UNVEILING RECEPTOR-PEPTIDE COMPLEXES USING ALPHAFOLD AND DEEP LEARNING

HANS CARLO VON VEGESACK / 202303410

Master Thesis

Advisor: Palle Villesen, Alexander Sebastian Hauser

Department of Bioinformatics
Faculty of Natural Sciences
Aarhus University

June 2025

Hans Carlo von Vegesack: *Unveiling Receptor-Peptide Complexes using AlphaFold and Deep Learning*, Master Thesis © June 2025

Recent advancements in deep learning (DL) have significantly improved our ability to predict protein structures with atomic accuracy. With the introduction of AlphaFold Multimer, it became possible to predict protein complexes involving multiple chains, opening new avenues in structural bioinformatics. Building on these developments, this thesis explores how DL models can be utilised to identify true interactions between G-protein-coupled receptors (GPCRs), the largest family of therapeutic targets, and their associated peptide hormones. We systematically benchmarked two state-of-the-art DL models, AlphaFold 3 and Chai-1, on a dataset designed to mimic realistic screening conditions. Model performance was evaluated using a first-pick classifier approach, where AlphaFold 3 and Chai-1 achieved AUC values of 0.83 and 0.77, respectively. Further analysis revealed that combining structural confidence metrics through forward feature selection and a random forest classifier significantly improved binder recovery. These findings demonstrate that DL models can reliably rediscover peptide binders and support peptide drug discovery. Moreover, they highlight the potential of tailored post-processing strategies to enhance DL predictions and streamline ligand screening workflows.

CONTENTS