universität wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

“Investigation Of The Interface Between Human Oxytocin Receptor And Human Dopamine D2 Receptor”

verfasst von / submitted by

Julian Gabler BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magister der Pharmazie (Mag.pharm.)

Wien, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 449
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Diplomstudium Pharmazie
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Thierry Langer
Mitbetreut von / Co-Supervisor:	Dr. Marcus Wieder, MSc MSc

TABLE OF CONTENTS

1. INTRODUCTION

Biomolecular interactions are fundamental for both metabolic and regulatory processes of life. Understanding these interactions in detail can be crucial for identifying bioactive substances to alter those processes. The knowledge of how those individual molecular interactions combine can lead to a more global understanding of processes in living organisms. For interaction analysis, the three-dimensional structural information of the molecules involved is required (Lengauer and Rarey, 1996).

It has been shown that heterocomplexes of G-protein coupled receptors (GPCRs) are involved in a variety of neurophysiological processes in the central nervous system (CNS), such as long-term memory (Fuxe et al., 2014) and social bonding (De la Mora et al., 2016). On the other hand, dysfunction of such complexes may lead to brain disorders, qualifying them as drug targets, e.g., for the therapy of Parkinson's disease, schizophrenia, anxiety and depression (Borroto-Escuela et al., 2017; De la Mora et al., 2016). Although many GPCR heterodimers have been described, it remains difficult to determine their physiological role, mainly because of the challenge to differentiate receptor crosstalk from direct protein-protein interactions in *in vivo* experiments (Salahpour and Caron, 2012). Allosteric receptor-receptor interactions taking place in GPCR heterocomplexes can affect ligand binding, signaling, mobility within the neuronal membrane or lead to the emergence of new allosteric binding sites (Borroto-Escuela et al., 2017).

Dopamine as a neurotransmitter has a variety of effects on the CNS, involving cognition, emotion, perception, motivation, reward and sleep, whereas disruption of the dopamine system may lead to autism, depression, anxiety, Parkinson's disease and schizophrenia (Baskerville and Douglas, 2010). The neuropeptide oxytocin has an important role in social attachment, affiliation and sexual behavior (Gimpl and Fahrenholz, 2001). Interactions of the central dopamine and oxytocin systems have been proposed, since disturbances of oxytocin levels in patients

with dopamine dependent disorders have been detected (Baskerville and Douglas, 2010).

This work focusses on the GPCR heterocomplex formed by oxytocin receptor and dopamine D2 receptor. It is known that oxytocin receptor (OTR) and dopamine D2 receptor (D2R) are able to form heteromer complexes. Facilitatory allosteric effects lead to reciprocal enhanced signaling of both D2R and OTR pathways upon ligand activation. It is assumed that OTR-D2R complexes play a role in social bonding (Romero-Fernandez et al., 2012). As disruption or dysfunction of those heteromers may play a role in the neuropathology of anxiety (De la Mora et al., 2016) and schizophrenia (Borroto-Escuela et al., 2017), these complexes are suitable drug targets. In the treatment of schizophrenia, the blockade of D2R in D2R-OTR heterocomplexes by D2R antagonistic drugs may lead to the failure of D2R antagonists to reduce negative symptoms of schizophrenia, such as social interaction deficits. This issue might be addressed by the development of a treatment strategy with agonistic heterobivalent drugs, specifically targeting these heterocomplexes. These compounds would comprise two pharmacophores, one targeting OTR and one D2R, both attached to a linker in appropriate length, connecting the orthosteric binding sites of the receptors. Heterobivalent ligands are also powerful tool for studying the quaternary structure of GPCR dimers and their functional relevance (Borroto-Escuela et al., 2017; Hübner et al., 2016).

Important methods to study heteroreceptor complexes in cells and tissue are fluorescence resonant energy transfer (FRET), bioluminescence energy resonance transfer (BRET) and *in situ* proximity ligation assay (Fuxe et al., 2014). These techniques primarily demonstrate the existence of those complexes rather than revealing structural details of the protein-protein interaction (Gurevich and Gurevich, 2008). There is little information about the structural characteristics of the OTR-D2R complex that would be needed for further neurophysiological and neuropathological research and drug design. Fortunately, the crystal structures of oxytocin receptor (Waltenspühl et al., 2020) and dopamine D2 receptor were released recently (Yin et al., 2020; Wang et al., 2018), opening new opportunities to study interactions of those receptors (Lengauer and Rarey, 1996).

The aim of this work is the modeling of the OTR-D2R heterocomplex via protein-protein docking, followed by investigation of the emerging interfaces. That was put into practice by the design and docking of a heterobivalent linker to the modeled complexes and the subsequent discussion of structural details and stability of the complexes after molecular dynamics simulation in a lipid bilayer environment. The determination of the linker length that is necessary to overcome the distance between the orthosteric ligand binding sites may serve as a basis for *in vitro* and *in vivo* experiments with heterobivalent drugs targeting the OTR-D2R heterocomplex.

Protein-protein docking is an *in silico* method which allows to study protein-protein interactions. It usually consists of a sampling stage, where a number of plausible protein-protein binding conformations is generated, and a scoring stage to rank those conformations (Zhang et al., 2016). Two docking techniques were used to generate candidates: the Rosetta MPDock program in a local docking approach (Alford et al., 2015), and the Memdock server (Hurwitz et al., 2016), which employs a global search for docked structures. After preparation of the single protein structures for the docking process, the generation of start structures for Rosetta MPDock was guided by a scientific literature search for similar complexes of known or supposed interfaces involving D2R or OTR. In agreement with the findings, the receptors were rotated into four orientations and then docked with Rosetta MPDock. For Memdock, prepared single structures of the receptors sufficed. The results were evaluated according to the number of amino acids at the interface to select one complex for further investigation. A bivalent linker was designed by docking small molecule agonists to the best protein complexes found by Rosetta MPDock and Memdock. This was followed by interconnection of the agonists by a linker of defined length. The protein-protein-linker complex was then integrated to a membrane environment and a molecular dynamics simulation was carried out. RMSD analysis of the protein backbone was performed to determine conformational stability of the complex. Further structural analysis was carried out by the PyRosetta Alanine Scan protocol, allowing the identification of residues of the interface and their approximate changes in binding energy upon mutation (Kortemme and Baker, 2002). It is hoped that the gain in knowledge provides the basis for further structural

investigations of the complex as well as for *in vitro* and *in vivo* studies on bivalent ligands that act on the OTR-D2R system, which in turn may eventually lead to development of new drugs targeting the OTR-D2R complex.

