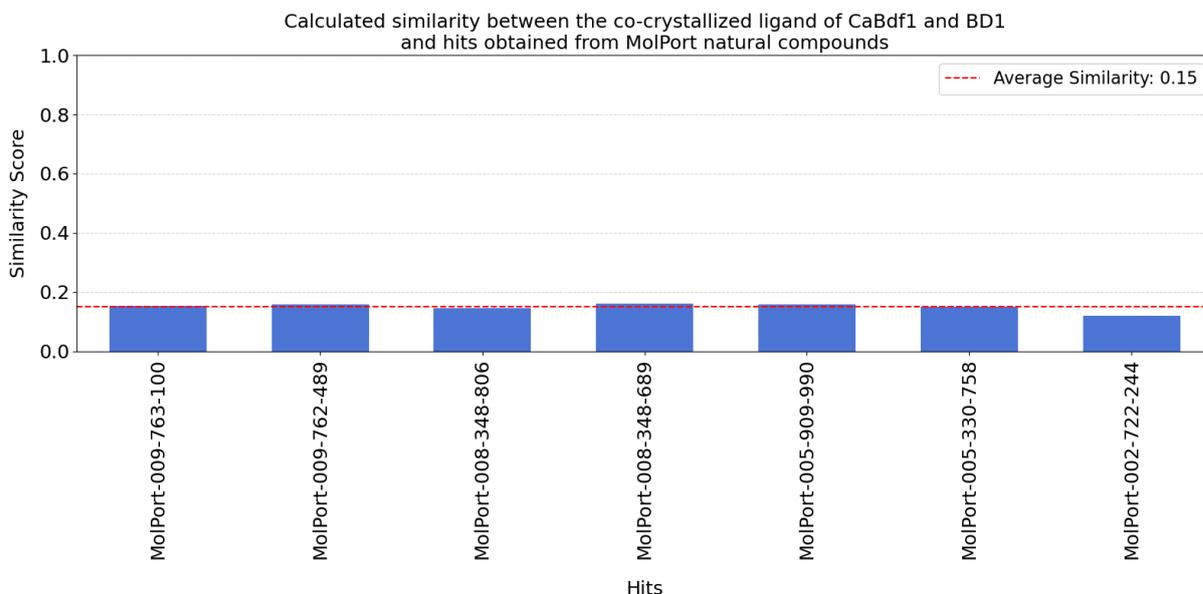


Figure 24 - The bar chart shows the similarity between our promising hits of the natural compound library and the co-crystallized ligand (8FK) of CaBdf1 BD1. Hits are represented as blue bars, and the average similarity as a red, dashed line.

4.6.2 Similarity analysis of the co-crystallized ligand of *Candida* bromodomain containing factor 1 bromodomain 2 and hits obtained from MolPort natural compounds

Figure 25 shows the calculation of the similarity between our promising hits of the natural compound library and the co-crystallized ligand (8HZ) of CaBdf1 BD2. The calculation was performed with RDKit based on Tanimoto similarity. Figure 25 represents a bar chart where each bar represents a molecule from our hit-list in blue color. The similarity score is located on the y-axis and indicates the similarity in percent with respect to the co-crystallized ligand.

We have also calculated an average similarity which is shown as a red dashed line in the chart. The average similarity score across all docking hits was approximately 12%, indicating a low structural resemblance to the reference ligand. Based on our docking predictions, this suggests that the structure of the co-crystallized ligand has not been fully optimized. This could be relevant for further ligand optimization and identifying more suitable scaffolds.

