# Proteins in harmony: Tuning selectivity in early drug discovery

Burggraaff, L.

# Glossary

## Glossary

| | |
|---|---|
| Affinity | Measure for the strength of substances to combine with each other (e.g. ligand-protein combination). |
| Allosteric | A non-orthosteric binding pocket on a protein. |
| Amino acids | The 'building blocks' of a protein or peptide. |
| Binding pocket | An area on a protein where substances (e.g. ligands, substrates) can bind. |
| Binding pose metadynamics | An automated, enhanced sampling method that assesses the compound stability in the binding pocket. A higher mobility indicates a less stable compound pose. |
| Bioactivity | A measure of effect caused by a compound when bound to a protein. |
| Clustering | Defining subgroups in a larger group (of molecules) based on structural similarity. |
| Compound | A (chemical) substance or small molecule. |
| Cryo-electron microscopy (cryo-EM) | A technique to determine the arrangement of atoms using cooled/frozen samples and electron microscopy. |
| Crystallography | A technique to determine the arrangement of atoms using crystals of the sample. |
| Protein crystallography | The determination of the arrangement and bonding of atoms of a protein using crystallography. |
| X-ray crystallography | A technique to determine the arrangement of atoms using crystals of the sample and X-rays refraction. |
| Descriptors | Identifiers used to describe a compound/protein; often chemical properties. |
| Docking | A computational technique to virtually place compounds into a rigid binding pocket of a protein. |
| Induced-fit docking | Similar to docking, but with (restricted) movement of the binding pocket amino acids enabled. |
| Ensemble scoring | A scoring method that uses multiple values and/or methods to score, e.g., the binding affinity of a compound. |
| Enzyme | A protein that carries out a biochemical reaction. |

| | |
|---|---|
| Homology model | A (protein) model that is based on a similar or related structure. |
| Inhibition | The opposite of activation; the action of inhibiting a process. |
| Ligand<br>Endogenous ligand | A compound that interacts with or binds to a protein.<br>The natural ligand that binds to the protein. |
| Machine learning | A computational technique that applies algorithms and statistical modeling to automatically learn (patterns) from data. |
| Molecular dynamics simulation | A (newtonian) physics-based technique to simulate the movements of atoms and molecules. |
| Orthosteric | The binding pocket on a protein where the endogenous ligand or substrate binds. |
| Pathophysiology | The disordered physiological processes associated with disease or injury. |
| Polypharmacology | Compounds that act on multiple proteins. |
| Protein | Large molecules composed of chains of amino acids of which the structure is encoded in the DNA. |
| Protein expression | The biologic synthesis of proteins. |
| Proteochemometrics (PCM) | Supervised machine learning of bioactivity based on explicit molecular and target descriptors (typically proteins). |
| Quantitative structure-activity relationship (QSAR) | The relation between the chemical structure of a compound and its activity on a protein. |
| Random forest | An ensemble learning method that applies multiple unpruned decision trees. |
| Receptor | A membrane bound protein that can activate second messengers after ligand recognition. |
| Selectivity | The property of affecting (e.g. inhibiting) a specific protein, while not affecting others at a similar level. |
| Statistical modeling | A modeling technique based on mathematics. The model is defined by an equation that approximates the data. |

| | |
|---|---|
| Structural protein-ligand interaction fingerprints (SPLIFs)[i] | An interaction fingerprint that explicitely encodes the interaction between ligand and protein in a structural way and therefore implicitly captures all types of ligand–protein interactions. |
| Substrate | A molecule that is used by enzymes in catalysis. |
| Target | A protein considered in ligand binding. |
| Anti-target | A protein to which ligand binding should be avoided. |
| Off-target | A protein that will result in undesired effects upon ligand binding. |
| On-target | A protein that will result in desired effects upon ligand binding. |
| T-Distributed Stochastic Neighbor Embedding (t-SNE) | A non-linear technique for dimensionality reduction to visualize high-dimensional data. |
| Text mining | Deriving data from text by automatically searching the text using the computer. |
| Transporter | A protein that transports substances across the cell membrane. |
| Virtual screening | A computational way of testing compounds against a protein. |

i Da, C.; Kireev, D. Structural Protein-Ligand Interaction Fingerprints (SPLIF) for Structure-Based Virtual Screening: Method and Benchmark Study. J. Chem. Inf. Model. 2014, 54, 2555-2561.

# List of Publications

## List of Publications

Burggraaff, L.; Oranje, P.; Gouka, R.; van der Pijl, P.; Geldof, M.; van Vlijmen, H. W. T.; IJzerman, A. P.; van Westen, G. J. P. *Identification of Novel Small Molecule Inhibitors for Solute Carrier SGLT1 Using Proteochemometric Modeling*. Journal of Cheminformatics **2019**, 11, 15. https://doi.org/10.1186/s13321-019-0337-8.