2. PROTEIN-PROTEIN INTERACTIONS

Per definition, protein-protein interactions (PPIs) are physical contacts between two proteins that occur in a cell or an organism *in vivo*. The interaction interface should be definable by specific selected biomolecular forces and events, rather than accidental contact. Also, the interaction interface should not be involved in generic functions such as protein synthesis or degradation that concern nearly every protein. Instead, it should be evolved for a specific purpose (De Las Rivas and Fontanillo, 2010).

Protein-protein interactions occur in protein complexes. A protein complex includes two or more proteins, each partner of a complex is called a protomer. Homocomplexes and heterocomplexes are distinguished. Homocomplexes consist of identical protomers and are usually permanent and optimized. Heterocomplexes are made of two different protein partners and they can be non-obligatory, i.e., their formation and separation depend on environmental factors and they also exist independently. Interfaces can be characterized by following properties: size of the interface, complementarity, residue interface propensities, hydrophobicity, hydrogen bonding, secondary structures, and complex changes on complex formation (Jones and Thornton, 1996).

Size of the interface

The size of an interface can be measured in Angstrom^2 (Å²) by the relative change of accessible surface area (∆ASA) on complex formation. It is also known as buried surface area (Khechinashvili et al., 2008). The ∆ASA is known to correlate with the hydrophobic free energy change that occurs during transition from a polar solvent to hydrophobic environment (Jones and Thornton, 1996).

Steric complementarity

The steric complementarity of permanent complexes is usually higher than non-obligate ones. Many interfaces have cavities, they can represent about 10% of the total interface surface (Veselovsky et al, 2002). The complementarity can be measured via defining a gap index, where the gap volume between the molecules is divided by the interface accessible surface area (Jones and Thornton, 1996).

Electrostatics

Electrostatic interactions enable the formation of salt bridges on complementary protein interface surfaces. Electrostatic attraction of opposite charges is possibly involved in long range promotion of complex formation, thus having an influence on complex formation rate and stability (Veselovsky et al., 2002).

Hydrophobicity

Hydrophobic regions of a soluble protein that are exposed to an aqueous medium are energetically unfavorable. That is why these regions are often involved in binding sites where they are shielded by bulk water (Eisenhaber, 1999). Hydrophobicity can be indicated by residue propensities, i.e., the distribution of different amino acids at the interface compared to distribution over the whole protein surface. Hydrophobic residues show a preference for interfaces of homodimeric complexes. Heterocomplexes have a more balanced propensity for polar and hydrophobic residues. This can be explained by the fact that homodimer complex partners rarely occur in their monomeric form, and their hydrophobic surfaces are permanently buried in the complex (Jones and Thornton, 1996).

Hydrogen Bonds

Hydrogen bonds are associations between a hydrogen atom that is bound to an electronegative atom (donor) and a second relatively electronegative atom with a lone pair of electrons (acceptor). It is primarily considered an electrostatic force acting between the hydrogen and the lone pair of the acceptor (Arunan et al., 2011). Hydrogen bonds are also present at protein-protein interfaces, usually one can find one hydrogen bond every 100-200 Å² of interface surface area. Amino acid side chains are responsible for 76% of the interchain hydrogen bonds. Water molecules can surround contacting surfaces, or they can be located in interface cavities, where they are able to form often highly coordinated hydrogen bond networks that are highly coordinated. Protein-protein binding involves partial desolvation of contacted surfaces (Veselovsky et al., 2002).

Conformational changes

Proteins usually change their conformation upon complex formation. These changes can affect side chains alone, segment movement involving the main chain, and domain movements (Jones and Thornton, 1996). The conformational changes of GPCRs upon dimerization may affect G-protein signaling or selectivity, internalization regulation and ligand binding cooperativity. Positive and negative cooperativity mean that, once one of the partners has bound to a ligand, the binding affinity of the other partner to the same ligand increases or decreases (Terrillon and Bouvier, 2004). Complex formation can also have allosteric effects, meaning that affinity of ligand binding can be altered by conformational change upon protein-protein interaction (Borroto-Escuela et al., 2018).

Membrane-specific properties

Membrane protein-protein interactions are also driven by the properties of their transmembrane domains, which play active roles in the oligomerization of membrane proteins. There are conserved motifs in transmembrane domains that are known to mediate protein-protein interactions. Interhelical interactions can be carried out by complementarity, maximizing the Van der Waals contacts count (Fink et al., 2011). Van der Waals forces occur when two molecules or atoms are in close proximity to each other, where fluctuations of electric charges can lead to weak attraction. If the distance between the partners is too small, repulsion will occur (Bruce Alberts et al., 1998). Hydrogen bonds can act as stabilizers, especially if they are buried in the membrane (Fink et al., 2011).

3. ABOUT THE PROTEINS OF INTEREST

3.1 G protein-coupled receptors

G protein coupled receptors (GPCRs) are membrane proteins that are involved in a variety of cellular physiological responses to stimuli, including hormones, neurotransmitters, and others. Nearly half of all prescribed drugs worldwide target members of the GPCR family, which comprises approximately 1000 members (Ritter and Hall, 2009). GPCRs can be divided into three subfamilies based on sequence similarity. The largest of them is the rhodopsin-like subfamily, which also includes dopamine D2 receptor and oxytocin receptor (Cho et al., 2012; York et al., 2017). All of them share a conserved seven-transmembrane alpha helix topology. The seven transmembrane helices (TMs) are connected by intracellular and extracellular loops (ICLs and ECLs). ICL3 varies widely in length and sequence among all GPCRs, ranging from 20 to 165 amino acids (AA). C- and N-termini may contain receptor-specific domains. Extracellular agonist binding induces conformational changes in the GPCR's transmembrane and intracellular domains, activating the receptor. This enables interaction with heterotrimeric G-Proteins, which are anchored to the inner plasma membrane. The activated receptor now catalyzes the release of GDP and binding of GTP to the G-Protein's Gα subunit, acting as a so-called guanine exchange factor. Additionally, activated GPCRs can associate with G protein coupled receptor kinases (GRKs), leading to GPCR phosphorylation and decreased interaction with the G protein. GPCR signaling can also be altered by arrestin binding. This turns off G protein signaling, while other, arrestin-mediated signal pathways are activated (Ritter & Hall, 2009; Gurevich & Gurevich, 2008).

3.2 Oxytocin receptor

The oxytocin receptor (OTR) belongs to the G-protein coupled receptor family. It acts as a receptor for the nonapeptide oxytocin, which acts as a hormone for uterine smooth muscle contraction during childbirth, is involved in lactation and male ejaculation. It is also playing a role as a neurotransmitter in the brain and is involved in complex social behavior, maternal care, stress and anxiety. OTR is a

potential therapeutic target for the treatment of a variety of mental diseases, e.g., schizophrenia, social anxiety disorders and autism. To date, oxytocin is administered to induce labor, as the antagonist atosiban is used to prevent preterm labor (Koehbach et al. 2013, Waltenspühl et al., 2020).