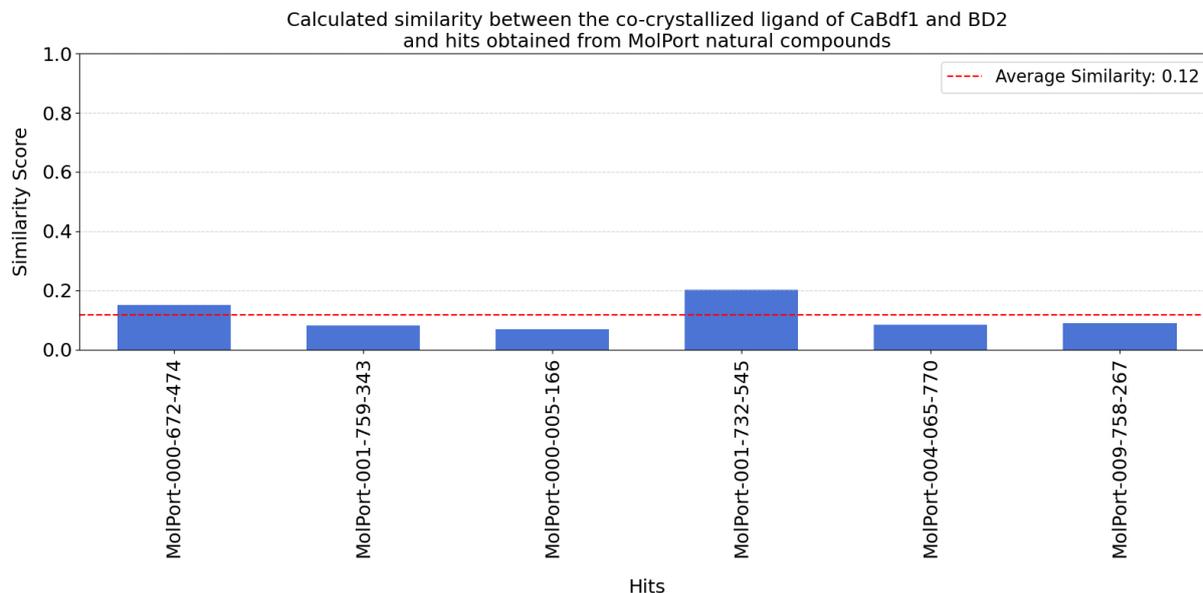


Figure 25 - The bar chart shows the similarity between between our promising hits of the natural compound library and the co-crystallized ligand (8HZ) of CaBdf1 BD2. Hits are represented as blue bars, and the average similarity as red dashed line.

Among all docking results with the MolPort natural product library, the average similarity scores showed no significant similarities between the co-crystallized ligands off both targets and their hits.

4.6.3 Similarity analysis of the co-crystallized ligand of *Candida* bromodomain containing factor 1 bromodomain 1 and hits obtained from MolPort compounds with molecular weight between 250 and 320 Daltons

Figure 26 illustrates the calculation of similarity between our promising hits from the subset of MolPort library compounds with molecular weights between 250 and 320 Daltons and the co-crystallized ligand (8FK) of CaBdf1 BD1. The calculation was performed using RDKit based on Tanimoto similarity. Figure 25 presents a bar chart where each bar represents a molecule from our hit list in blue bars, with the similarity score shown on the y-axis indicating the percent similarity to the co-crystallized ligand.

We have also computed an average similarity, depicted as a red dashed line in the chart. The average similarity score across all docking hits was approximately 15%, indicating a low structural resemblance to the reference ligand. According to our docking predictions, this suggests that the structure of the co-crystallized ligand may not be fully optimized. This finding could be significant for further ligand optimization and the identification of more suitable scaffolds. Notably, no docking result exhibited significant similarity with the co-crystallized ligand.

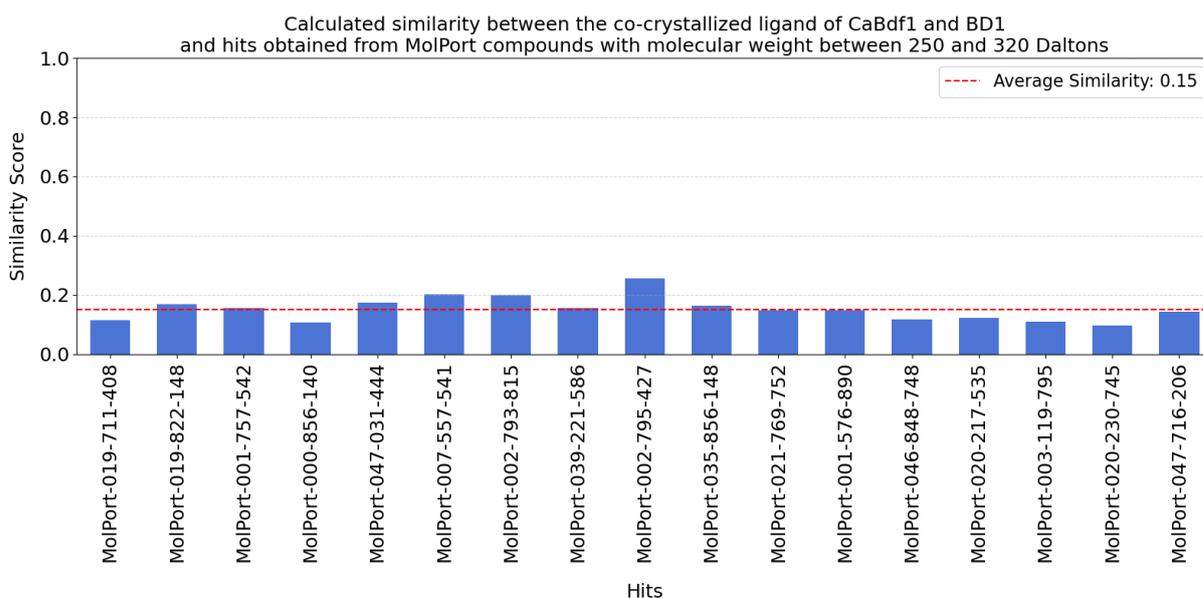


Figure 26 - The bar chart shows the similarity between our promising hits of subset of MolPort library and the co-crystallized ligand (8FK) of CaBdf1 BD1. Hits are represented as blue bars, and the average similarity as red dashed line.