Burggraaff, L.; van Vlijmen, H. W. T.; IJzerman, A. P.; van Westen, G. J. P. *Quantitative Prediction of Selectivity between the $A_1$ and $A_{2A}$ Adenosine Receptors*. Journal of Cheminformatics **2020**, 12, 33. https://doi.org/10.1186/s13321-020-00438-3

Burggraaff, L.; Lenselink, E. B.; Jespers, W.; van Engelen, J.; Bongers, B. J.; Gorostiola González, M.; Liu, R.; Hoos, H. H.; van Vlijmen, H. W. T.; IJzerman, A. P.; van Westen, G. J. P. *Successive Statistical and Structure-Based Modeling to Identify Chemically Novel Kinase Inhibitors*. Journal of Chemical Information and Modeling **2020**. https://doi.org/10.1021/acs.jcim.9b01204.

Burggraaff, L.; van Veen, A.; Lam, C. C.; van Vlijmen, H. W. T.; IJzerman, A. P.; van Westen, G. J. P. *Annotation of allosteric compounds to enhance bioactivity modeling for class A GPCRs*. submitted.

Sydow, D.*; Burggraaff, L.*; Szengel, A.; van Vlijmen, H. W. T.; IJzerman, A. P.; van Westen, G. J. P.; Volkamer, A. *Advances and Challenges in Computational Target Prediction*. Journal of Chemical Information and Modeling **2019**. https://doi.org/10.1021/acs.jcim.8b00832.
*these authors contributed equally

Oranje, P.; Gouka, R.; Burggraaff, L.; Vermeer, M.; Chalet, C.; Duchateau, G.; van der Pijl, P.; Geldof, M.; de Roo, N.; Clauwaert, F.; Vanpaeschen, T.; Nicolaï, J.; de Bruyn, T.; Annaert, P.; IJzerman, A. P.; van Westen, G. J. P. *Novel Natural and Synthetic Inhibitors of Solute Carriers SGLT1 and SGLT2*. Pharmacology Research & Perspectives **2019**, 7, e00504. https://doi.org/10.1002/prp2.504.

Hameed, D. S.; Sapmaz, A.; Burggraaff, L.; Amore, A.; Slingerland, C. J.; van Westen, G. J. P.; Ovaa, H. *Development of Ubiquitin-Based Probe for Metalloprotease Deubiquitinases*. Angewandte Chemie International Edition **2019**, 58, 14477–14482. https://doi.org/10.1002/anie.201906790.

Ruano-Ordás, D.; Burggraaff, L.; Liu, R.; van der Horst, C.; Heitman, L. H.; Emmerich, M. T. M.; Mendez, J. R.; Yevseyeva, I.; van Westen, G. J. P. *A Multiple Classifier System Identifies Novel Cannabinoid CB2 Receptor Ligands*. Journal of Cheminformatics **2019**, 11, 66. https://doi.org/10.1186/s13321-019-0389-9.

Koenders, S. T. A.; Wijaya, L. S.; Erkelens, M. N.; Bakker, A. T.; van der Noord, V. E.; van Rooden, E. J.; Burggraaff, L.; Putter, P. C.; Botter, E.; Wals, K.; van den Elst, H.; den Dulk, H.; Florea, B. I.; van de Water, B.; van Westen, G. J. P.; Mebius, R. E.; Overkleeft, H. S.; Le Dévédec, S. E.; van der Stelt, M. *Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells*. ACS Central Science **2019**. https://doi.org/10.1021/acscentsci.9b01022.

Yang, X.; Dilweg, M.; Osemwengie, D.; Burggraaff, L.; Es, D.; Heitman, L.; IJzerman, A. *Design, Synthesis and Pharmacological Profile of LUF7746, a Novel Covalent Partial Agonist for the Adenosine $A_1$ Receptor*. **2020**. https://doi.org/10.26434/chemrxiv.11903286.

Zhou, J.; Mock, E. D.; Martella, A.; Kantae, V.; Di, X.; Burggraaff, L.; Baggelaar, M. P.; Al-Ayed, K.; Bakker, A.; Florea, B. I.; Grimm, S. H.; den Dulk, H.; Li, C. T.; Mulder, L.; Overkleeft, H. S.; Hankemeier, T.; van Westen, G. J. P.; van der Stelt, M. *Activity-Based Protein Profiling Identifies α-Ketoamides as Inhibitors for Phospholipase A2 Group XVI*. ACS Chemical Biology **2019**. https://doi.org/10.1021/acschembio.8b00969.

Zhou, J.; Mock, E. D.; Al Ayed, K.; Di, X.; Kantae, V.; Burggraaff, L.; Stevens, F.; Martella, A.; Mohr, F.; Jiang, M.; van der Wel, T.; Wendel, T. J.; Ofman, T.; Tran, Y.; de Koster, N.; van Westen, G. J. P.; Hankemeier, T.; van der Stelt, M. *Structure-Activity Relationship Studies of α-Ketoamides as Inhibitors of the Phospholipase A and Acyltransferase (PLAAT) Enzyme Family*. Journal of Medicinal Chemistry **2020**. https://doi.org/10.1021/acs.jmedchem.0c00522.