In 2020, the OTR crystal structure was determined by Wattenspühl et al. It was co-crystallized with the OTR-selective antagonist retosiban. Eight mutations were introduced by directed evolution in *Saccharomyces cerevisiae*, furthermore the 34-residue intracellular loop 3 (ICL3) was replaced with thermostable glycogen synthase from *Pyrococcus abyssi* to allow for crystallization in lipidic cubic phase (LCP). The flexible C-terminus (residues 360 to 389) were truncated, also D153 and S224 were substituted with alanine for increased thermostability. The crystals achieved a resolution of 3.2 Angstrom. A cholesterol molecule binding between helices IV and V that might be responsible for an allosteric effect on OTR function was also observed (Waltenspühl et al., 2020).

3.3 Dopamine D2 receptor

The Dopamine D2 receptor (D2R) is a G-protein coupled receptor that is activated by the neurotransmitter dopamine, or synthetic agonists such as bromocriptine. Dopamine is involved in multiple physiological and cognitive functions, e.g., movement, reward, and emotion. The D2R signaling pathway can be impaired in Parkinson's disease, where the damage of dopaminergic neurons in combination to a deficit of dopamine leads to psychiatric and motoric deficits. There are five known GPCR dopamine receptors, D1-D5, and their physiological functions overlap partially. D1 and D5 are highly homologous in terms of identity and similarity, also D2, D3 and D4 are closely related to each other. However, the dopamine binding residues are highly conserved among all of them (Yin et al., 2020).

In 2018, Wang et al. found a crystal structure for D2R with bound risperidone via x-ray diffraction experiments in their paper "*Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone*". The T4 Lysozyme residues 2-161 were cloned into the intracellular loop3 of the structure, replacing D2Rs residues 223-361. The N-terminus was truncated (residues 1-34).

Thermostabilization mutations were introduced. The model represents the inactive form of the receptor (Wang et al., 2018).

In 2020, Yin et al. determined the structure of D2R in complex with its G-protein in a lipid membrane via cryo electron microscopy. A resolution of 3.7 Angstrom was achieved. For expression, five point mutations were introduced to the wild type gene, also residues 1-29 were replaced with T4 lysozyme and a two residue alanine linker. The flexible, 120 amino acid intracellular loop 3 (ICL3) was replaced with the residues LVNTN (Yin et al., 2020). The modifications were done for crystallization and to stabilize the active form of the receptor. The D2R model created in this paper was used for the docking experiments in this work.

3.4 OTR-D2R heteromers

There is evidence for the existence of oxytocin receptor (OTR) dopamine D2 receptor (D2R) heteromer complexes located in the striatum of rat. It was experimentally determined via bioluminescence resonance energy transfer (BRET) and *in situ* proximity ligation assay. The results indicated facilitatory allosteric receptor-receptor interaction: oxytocin may make the orthosteric binding site of D2R more available for binding, increasing the affinity in the high affinity state of the receptor. Moreover, oxytocin increased the coupling of D2R to its G-protein (Romero-Fernandez et al., 2012). Activation by oxytocin furthermore leads to enhanced MAPK and PKC signaling and inhibition of the CREB pathway, whereas D2R stimulation enhances oxytocin activation and Calcineurin pathway. It is assumed that facilitatory OTR-D2R interactions are involved in pair bonding and social behavior, while lack of their interaction in the central amygdala of rat may lead to anxiety via dysfunction or disruption of OTR-D2R heterocomplexes (De la Mora et al., 2016).

3.5 Structure of GPCR Dimers

The aim of this work is to investigate the interface that forms in the complex of two different GPCRs, making the complex a GPCR heterodimer. Rosetta MP docking does not *per se* provide a global search approach that is able to seek out the global minimum of the complex. For that reason, a literature search was carried out with the goal of finding plausible interfaces for the oxytocin receptor /

dopamine D2 receptor complex. Papers examining the interfaces of homo- or heterodimers of GPCRs similar to D2R or OTR were included. The outlines of these papers were used as a basis to decide which start positions should be chosen for the Rosetta MP docking runs. Due to the high structural conservation between the D2R and OTR seven helix topology, known interfaces of one of the receptors were considered relevant for both.

OTR Homodimers

In the paper *"Design and Characterization of Superpotent Bivalent Ligands Targeting Oxytocin Receptor Dimers via a Channel-Like Structure"* by Busnelli et al. from 2016, homobivalent oxytocin analogs were used to bind dimerized oxytocin receptors, resulting in a boost of the G-protein signaling *in vitro* and *in vivo*. The ligand consisted of two molecules of the oxytocin receptor agonist and oxytocin analog dOTK, connected by linkers short aliphatic linkers with a length ~25 Angstrom. The group used models of κ-opioid receptor (KOR) and CXCR4 dimers as a template for homology modeling of the complex. KOR dimers are known to interact via transmembrane helix (TM) 1 and 2, and their amphiphilic helix 8, CXCR4 dimers via TM5. The OTR dimer-ligand interaction was hypothesized to be based on a TM1-TM2 symmetrical interface with helix 8 participation, since the ligand precisely fits a channel-like passage which is formed by this complex conformation.

Dopamine D1 and D2 receptor heterodimers

In the paper of O'Dowd et al., *"Two amino acids in each of D_1 and D_2 dopamine receptor cytoplasmic regions are involved in D_1-D_2 heteromer formation"* of 2012 a radioligand binding assay was performed to determine the interaction between D1R and D2R. There was no evidence of transmembrane-helical interaction, however two neighboring arginine residues in the ICL3 region of D2R were found to interact with two glutamate residues located in the carboxy tail of D1R. If this interaction was responsible for the OTR-D2R interface, two neighboring negatively charged residues would be necessary in the OTR amino acid sequence. Since they are not present, this interaction is probably negligible for

an OTR-D2R interface. This paper is particularly interesting because it states the existence of dopamine receptor interaction via unordered regions.

Dopamine D2 receptor homodimers

“The dopamine D2 receptor dimer and its interaction with homobivalent antagonists: homology modeling, docking and molecular dynamics” is a paper by Kaczor et al. from 2016 where a TM4/5 - TM1/7 interface is determined for D2R homodimers with Rosetta docking. ICL3 and C-terminus involvement due to other papers is stated, although these regions were not modeled.