4.6.4 Similarity analysis of the co-crystallized ligand of *Candida* bromodomain containing factor 1 bromodomain 2 and hits obtained from MolPort compounds with molecular weight between 250 and 320 Daltons

Figure 27 depicts the calculation of similarity between our promising hits from the subset of MolPort library compounds with molecular weights between 250 and 320 Daltons and the co-crystallized ligand (8HZ) of CaBdf1 BD2. The calculation was conducted using RDKit based on Tanimoto similarity. In Figure 27, each bar represents a molecule from our hit-list in blue color, with the similarity score on the y-axis indicating the percent similarity to the co-crystallized ligand. The average similarity score across all docking hits was approximately 24%, indicating a moderate structural resemblance to the reference ligand. According to our docking predictions, this implies that the structure of the co-crystallized ligand may not be fully optimized, which could be pertinent for further novel ligand finding optimization.

The compound MolPort-007-611-328 exhibited outstanding similarity, with a score of approximately 65% or higher. This could suggest a similar binding mode as the co-crystallized ligand and can be used for further optimization.

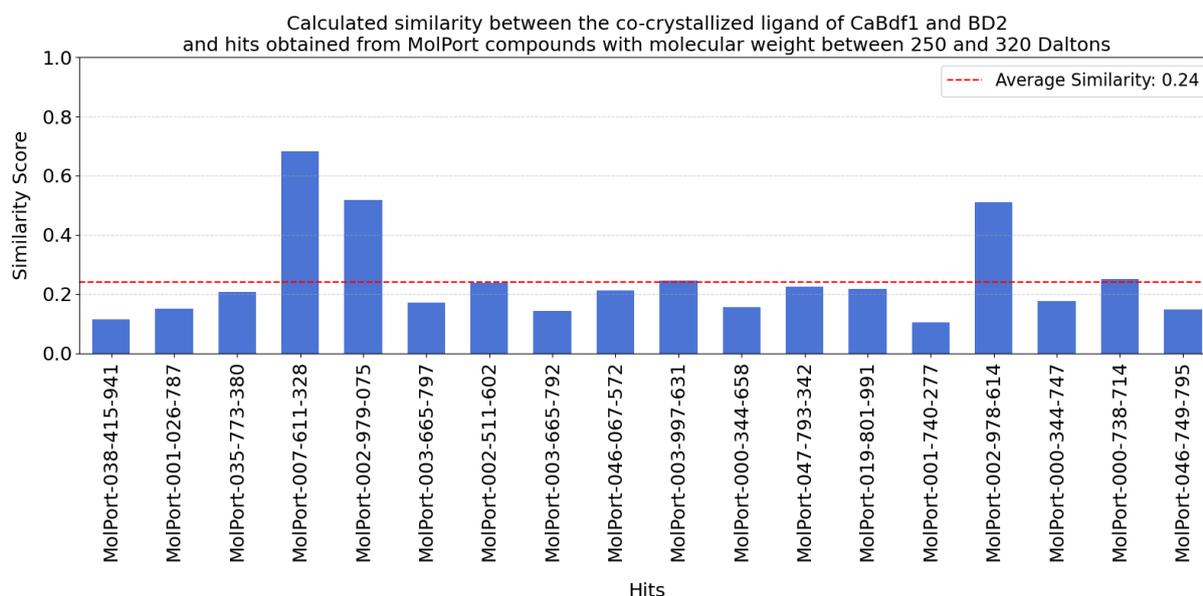


Figure 27 - The bar chart shows the similarity between our promising hits of subset of MolPort library and the co-crystallized ligand (8HZ) of CaBdf1 BD2. Hits are represented as blue bars, and the average similarity as red dashed line.

Similarity analysis for *C. auris* was not performed because no existing co-crystallized structure or information about active small molecules existed.

5 Conclusions

Multidrug-resistant *Candida* strains pose increasing challenges to public health. We employed structure-based virtual screening techniques to identify potential modulators for the bromodomains of *Candida* species.

PCA revealed the diverse chemical space of the investigated compound libraries and further provided insights into the distribution and physicochemical properties of the 65 compounds identified as promising hits. Similarity analysis revealed compounds with structural similarity to co-crystallized ligands. This allowed us to determine potential binding modes and indicate dissimilar structures or evaluate different molecules with potentially good binding features that can be used for affinity studies, further optimization, and ligand discovery.

The scaffold analysis identified common structural motifs that may be important for ligand binding, providing valuable insights into the possibility of deeper ligand optimization and drug discovery.