Mladic, M.; de Waal, T.; Burggraaff, L.; Slagboom, J.; Somsen, G. W.; Niessen, W. M. A.; Manjunatha Kini, R.; Kool, J. *Rapid Screening and Identification of ACE Inhibitors in Snake Venoms Using At-Line Nanofractionation LC-MS*. Analytical and Bioanalytical Chemistry **2017**, 409, 5987–5997. https://doi.org/10.1007/s00216-017-0531-3.

# Curriculum Vitae

## Curriculum Vitae

Lindsey Burggraaff was born on the 17th of November 1990 in Purmerend, The Netherlands. After graduating from pre-university education at the Da Vinci College in Purmerend in 2010, she started her study in pharmaceutical sciences at the VU in Amsterdam. After obtaining her bachelor degree, she continued with her master in Drug Discovery and Safety. She completed the curriculum of two specializations: "Drug Discovery and Target Finding" and "Computational Medicinal Chemistry and Toxicology". Her master studies were finalized with an internship in the computational medicinal chemistry group where she modeled selectivity of compounds between human and viral phosphodiesterases in the framework of medicines against neglected parasitic diseases.

In 2015 she pursued her interest in computational modeling by performing a PhD at Leiden University, during which she became an expert in data science and 3D modeling of ligand-protein interactions. Using machine learning, she successfully modeled glucose transporter proteins for a project with Unilever. She provided computational expertise to different fields of research and closely collaborated with multiple institutes including: the department of Cell and Chemical Biology, Oncode Institute, Leiden University Medical Center; the Institute of Physiology, Charité Universitatsmedizin, Berlin, Germany; and Karlsruhe Institute of Technology, Karlsruhe, Germany. Moreover, she visited the Cell and Molecular Biology group in Uppsala, Sweden, as guest researcher to master additional 3D modeling techniques.

Lindsey frequently presented her work at national and international conferences. She presented at the Centre for Computational Life Sciences, Leiden; EuroQSAR, Verona, Italy; ULLA, Helsinki, Finland; and the FIGON Dutch Medicines Days, Ede, of which the latter was awarded with the best poster award. Moreover, she obtained a travel grant to support her attendance at the Eight Joint Sheffield Conference on Chemoinformatics in Sheffield, UK, where she presented her research to experts in the field. Next to scientific outreach, she focused on conveying her research to a broader audience - the general public. With her participation in FameLab, a renowned international science communication competition, she reached the national final with a pitch that was praised for its clarity. Furthermore, she presented her research to a laymen audience at "the night of discoveries", a yearly cultural event in Leiden that showcases the wonders of science to thousands of visitors.

She continues her career in digital acceleration and currently takes on new data-driven challenges in industry.

# Acknowledgements

## Acknowledgements

This thesis is a result of four years of research, during which I met many inspiring people and gratefully experienced the compassion of many colleagues and peers. I would like to thank everyone who contributed to this thesis in any way, be it scientifically or socially.

Thank you, Ad and Herman for your guidance, advice and kindness during my research – I am grateful to have had you as my promotors. Ad, I appreciate our scientific and social interactions. Herman, thank you for your critical eye and your refreshing opinions from an industry point of view.

Gerard, my co-promotor, thank you for always being there to provide me with feedback on my scientific work. I am thankful for the enjoyable social activities we shared. You made me feel trusted and appreciated.

Thank you to all my collaborators from the other faculties, industry and institutes, for providing me with the opportunity to broaden my perspective on the applicability of computational methods in other research fields. Michael, thank you for you advice on computer science matters.

Furthermore, I would like to thank all my (former) colleagues who contributed to the pleasant working environment. Especially my computational colleagues Bart, Willem, Shardul and Olivier. Bart, the last time I checked there was still a physical trace at our office wall reminding me of your 4 o'clock grapefruit time. May it live on through many generations of PhDs. Willem, you are awesome and you know it! I want to point out to Shardul that the key you broke is still blamed on me. Being a boxer has not been in my favour, especially when the window spontaneously cracked shortly after (Ad, I swear it was just coincidental). Nevertheless, you are my friend in business and beyond. Olivier, you are truly one of a kind and the most soulful dancer I have ever met; you are glorious.

Thank you lab-ladies Anna, Xue, Natalia, Rongfang, Xuesong and Tasia, who made coffee breaks interestingly fun and with whom I experienced many social events outside of working hours. We had a lot of fun together and I will definitely miss working with you.

To the students who performed their research internships under my supervision: thank you for your hard work and for making me the supervisor I am today. This goes especially for my master students Amber, Jenke, Florian, Marina and Martijn. Although your personalities could not differ more, I enjoyed supervising every single one of you. You are great minds and I surely believe you will do great in your future careers.

My paranymphs, Nadine and Willem, thank you for your support and for being by my side on such a special day. Nadine, thank you for being such a dear friend over the past years, and for the many more years to come! Willem, you were a valued office neighbour and are an even more valued friend. With you, laughter can always be found right around the corner.

Special thanks to my parents, Hans and Ellen, of whom I received unconditional love and support. Sandy, Dana, Narita, Bradley and Rozemarijn, I am grateful for the bond we have - I can praise myself lucky with such amazing siblings.