Dopamine D2 / Adenosine 2A (A2AR) receptor heterodimers

Borroto-Escuela et al. stated in their paper “*Mapping the Interface of a GPCR Dimer: A Structural Model of the A_{2A} Adenosine and D_2 Dopamine Receptor Heteromer*” that A2AR and D2R form a complex with via TM4 and TM5 interaction of both protomers. Their insights were gained by bioluminescence resonance energy transfer (BRET) assays, where dimerization was blocked by adding TM4 and TM5 peptides of A2AR. Also, protein-protein docking was used to model a complex with a TM4/TM5 interface, revealing participating residues. BRET experiments were in turn carried out where predicted residues of the interface were mutated, and reduced interaction was observed. Also, allosteric effects of A2AR upon D2R agonist binding were reduced. It was stated that ICL3 of D2R is probably close to the TM4/TM5 interface. Also, other interfaces proposed by earlier studies were discussed, involving D2R TM 5, TM6 and TM7. ICL3 was not modeled in this experiment due to lack of reliable templates.

4. MOLECULAR MODELLING

In order for the Rosetta MP-dock protocol to work and deliver meaningful results, it was necessary to make preparations and adaptations to the original protein structures.

Sources of the protein structures

Crystal structures of the active (Yin et al., 2020) and inactive (Wang et al., 2018) conformation of human dopamine D2 receptor, as well as the inactive form of human Oxytocin receptor (Waltenspühl et al. 2020) were obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (available under www.rcsb.org) as coordinate files (subsequently called PDB files). The structure of the active conformation of oxytocin receptor was taken from the supplementary material of the paper from Busnelli et al, 2016.

Restoring mutated and missing amino acids / removing unordered regions

The crystal structures contained many modifications necessary for the crystallization process which had to be edited or deleted before the docking process. The protein-protein docking techniques that will be used have to treat the protomers as rigid bodies. Since unordered regions cannot be represented in a meaningful way in such an approach, they had to be skipped or modified. First, the native amino acid (AA) sequences were downloaded from the Uniprot database (https://www.uniprot.org, UniProt Consortium, 2018), P14416 for D2R, and P30559 for OTR (The Uniprot Consortium, 2020a and 2020b).

The AA sequence of D2R was altered so that the unordered 120 AA ICL3 domain was replaced with a 7 AA fragment (KLVNTNR). In case of the pdbs from the Protein Database, the pdb files were merged with the corresponding AA-sequence by the SWISS-model homology modeling server (https://swissmodel.expasy.org/, Waterhouse et al., 2018).

Since SWISS-MODEL did not find any templates for AA 1-34 and 334-359 in case of OTR 6tpk, 1-36 and D2R 6cm4, for AA 1 - 31 in case of D2R 6vms, these regions could not be modeled. The resulting models were corrected using the original amino acid sequence (except ICL3 and N-termini and C-terminus of OTR).

Determination of protonation state: The PDB2PQR-Server

The PDB2PQR Server is a web service that allows automated reconstruction of missing heavy atoms and hydrogens as well as estimation of protonation states of biomolecules in a way that hydrogen bonding is favored. It can also assign atomic partial charges and radius parameters from various force fields (Dolinsky et al., 2004).

Protein-protein interaction is often associated with changes in protonation states of ionizable residues (Mason & Jensen, 2007). Since these changes cannot be accounted for in the most docking programs including Rosetta, choosing the right protonation state is crucial. The PDB2PQR server was used to assign protonation states at pH 7.0, which were written to the pdb coordinate files.

Determination of membrane orientation: The PPM Server

The PPM Server is a web service that is capable of calculating the rotational and translational position of membrane proteins in membranes (Lomize et al., 2011).

The PDB coordinate files were used as input; the output was a PDB coordinate files with updated positions of the membrane proteins. Also, membrane-dummy atoms were now included, adding information about membrane orientation, thickness, and boundaries.

Structure selection

It was taken into consideration to build models of all four combinations of inactive and active (both active, D2R active – OTR inactive, D2R inactive – OTR active, and both inactive) and then dock all of them, but that would have resulted in excess of computation time. Usage of the inactive structures would have had the advantage of more structural accuracy, since they are both originate from crystal structures, whereas the oxytocin receptor model in its active form was based on

a homology model of β2-adrenergic crystal structure (Busnelli et al., 2016). However, since it is known that the complex formation has enhancing effects on G-protein coupling (Romero-Fernandez et al., 2012), it seems reasonable to assume that the bound conformations of the receptors are in line with the active conformations rather than with the inactive ones. Hence, the active structures were chosen for docking (see chapter “Sources of protein structures”).

Preparations for Rosetta MPDock

The Rosetta MP Docking protocol (Alford et al., 2015) was used to perform the protein-protein docking step. Following inputs are necessary to run it: a span file and “prepacked” PDB coordinate file containing both proteins transformed into membrane coordinates. The Rosetta MPDock protocol will be described in detail in chapter 6.1.

Setting up a start file for Rosetta MPDock via Pymol

“The Pymol Molecular Graphics System” (Schrödinger LLC, 2020) was used to create start files for protein-protein docking. The active forms of OTR and D2R were loaded and, at first, aligned via the “align” function of Pymol. Then, with the help of the dummy membrane atoms of the previous step, the correct alignment of the two proteins was verified. The next step was to translate the D2R receptor by 45 Angstrom along the x-axis of the coordinate system (Fig. 4a).

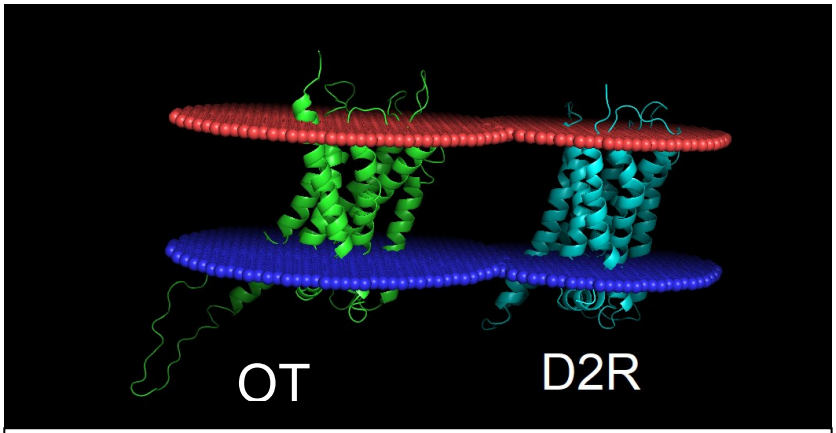


Figure 4a. An alignment of both receptors after translation, ready for rotation. The membrane dummy atoms created by the PPM server were used to orientate the proteins correctly

Transmembrane helix orientation in Pymol

Pymol was used to rotate both protein structures in a way that certain transmembrane helices (TM) were facing each other. According to literature research these orientations were expected to form complexes upon docking, hopefully with those transmembrane regions forming an interface. Following combinations of transmembrane helices were aligned:

- OTR TM 4/5 facing D2R TM 4/5
- OTR TM 1/2 facing D2R TM 4/5
- OTR TM 1/2 facing D2R TM 1/7
- OTR TM 4/5 facing D2R TM 1/7

Within the Rosetta docking protocol, the partners are later brought into proximity again by the “slide into contact”-mover, which, after even increasing the distance, slides on of the proteins until it has contact with the other one.