The predicted interactions provide detailed mechanistic insights into the potential binding modes of the identified hits and elucidate the potential fundamental interactions that govern ligand binding to the target proteins. This analysis highlights the importance of specific residues and structural features for ligand recognition and binding affinity.

The results should be tested and validated in the laboratory to confirm or deny the predictions. Suppose the tests with the compounds are positive, and the interactions can thus be confirmed. In that case, these results can accelerate further NMR-studies on these target molecules, as the scaffolds could be used to focus on specific molecules with specific properties. However, reprocessing the docking screen should be necessary if the laboratory studies yield negative results and the predicted interactions are not confirmed.

Future research could focus on validating the *in vitro* stability of the water molecules in the binding pockets, thus investigating the influence of ligands on water molecules during the binding process. In addition, our collaboration partner, Assoc.-Prof. Julien Orts, from the University of Vienna, will perform experimental validation of the hits by NMR.

6 References

1. Sayou, C. & Govin, J. Functions and inhibition of BET bromodomains in pathogenic fungi. *Curr. Opin. Green Sustain. Chem.* **34**, 100590 (2022).
2. Su, H., Han, L. & Huang, X. Potential targets for the development of new antifungal drugs. *J. Antibiot. (Tokyo)* **71**, 978–991 (2018).
3. Mietton, F. *et al.* Selective BET bromodomain inhibition as an antifungal therapeutic strategy. *Nat. Commun.* **8**, 15482 (2017).
4. Batool, M., Ahmad, B. & Choi, S. A Structure-Based Drug Discovery Paradigm. *Int. J. Mol. Sci.* **20**, 2783 (2019).
5. Hamza, A., Wei, N.-N. & Zhan, C.-G. Ligand-Based Virtual Screening Approach Using a New Scoring Function. *J. Chem. Inf. Model.* **52**, 963–974 (2012).
6. Lopes, J. P. & Lionakis, M. S. Pathogenesis and virulence of *Candida albicans*. *Virulence* **13**, 89–121 (2022).
7. WHO fungal priority pathogens list to guide research, development and public health action. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.
8. Da Silva Dantas, A. *et al.* Cell biology of *Candida albicans*–host interactions. *Curr. Opin. Microbiol.* **34**, 111–118 (2016).
9. Jabra-Rizk, M. A. *et al.* *Candida albicans* Pathogenesis: Fitting within the Host-Microbe Damage Response Framework. *Infect. Immun.* **84**, 2724–2739 (2016).
10. Chang, P., Fan, X. & Chen, J. Function and subcellular localization of Gcn5, a histone acetyltransferase in *Candida albicans*. *Fungal Genet. Biol.* **81**, 132–141 (2015).
11. Ahmad, S. & Alfouzan, W. *Candida auris*: Epidemiology, Diagnosis, Pathogenesis, Antifungal Susceptibility, and Infection Control Measures to Combat the Spread of Infections in Healthcare Facilities. *Microorganisms* **9**, 807 (2021).
12. Castro, L. A., Giusiano, G. & Martínez, E. *Candida auris* infection in the central catheter of a patient without sepsis symptoms. *Colomb. Medica* **50**, 293–98 (2019).
13. Chowdhary, A., Jain, K. & Chauhan, N. *Candida auris* Genetics and Emergence. *Annu. Rev. Microbiol.* **77**, 583–602 (2023).
14. Oliva, A. *et al.* Invasive *Candida* infection: epidemiology, clinical and therapeutic aspects of an evolving disease and the role of rezafungin. *Expert Rev. Anti Infect. Ther.* **21**, 957–975 (2023).
15. Du, H. *et al.* *Candida auris*: Epidemiology, biology, antifungal resistance, and virulence. *PLOS Pathog.* **16**, e1008921 (2020).
16. Chaabane, F., Graf, A., Jequier, L. & Coste, A. T. Review on Antifungal Resistance Mechanisms in the Emerging Pathogen *Candida auris*. *Front. Microbiol.* **10**, 2788 (2019).
17. Watkins, R. R., Gowen, R., Lionakis, M. & Ghannoum, M. Update on the Pathogenesis, Virulence, and Treatment of *Candida auris*. *Pathog. Immun.* **7**, 46–65 (2022).
18. Chatterjee, S. *et al.* Draft genome of a commonly misdiagnosed multidrug resistant pathogen *Candida auris*. *BMC Genomics* **16**, 686 (2015).
19. Sanglard, D. Emerging Threats in Antifungal-Resistant Fungal Pathogens. *Front. Med.* **3**, (2016).