The span file: determination of intra-membrane regions

The Rosetta MP-Docking protocol demands a “span file”, which is a topology file that is read into Rosetta. It contains information about the number, length and sequence position of the proteins’ transmembrane helices within the membrane.

The Rosetta application “Span from PDB” was used to generate one file for each transmembrane helix orientation. Secondary structure residues with z-coordinates between 15 and -15 are considered to lie within the membrane (Alford et al., 2015). This step requires a membrane coordinate frame, in this case they were created by the PPM server before.

Prepacking

In this step, the side chains of the docking partners are repacked, to make sure that the side chains of all amino acids are in a low energy conformation: the docking partners are moved out of contact, side chains are optimized and finally the original orientation is restored. (Gray et al., 2003).

10 poses and a scoring file were created for each orientation. The pose with the best score, according to the scoring file, was used for the docking run.

5. DOCKING AND SCORING IN THEORY

Generally speaking, docking and scoring are *in silico* methods that are combined to examine two or more molecules and iteratively predict new coordinates for one or more of them, under the restriction that a complex is formed. In the “docking stage”, a search for candidate coordinates is performed. Each new orientation and conformation of all constituents - together referred to as a pose - are then analyzed at the “scoring stage” to evaluate the quality for each candidate structure (Torres et al. 2019, Morris and Lim-Wilby, 2008).

5.1. Molecular docking

Molecular docking refers to a method to generate conformations of small molecules and analyzes the orientation of that molecules in macromolecular binding sites. (Torres et al., 2019). In molecular docking, which is also called protein-ligand docking, the molecules to be docked into the macromolecule are small, mostly organic molecules. This can be useful for drug design, but also for the understanding of biological processes. Databases of ligands can be screened to find individual molecular interactions with the target structure, in order to obtain leads for further drug development (Lengauer and Rarey, 1996; Trott and Olson, 2010).

The target structure is usually a protein, DNA or RNA. Its atomic coordinates can often be accessed by the Protein Data Bank (rcsb.org), there are also “ab initio” and “comparative” modelling techniques to predict the 3D-structure of a protein. The ligand structure can be obtained from small ligand databases or calculated from 2D structures. If unknown, the probable binding region can either be predicted algorithmically, or a computationally expensive “blind docking” simulation can be performed, which considers the whole target structure as a search area in the docking process. For docking calculations, the target binding site is discretized by representing it on a grid with precalculated potential energies for interaction. Then, interactions are calculated for each grid point based on Lennard-Jones and electric potentials. The protonation states and charge assignments of the target amino acids are mostly assumed to remain unchanged

between bound and unbound states. Since charge plays a major role in ligand interactions, setting protonation states in the right way is critical to binding affinity prediction. This can be achieved by software, e.g. PropKa or H++ (Torres et al., 2019).

Molecular docking scoring functions try to approximate standard chemical potentials of the system, with the goal of determining the bound conformation preference and the free energy of binding. In contrast, molecular dynamics simulations use force fields and solvent molecules to reproduce physical energies directly. This means docking scoring functions trade their representational detail for computational speed. (Trott and Olson, 2010).

Not only minima in the energy landscape, but also its shape and temperature are considered in molecular docking scoring functions. When used, superficially physics-based terms of the scoring function need to be empirically weighted. This might account for e.g., the entropic penalty in binding energy that is inflicted because of the loss of degrees of freedom upon binding. Another solution for estimating entropic contributions would be to consider the loss of accessible rotamers upon binding (Trott and Olson 2010 p. 2; Chang et al., 2006).

However, many challenges remain in the field of molecular docking since often neither true ligands among a set of molecules nor the correct ligand conformation are identified correctly (Torres et al., 2019).

5.2 Protein-protein docking

Protein-protein docking is an *in silico* technique that aims to predict the three-dimensional structure of a complex of at least two proteins A and B with given structures. Protein A is often referred to as the receptor, Protein B as the ligand. Based on the notion of steric and physicochemical complementarity at the interface it can provide knowledge of the protein-protein interactions taking place, which might also serve as a basis for drug design (Vakser, 2014).

The process of protein-protein docking consists of a docking and a ranking stage; the docking stage usually starts with a low-resolution step in which the ligand protein is translated and rotated to get the relative binding orientations. Then, a

more local, high resolution step is employed for refinement of the first step, some approaches allowing for flexible moves of side chains, backbone, or both, especially of the interface to minimize the energy of the complex. In the ranking stage, a scoring function is applied to the structures sampled in the docking stage, generating a ranked list of complexes (Zhang et al., 2016).

Since Rosetta dock is an important element of the Rosetta membrane protein-protein docking protocol, which was used for this work, the Rosetta Dock stages will be compared with other popular methods.

5.2.1 Low resolution docking stage

The concept of rigid body docking

The first step of protein-protein docking is the search for relative orientations of the binding partners in the complex. It is usually based on the “rigid body” assumption, meaning that both proteins are treated as solid bodies without flexibility during this step. (Vakser, 2014, Zhang et al. 2016).

Rosetta Dock: Monte Carlo search algorithm

At first, for initial perturbation translational and rotational moves of the ligand are generated randomly. In the next step, probability models decide whether to accept or to reject this move. In the low-resolution docking stage, Rosetta Dock makes use of a 500-step Monte Carlo search that is adjusted dynamically to reach an acceptance rate of 25% (Zhang et al., 2016).

At this point, Rosetta Dock makes use of centroids, simplifying the side chains to unified pseudo atoms. The docking function consists of a “bump” term, a contact term, a docking-specific statistical environment, and residue-residue pairwise potentials (Chaudhury et al., 2011). According to Kozakov et al, 2017, this method is best suited for localized searches over selected regions.

Global systematic search

If no *a priori* information of the complex is available, global systematic search is the most powerful and most popular approach. With this method the whole six dimensional conformational space of the complex is sampled. The ligand protein is rotated and translated each over three dimensions on a discretized grid. If the

search is exhaustive, billions of structures will be generated. Since this is computationally expensive, finding meaningful structures is accelerated by search algorithms. For example, FFT-based geometric algorithms systematically evaluate the geometric fit between two proteins using correlation type scoring functions, optimized by the Fast Fourier Transformation method. It allows to reduce the calculation time cost of the shift vector by which protein B is moved. Geometric matching algorithms aim to detect spatial complementarity patterns by local shape feature matching. (Zhang et al., 2016, Kozakov et al., 2017).