20. Maubon, D., Garnaud, C., Calandra, T., Sanglard, D. & Cornet, M. Resistance of *Candida* spp. to antifungal drugs in the ICU: where are we now? *Intensive Care Med.* **40**, 1241–1255 (2014).
21. Lakhani, P., Patil, A. & Majumdar, S. Challenges in the Polyene- and Azole-Based Pharmacotherapy of Ocular Fungal Infections. *J. Ocul. Pharmacol. Ther.* **35**, 6–22 (2019).
22. Marak, M. B. & Dhanashree, B. Antifungal Susceptibility and Biofilm Production of *Candida* spp. Isolated from Clinical Samples. *Int. J. Microbiol.* **2018**, 1–5 (2018).
23. Kale, S., Goncareenco, A., Markov, Y., Landsman, D. & Panchenko, A. R. Molecular recognition of nucleosomes by binding partners. *Curr. Opin. Struct. Biol.* **56**, 164–170 (2019).
24. Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 178–189 (2015).
25. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487–500 (2016).
26. Filippakopoulos, P. *et al.* Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **149**, 214–231 (2012).
27. Domínguez-Andrés, J. *et al.* Bromodomain inhibitor I-BET151 suppresses immune responses during fungal–immune interaction. *Eur. J. Immunol.* **49**, 2044–2050 (2019).
28. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug Discov.* **13**, 337–356 (2014).
29. Zhou, Y. *et al.* Toward more potent imidazopyridine inhibitors of *CANDIDA ALBICANS* Bdf1: Modeling the role of structural waters in selective ligand binding. *J. Comput. Chem.* **43**, 2121–2130 (2022).
30. Gorgulla, C. Recent Developments in Ultralarge and Structure-Based Virtual Screening Approaches. *Annu. Rev. Biomed. Data Sci.* **6**, 229–258 (2023).
31. Ferreira, L., Dos Santos, R., Oliva, G. & Andricopulo, A. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules* **20**, 13384–13421 (2015).
32. Jahn, A., Hinselmann, G., Fechner, N. & Zell, A. Optimal assignment methods for ligand-based virtual screening. *J. Cheminformatics* **1**, 14 (2009).
33. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
34. Song, C. M., Lim, S. J. & Tong, J. C. Recent advances in computer-aided drug design. *Brief. Bioinform.* **10**, 579–591 (2009).
35. Scior, T. *et al.* Recognizing Pitfalls in Virtual Screening: A Critical Review. *J. Chem. Inf. Model.* **52**, 867–881 (2012).
36. Agu, P. C. *et al.* Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Sci. Rep.* **13**, 13398 (2023).
37. Fan, J., Fu, A. & Zhang, L. Progress in molecular docking. *Quant. Biol.* **7**, 83–89 (2019).
38. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
39. Friesner, R. A. *et al.* Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* **49**, 6177–6196 (2006).
40. Warren, G. L. *et al.* A Critical Assessment of Docking Programs and Scoring

- Functions. *J. Med. Chem.* **49**, 5912–5931 (2006).
41. Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
 42. Dickerhoff, J., Warnecke, K. R., Wang, K., Deng, N. & Yang, D. Evaluating Molecular Docking Software for Small Molecule Binding to G-Quadruplex DNA. *Int. J. Mol. Sci.* **22**, 10801 (2021).
 43. Cheng, T., Li, Q., Zhou, Z., Wang, Y. & Bryant, S. H. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *AAPS J.* **14**, 133–141 (2012).
 44. Schrödinger Release 2024-1: LigPrep, Schrödinger, LLC, New York, NY, 2024.
 45. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).
 46. Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
 47. Abdi, H. & Williams, L. J. Principal component analysis. *WIREs Comput. Stat.* **2**, 433–459 (2010).
 48. Bro, R. & Smilde, A. K. Principal component analysis. *Anal Methods* **6**, 2812–2831 (2014).
 49. Sander, T., Freyss, J., Von Korff, M. & Rufener, C. DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. *J. Chem. Inf. Model.* **55**, 460–473 (2015).

APPENDIX

Summary of all predicted hits for the target *Candida albicans* bromodomain factor 1 bromodomain 1

Table S1 shows the most promising potential docking hits for the CaBdf 1 BD1 target that were identified through molecular docking studies conducted with a subset of the MolPort library. Each hit includes its 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, and ring fragments. The DataWarrior platform was used to determine the cLogP values and cluster affiliations by ring fragments.

Structure	MolPort-ID	SMILES	Docking score	cLogP	Ring fragments
	MolPort-047-031-444	<chem>O[C@H](CCC[C@H]1NC(C[n]2c(cccc3)c3nc2)=O)[C@@H]1O</chem>	-10,148	0,177	
	MolPort-021-769-752	<chem>CC(C)(CCc1cccc(C(N2Cc3c[nH]nc3CC2)=O)c1)O</chem>	-9,741	1,854	
	MolPort-002-793-815	<chem>O=C(c(cccc1)c1N[C@@H]1c2ccc[o]2)N1c(cc1)ccc1F</chem>	-9,601	2,827	
	MolPort-001-757-542	<chem>Cc(c(N(C(c(cccc1)c1N1)=O)C1=S)ccc1)c1Cl</chem>	-9,158	3,571	
	MolPort-007-557-541	<chem>Cc(ccc(N(C(Nc1c2c(ccc1)=O)N2O)c1)c1)Cl</chem>	-9,151	2,819	
	MolPort-019-711-408	<chem>O=C(c([s]cc1)c1N1)N(c2cc(F)cc(F)c2)C1=S</chem>	-9,143	2,689	
	MolPort-001-576-890	<chem>CCc(cccc1)c1N(C(c(cccc1)c1N1)=O)C1=S</chem>	-9,093	3,380	
	MolPort-039-221-586	<chem>C[n]1nc(CCCC2)c2c1C(N(CC1)CC2=C1N=CNC2=O)=O</chem>	-9,081	0,165	
	MolPort-046-848-748	<chem>COc(cc1)ccc1N(Cc(cccc1)c1N1)C1=S</chem>	-9,062	2,380	
	MolPort-002-795-427	<chem>ON([C@@H](c(ccc(Cl)c1)c1Cl)Nc1c2ccc1)C2=O</chem>	-9,051	2,342	

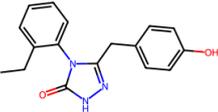
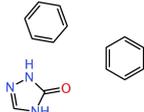
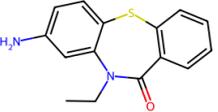
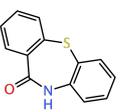
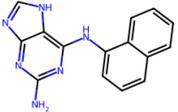
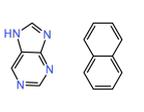
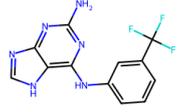
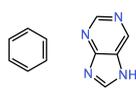
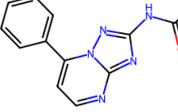
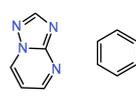
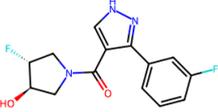
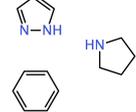
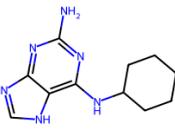
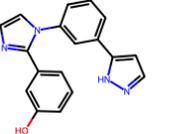
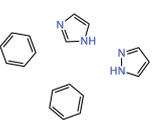
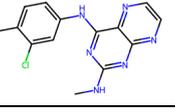
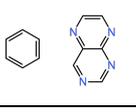
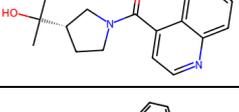
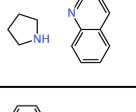
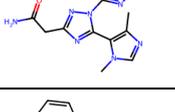
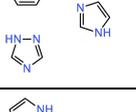
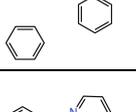
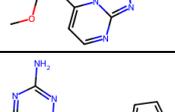
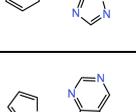
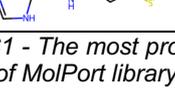
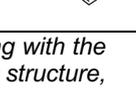
	MolPort-019-822-148	<chem>CCc(cccc1)c1N(C(Cc(cc1)ccc1O)=NN1)C1=O</chem>	-8,743	3,209	
	MolPort-015-136-309	<chem>CCN(c(cc(cc1)N)c1Sc1c2cccc1)C2=O</chem>	-8,675	3,110	
	MolPort-046-848-539	<chem>Nc1nc(Nc2cccc3ccc cc23)c2[nH]cnc2n1</chem>	-8,569	2,747	
	MolPort-000-856-140	<chem>Nc(nc1Nc2cccc(C(F)(F)F)c2)nc2c1[nH]cn2</chem>	-8,397	2,401	
	MolPort-002-003-653	<chem>CC(Nc1n[n]2c(-c3cccc3)ccnc2n1)=O</chem>	-8,381	1,083	
	MolPort-046-188-559	<chem>O[C@H](CN(C1)C(c2c[nH]nc2-c2cccc(F)c2)=O)[C@@H]1F</chem>	-8,298	0,897	
	MolPort-006-821-342	<chem>Nc1nc(NC2CCCCC2)c2[nH]cnc2n1</chem>	-8,266	1,225	
	MolPort-020-230-745	<chem>Oc1cc(-c2ncc[n]2-c2cc(-c3ccn[nH]3)ccc2)cc c1</chem>	-8,254	2,223	
	MolPort-003-119-795	<chem>Cc(ccc(Nc1nc(NC)n c2nccnc12)c1)c1Cl</chem>	-8,172	2,665	
	MolPort-035-856-148	<chem>CC(C)([C@@H](CC1)CN1C(c1ccnc2ccc c12)=O)O</chem>	-8,139	2,554	
	MolPort-047-716-206	<chem>Cc1c(-c2nc(CC(N)=O)n[n]2-c2ncccc2)[n](C)cn1</chem>	-8,119	0,544	
	MolPort-020-217-535	<chem>OCCc(cccc1)c1-[n]1c(-c2cccc(F)c2F)ncc1</chem>	-8,113	2,960	
	MolPort-028-913-593	<chem>CC(C)c1n[n]2c(-c(ccc(F)c3)c3OC)cc nc2n1</chem>	-8,090	2,117	
	MolPort-006-808-255	<chem>Nc(nc1NCCc2ccc[s]2)nc2c1[nH]cn2</chem>	-8,056	1,384	

Table S1 - The most promising docking results for CaBdf 1 BD1 were identified by molecular docking with the subset of MolPort library. In the Table, the following information is contained: compound number, 2D structure,

MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.