The scoring functions for global approaches used to focus on shape complementarity in their beginnings in the 1990s, but over time electrostatic and desolvation contributions as well as pairwise interaction terms have been introduced to improve accuracy (Kozakov et al., 2017, Zhang et al. 2016).

The molecular flexibility problem

There is a difference between the conformation of the proteins in the complex (bound) and outside the complex (unbound) *in vivo*, which contributes to the functional change upon partner binding. This complicates the search for the correct complex conformation (Vakser, 2014).

To deal with this problem, following strategies are pursued:

Coarse-graining

The lack of flexibility in the rigid body docking approach can in part be dealt with by coarse-graining: by lowering the structural resolution, the difference between unbound and bound state is less significant. The simplified, coarse-grained representation of the docking partners usually suffices for the prediction of the binding site, i. e. the interface. If one is interested in binding affinity prediction or inhibition of binding, high resolution structures are needed (Vakser, 2014).

This method is also used by the Rosetta Dock low resolution stage by replacing side chains with centroids, which was discussed earlier (Chaudhury et al., 2011).

Tolerance of overlaps

Most rigid docking approaches rely on shape complementarity scoring functions. If the shape complementarity term of the scoring function allows for some overlap, moderate differences between bound and unbound structures can be tolerated. If steric clashing appears, it can be resolved by energy minimization in the next step (Kozakov et al., 2017).

Flexible docking

Flexible docking takes the protein flexibility into account by involving a conformational transition between bound and unbound states of one or both partners. Together with higher resolution input structures they can help to reveal more structural details of the interface. The conformational space can also be discretized into rotameric states of the side chains, with transitions between the bound and unbound state. There is a huge number of internal coordinates to be sampled (Vakser, 2014).

Compared to other surface residues, the side chains of the interface have been shown to undergo smaller structural fluctuations, pointing to less flexibility than in non-interface regions (Ruvinsky and Vakser, 2010).

5.2.2 High resolution docking stage / refinement

In this step, the complex structures found in the low-resolution docking step are refined by a high-resolution docking step with energy minimization. (Zhang et al., 2016).

Rosetta Dock high resolution docking stage

After the low-resolution stage, the centroid pseudo-atoms are changed back to their original side chain atoms. This is followed by a 50-step Monte Carlo minimization, where at first random perturbation steps around 0.1 Å (translations) and 3.0 degrees (rotations) are performed. Second, the rigid body orientation is energy minimized, and in the third step the side-chain conformations are optimized by running “RotamerTrials”, where each residue is placed into its optimal rotamer state given the neighboring residue’s conformations. Fourth, the

Metropolis criteria are tested, deciding whether to accept the new structure or not (Chaudhury et al., 2011).

Genetic algorithms

Another example of how to refine the initial low resolution docking results are genetic algorithms. These are non-deterministic search algorithms that can optimize the dihedral angles of specified residues. At first, sets of dihedral angles are randomly generated, each one representing a conformation (e.g., a sequence of main chain dihedral angles). Then, a fitness function will select some of those sets, with higher fitness values more likely to be selected. In the following “Crossover”-step, the values for angles are exchanged randomly among different sets. Then, for further randomness, some will be mutated to a totally different value. At last, the fitness function will be invoked again and compare the fitness values with the previous ones. If the fitness values are worse than before, other sets will be selected, otherwise the new structure will be stored (Zhang et al., 2016).

5.2.3 Ranking stage

The refined structures of the complex are now ready to be rated by a scoring function. Ideally, near-native structures score higher than false positive matches. The scoring functions of the docking stage often have to make use of approximations for reasons of computational efficiency, which in turn might impair the prediction of affinity. Thus, a more rigorous scoring function for ranking of the structures can be useful. Important criteria for ranking methods are the f_nat value, which is the fraction of ligand-receptor contacts that can also be found in the native structure. The RMSD (root mean square deviation) compares the native and the generated conformation of the complex by the root mean square of the relative atomic distances. f_nat and RMSD values are also used to benchmark docking programs. In general, there are two basic scoring approaches: force field based and knowledge-based scoring functions (Zhang et al., 2016).

Force field based scoring functions

Force fields estimate different potentials by taking into account certain sets of physical parameters, e.g. standard bond angles or Van der Waals radii. The scoring function calculates potential energy terms as a function of the atomic coordinates for those parameter sets. The sum of those energy terms leads to the energy of the conformation. Since these functions try to account for physical forces that exist *in vitro* and *in vivo*, they are also called “physical” scoring functions. Yet, different force fields use different sets of parameters in order to optimize the correlation between RMSD and energy, also their energy formulas differ. Lower energy of a conformation will usually mean a better rank (Zhang et al., 2016).

Knowledge based scoring functions

Instead of optimizing specified terms themselves, knowledge-based scoring functions aim to find the best weight of each scoring term based on training sets. These training sets can take into account experimental knowledge that is usually based on matches of bound and unbound forms of proteins. This method is also called “informatics-driven”. For example, residue-based pair potentials can be used to optimize the f_nat value. (Zhang et al. 2016, Liu & Vakser, 2011). Such pairwise potentials can be found in several docking programs. A good example is the “Decoys As the Reference State” (DARS) approach. It uses docked conformations with good shape complementarity as reference state for the extraction of the interaction frequency on an atomic level. This allows statistical potentials for different atom type pairs to be defined. They depend on the probability of contact between two atoms in the native complex structure and the probability of contact in an appropriate reference state. This resulting potential can help to find near-native conformations (Chuang et al., 2008).

Rosetta Dock scoring function

After the high-resolution docking stage, the complex is now rated by a full atom scoring function, by standard consisting of Van der Waals attractive and repulsive terms, a solvation term, an explicit hydrogen bonding term, statistical residue-residue pairwise interaction term, a side chain conformational energy term and an electrostatic term (Chaudhury et al., 2011).

6. MEMBRANE DOCKING METHODS

6.1 Rosetta MPDock Protocol

Since membrane proteins are difficult to overexpress and reconstitute into membrane-like structures *in vitro*, it is hard to determine their structures with methods involving crystallization techniques. Hence, computational tools such as RosettaMP are important to examine membrane protein structures. RosettaMP is a framework of the Rosetta software, providing scoring, conformational sampling and mutation routines for membrane proteins that can be combined to create new protocols, for example membrane protein-protein docking or prediction of free energy changes upon mutation. There are advantages of membrane proteins compared to soluble molecules that help to make sampling more efficient: their structural motifs are limited to either alpha-helix bundles or beta-barrels, and these secondary structures adopt preferred orientations within the membrane. Docking search space is also limited to two dimensions of the lipid bilayer, imposing structural constraints on the protein's position and orientation. However, membrane proteins are typically larger than soluble proteins, raising the expense of computational power in conformational space search. Moreover, there is a need for a representation of the lipid bilayer. This can be achieved by implementing a layered continuum solvation model, which is computationally inexpensive but cannot account neither for membrane fluctuations nor interactions between lipids and the proteins. Also, the precise location of the lipid bilayer is experimentally inaccessible, which might also be challenging for the scoring function's prediction ability (Alford et al., 2015).