Table S2 shows the most promising potential docking hits for the CaBdf 1 BD1 target that were identified through molecular docking studies conducted with the MolPort natural products library. Each hit includes its 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, and ring fragments.

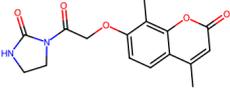
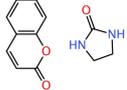
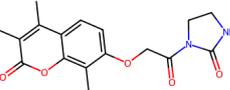
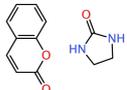
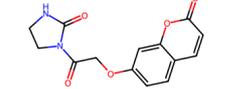
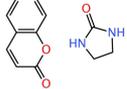
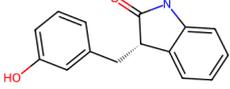
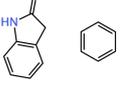
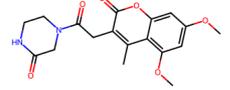
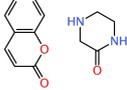
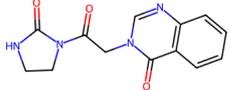
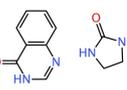
Structure	MolPort-ID	SMILES	Docking score	cLogP	Ring fragments
	MolPort-009-763-100	<chem>CC(c(ccc(OC C(N(CCN1)C1=O)=O)c1C)c1O1)=CC1=O</chem>	-8,234	1,039	
	MolPort-009-762-489	<chem>Cc(c(O1)c(cc2)C(C)=C(C)C1=O)c2OCC(N(CCN1)C1=O)=O</chem>	-8,201	1,479	
	MolPort-008-348-806	<chem>O=C(COC(cc1)cc(O2)c1C=CC2=O)N(CCN1)C1=O</chem>	-8,332	0,617	
	MolPort-008-348-689	<chem>Oc1cccc(C[C@@H](c(ccc2)c2N2)C2=O)c1</chem>	-9,198	2,544	
	MolPort-005-909-990	<chem>CC(c(c(OC)cc(OC)c1)c1O1)=C(CC(N(CCN2)CC2=O)=O)C1=O</chem>	-8,447	0,533	
	MolPort-005-330-758	<chem>O=C(CN(C=Nc1c2c ccc1)C2=O)N(CCN1)C1=O</chem>	-8,491	0,003	

Table S2 - The most promising docking hits for CaBdf 1 BD1 were identified by molecular docking with the MolPort natural products library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.

Summary of all predicted hits for the target *Candida albicans* bromodomain factor 1 bromodomain 2

Table S3 shows the most promising potential docking hits for the CaBdf 1 BD2 target that were identified through molecular docking studies conducted with a subset of the MolPort library. Each hit includes its 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, and ring fragments.

Structure	MolPort-ID	SMILES	Docking score	cLogP	Ring fragments
	MolPort-038-415-941	<chem>Nc1nc([C@H](C2)C[C@H]2O)cc(NC(C(F)(F)F)n1)</chem>	-10,002	1,332	
	MolPort-001-026-787	<chem>NC(c1ccccc1)c1NC(Nc1ccccc1)=S=O</chem>	-9,890	2,285	
	MolPort-007-611-328	<chem>Cc(cc1)ccc1Nc1c(-c(cc2)ccc2O)nc2[n]1cccc2</chem>	-9,706	4,232	
	MolPort-002-978-614	<chem>Oc(cc1)ccc1-c1c(Nc2ccccc2)[n](ccnc2)c2n1</chem>	-9,599	2,888	
	MolPort-002-979-075	<chem>Cc1cccc(Nc2c(-c(cc3)ccc3O)nc3ncc[n]23)c1</chem>	-9,356	3,491	
	MolPort-002-511-602	<chem>NC(c1cc(-c(cc2)ccc2O)nc2ccc12)=O</chem>	-9,325	2,467	
	MolPort-035-773-380	<chem>CN(C)c(nc(cc1C(F)(F)F)c(cc2)ccc2O)c1Cl</chem>	-9,204	3,765	
	MolPort-001-740-277	<chem>O[C@@H](C1)[C@@H](c(cc2)cc(O)c2O)Oc2c1c(O)cc(O)c2</chem>	-8,969	1,509	
	MolPort-003-997-631	<chem>Oc(cc1)ccc1-c1c[n](CCS2)c2n1</chem>	-8,908	1,845	
	MolPort-019-801-991	<chem>CC(C)c1nc(C(F)(F)F)nc(-c(cc2)ccc2O)c1</chem>	-8,863	3,480	
	MolPort-003-665-792	<chem>O[C@H](C1)[C@@H](c(cc2)ccc2O)Oc2c1c(O)cc(O)c2</chem>	-8,786	1,854	
	MolPort-003-665-797	<chem>Oc(cc1)ccc1C(Oc1c2ccc(O)c1)=CC2=O</chem>	-8,758	2,681	
	MolPort-000-344-658	<chem>Oc(cc1)ccc1N(C(/C(\S1)=C/c2ccccc2)=O)C1=O</chem>	-8,739	2,892	
	MolPort-000-344-747	<chem>Oc(cc1)ccc1N(C(/C(\S1)=C/c2cnccc2)=O)C1=O</chem>	-8,718	1,891	