According to Alford and her colleagues, the protein-protein docking algorithm of Rosetta MP, MPDock, is a combination of RosettaMP with the standard RosettaDock algorithm, designed for docking proteins in a lipid bilayer. As mentioned before, the MPDock protocol requires a prepack step to repack side chains using rotamer trials. In the docking step, as in the RosettaDock protocol, there is a low-resolution centroid stage that starts with initial perturbation moves of 3 Å translations and 8-degree rotations, followed by a rigid body Monte Carlo

search. A high-resolution refinement stage for energy minimization and side chain optimization is performed next. In each stage different representations of the membrane are used: the membrane is considered to consist of five discrete layers that describe the hydrophobic core, the interface, and solvent regions in the low-resolution stage. In the full-atom stage the membrane is represented by an implicit membrane model, describing a continuous dielectric gradient from the hydrophobic core to the solvent region. However, this model also comprises a hydrophobic layer, a transition region, and a soluble region. The scoring functions of both stages are combinations of the standard RosettaDock scoring functions mentioned before and RosettaMP membrane score terms. In case of the low resolution steps, the RosettaMP membrane energy terms include three different knowledge based potentials: the first accounts for the propensity of a single residue to reside at a given depth in the membrane (referred to as the environment scoring term mp_env), the second for the pairwise interactions of two residues with given depth in the membrane, and the third for residue density; also there are scores for lipophilicity and penalties for non-helical secondary structures in the membrane, for residues outside the hydrophobic layer of the lipid membrane, and for transmembrane helices outside of the hydrophobic bilayer. At the high-resolution level, there are energy terms for atomic solvation free energy, knowledge-based potentials that describe the propensity of a single atom to be located at the given membrane depth, and a hydrogen bonding term that is adjusted to be stronger inside of the membrane (Alford et al., 2015).

6.2 Memdock Server

The Memdock server (http://bioinfo3d.cs.tau.ac.il/Memdock/) is designed for alpha-helical membrane proteins to be docked taking the lipid bilayer into account at the docking stage as well as for ranking and refining (Hurwitz et al., 2016). The docking process is based on the PatchDock algorithm, which uses a rigid body soft docking approach and ranks the candidates mainly by geometric shape complementarity (Schneidman-Duhovny et al., 2005). The Memdock algorithm is modified to enforce the placement of the docked proteins to be a membrane consistent orientation. This is achieved by constraining the angle change between the protein helical axis and the membrane normal to 0.4 radians and

limiting the translations along the Z-axis to 8 Angstrom. The refinement step is carried out by a modified Fiber Dock algorithm, capable of backbone and sidechain flexibility. Ligand orientation is optimized by a rigid body Monte Carlo minimization. The membrane aspect is also taken into account by constraining the ligand perturbations to deliver membrane consistent orientation results. Finally, in the re-ranking stage the results are scored and re-ranked by the MemScore energy function, which was developed based on statistical knowledge from known 3D structures of alpha-helical membrane proteins. Hydrophobic and polar layers of the lipid bilayer are taken into account. The energy terms comprise environmental energy, softened van der Waals energy, partial electrostatics energy and estimations of the binding free energy, which on its part considers hydrogen bonds, aliphatic interaction, and pi-pi interaction energies. The environmental term is analogous to the Rosetta membrane environment scoring term (Hurwitz et al., 2016).

7. DOCKING RESULTS

7.1 Results of Rosetta MPDock

Four protein coordinate files, containing OTR and D2R in an orientation where certain transmembrane helices face each other, were prepared in chapter 4, including their span files. These were now used as input for a docking full run in the Rosetta MP protein-protein docking protocol. 1000 poses and one score file were generated for each orientation. This score file contains information of energy terms of the scoring function and, most importantly, the total energy score of the whole complex, and the I_Sc term, which means interface score. I_Sc is calculated by the difference between the total energy score of the whole complex and the total energy scores of the isolated docking partners (Gray et al., 2003).

Since the RosettaMP docking protocol does not rank the poses in the scoring file, the results had to be evaluated manually. The complexes with an I_Sc below -75.000 were considered low energy complexes. Because the lowest total energy did not always correspond with the lowest I_Sc, the author created an arbitrary value, accounting for both terms:

$$[\text{weighted minimum score}] = ([\text{I\_Sc}] - 37.5) \times [\text{total energy}]$$

The I_Sc value was considered most important for the selection of the most promising result of each orientation, since intermolecular forces and shape complementarity are best depicted by it. Moreover, the population of low energy interfaces with I_Sc below -75 of each orientation was considered of interest too, since in a way it represents the probability of a meaningful complex to form from a given orientation. This is supported by the hypothesis that the global minimum in the energy landscape of a complex is not only deeper than local minima by a significant margin, but also its size and density are often bigger than local minima; the size is determined by the number of matches in a cluster, the density is defined by the number of matches in a cluster divided by its volume (O'Toole & Vakser, 2008).

ORIENTATION / VALUE	OTR 4/5 D2R 4/5	OTR 1/2 D2R 4/5	OTR 1/2 D2R 1/7	OTR 4/5 D2R 1/7
Low I_Sc population (I_Sc < -75.000)	96 of 1000	33 of 1000	24 of 1000	48 of 1000
Lowest I_Sc of dataset	-84.267	-79.041	-81.669	-83.569
I_Sc average whole dataset	-61.191	-61.661	-64.623	-64.863
Lowest total energy of dataset	-1980.913	-2006.023	-2006.675	-2001.043
Total energy average whole Dataset	-1933.928	-1937.138	-1901.012	-1928.268
Best weighted min value/ corresponding I_SC	241209 / -84.627	232437 / -78.325	232825 / -78.814	242264 / -83.569

Table 7.1a contains values from the score file generated by the Rosetta MPDock program, as well as derived and statistical values for interface evaluation.