	MolPort-047-793-342	<chem>Oc(cc1)ccc1-c1ncc[n]1[C@@H]1CSCC1</chem>	-8,657	1,995	
	MolPort-046-749-795	<chem>OCC1(CO)[C@@H](C2)[C@H]1CN2C(c(ccc(O)c1)c1F)=O</chem>	-8,644	0,562	
	MolPort-000-738-714	<chem>Oc(cc1)ccc1-c1nc2ccccc2nc1</chem>	-8,548	2,502	

Table S3 - The most promising docking hits for CaBdf 1 BD2 were identified by molecular docking with the subset of MolPort library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.

Table S4 shows the most promising potential docking hits for the CaBdf 1 BD2 target that were identified through molecular docking studies conducted with the MolPort natural compound library. Each hit includes its 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, and ring fragments.

Structure	MolPort-ID	SMILES	docking score	cLogP	Ring fragments
	MolPort-001-732-545	<chem>Oc1cc(-c2nc(-c3ccccc3)c(-c3ccccc3)[nH]2)ccc1</chem>	-9,024	4,618	
	MolPort-000-672-474	<chem>Oc1ccc([C@@H](Nc2c3ccccc2)NC3=O)cc1</chem>	-9,022	1,444	
	MolPort-004-065-770	<chem>COc(c(OC)c1)cc2c1N=CN(CC(N)=O)C2=O</chem>	-8,999	-0,594	
	MolPort-009-758-267	<chem>C=C1N(CC(N)=O)CCc2c1cccc2</chem>	-8,965	1,203	
	MolPort-000-672-474	<chem>Oc1ccc([C@H](Nc2c3ccccc2)NC3=O)cc1</chem>	-8,935	1,444	
	MolPort-001-759-343	<chem>NC(Cc1c[nH]c2c1ccc2)=O</chem>	-8,927	0,784	
	MolPort-000-005-166	<chem>[O-][C](c(cc1)cc(N2)c1N(CCC1)[C@@H]1C2=O)=O</chem>	-8,890	-1,731	

Table S4 - The most promising docking hits for CaBdf 1 BD2 were identified by molecular docking with the MolPort natural compound library. The Table contains the following information: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.

Summary of all predicted hits for the target *Candida auris* bromodomain

Table S5 shows the most promising potential docking hits for the *C. auris* BD that were identified through molecular docking studies conducted with the MolPort natural compound library. Each hit includes its 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, and ring fragments.

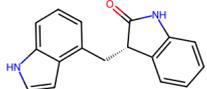
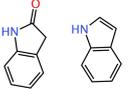
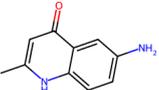
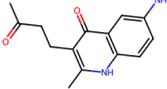
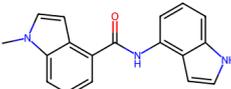
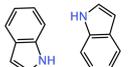
Structure	MolPort-ID	SMILES	docking score	cLogP	Ring fragments
	MolPort-008-348-673	<chem>O=C([C@H]1Cc2c(cc[nH]3)c3ccc2)Nc2c1cccc2</chem>	-9,104	2,929	
	MolPort-000-840-542	<chem>CC(Nc(cc1)c2cc1N)=CC2=O</chem>	-8,935	0,851	
	MolPort-002-521-114	<chem>CC(CCC1=C(C)Nc(ccc(N)c2)c2C1=O)=O</chem>	-8,864	1,601	
	MolPort-019-950-253	<chem>C[n]1c2cccc(C(Nc3c(cc[nH]4)c4ccc3)=O)c2cc1</chem>	-8,856	3,019	

Table S5 - The most promising docking hits for *C. auris* were identified by molecular docking with the MolPort natural compound library. In the Table, the following information is contained: compound number, 2D structure, MolPort-ID, docking score, SMILES code, cLogP value, 2D ring fragment structure, and cluster by ring fragments.