The data in table 7.1a show that the lowest I_Sc values of each orientation are quite similar, with the “OTR 4/5 - D2R 4/5” orientation scoring best, followed very closely by the “OTR 4/5 - D2R 1/7” orientation. “OTR 1/2 - D2R 1/7” and “OTR 1/2 - D2R 4/5” performed a little worse. The values for total energy score are quite consistent in all constellations, with small deficits of the “OTR 4/5 - D2R 4/5” orientation. The “weighted minimum score” was created to quantify the relevance of the total energy scores in relation to the I_Sc values. For example, if a pose with a low I_Sc value had a compromised total energy, e.g., due to steric clashing, the weighted minimum score would help to avoid selection of such interfaces. The best poses of the “45_45” and “45_17” orientations reached the highest weighted minimum scores, also the best “weighted minimum score-structures” were identical with the “lowest I_Sc-structure” of that orientation. All these data did not give a definitive answer to the question of the correct partner orientation, not to mention the correct global minimum for this complex. However, the docking results with the best weighted minimum scores of each orientation were selected for further analysis and are referred to as “best results” (see Chapter 8 and 9).

7.2 Memdock Results

The prepacked models of the active forms of the proteins (generated in chapter 5) were separately uploaded to the Memdock server. 20 results and a scoring file were received, the scoring file contained overall scores from the Memdock and the FibreDock scoring functions. Then all poses were visually inspected. Many of them involved amphiphilic helix 8 of OTR as part of the interface. Figure 7.2a shows an overlay of the four poses with the best Memdock scores, indicating that D2R, in all structures, lies in a similar direction from OTR's point of view.

Figure 7.2a: overlay of the best Memdock results. It appears this docking algorithm has a preference for D2R to be positioned near OXT's TM1/2/3 region.

The result with the highest score is almost a rotationally symmetrical arrangement, where the same transmembrane helices from each partner face each other, meaning that for example amphiphilic helix8 of OTR faces TM4 of D2R and vice versa. Both TM2 are in the middle of the interface, flanked by interacting TM1 and TM3. This level of shape complementarity could not be reached by any other interface of this run. With a score superiority by a margin of

14% (-669.5 vs. -590.6) to the second-best structure, this pose was selected for further analysis (Fig 7.2b).

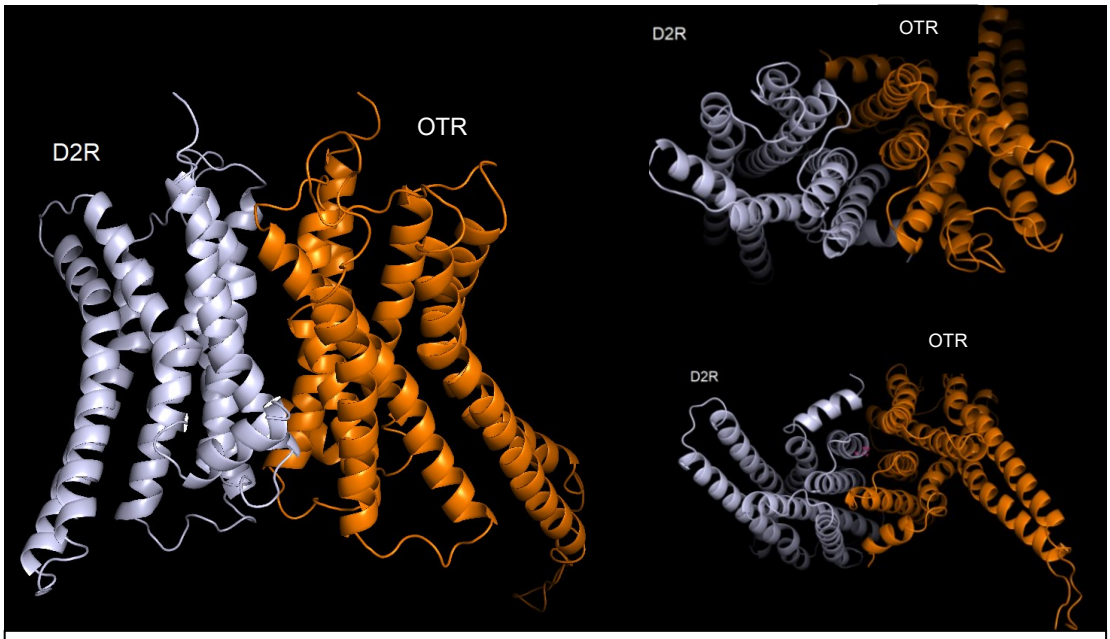


Figure 7.2b: Memdock result with best score in total perspective (left), in top view (top right) and bottom view (bottom right)

A more detailed analysis of the best Memdock result follows in chapter 9 after performing molecular dynamics simulation.

8. ROSETTA MP MODEL ANALYSIS

8.1 Methods

Pymol visual inspection and hydrogen bonds analysis

The structures of the docking results with the lowest weighted minima of each orientation were first visually inspected via Pymol software (Schrödinger LLC, 2020) to estimate plausibility of the orientation and shape complementarity. To gain more information, the “polar contacts” feature of Pymol was used to identify intermolecular hydrogen bonds.

PyRosetta Alanine Scan

The “D090 PyRosetta Alanine Scan” script mimes a technique from molecular biology, where the function, form or stability of proteins can be analyzed by introducing iterative mutations of each amino acid to alanine. If the function is impaired by the mutation, one can assume that the altered residue is relevant for the examined property. In our case, the Alanine Scan script is used to identify contributions of individual amino acids at the interface to the overall protein interaction by mutating each of them to alanine and compute free energy changes for each amino acid. This included energy terms for the attractive part of the Lennard-Jones potential, a linear distance-dependent repulsive term, hydrogen bond potentials, Coulomb electrostatics, and an implicit solvation model. By calculating the difference for the isolated proteins and the bound complex, the binding energy change (ddG) can be computed. The side chains of the mutants are repacked while the backbone remains static. Since the resulting free energy changes are based on knowledge-based potentials that do not represent physical energies, the results of the Alanine Scan can depict purely qualitative information on amino acids involved in interface formation. Free energy shifts towards a more positive value by mutation indicate poorer binding of the mutated residue, meaning that the original amino acid contributes to a more negative free energy of binding value (Kortemme and Baker, 2002).

It was observed that mostly lipophilic amino acids like valine, leucine or isoleucine contributed to a more negative value as described, whereas polar and large

residues shifted towards a more positive binding free energy value. Why this is the case remains elusive, yet it can be assumed that they are in a way involved in the protein-protein interface of the docking result, respectively.

8.2 Results

OTR 4/5 - D2R 4/5 orientation

Pymol found one intermolecular hydrogen bond between OTR glutamine 229 and D2R serine 147. The PyRosetta Alanine Scan protocol found 48 amino acids at the interface of the best docking result. Based on interacting amino acids in proximity to interacting residues of the docking partner, the transmembrane helices (TM) presumed to be involved in the interface formation are TM3, TM5 and TM6 of OTR, and TM3, TM4, TM5 and TM6 of D2R. To be exact, the interface should be called “OTR 3/5/6 – D2R 3/4/5/6” interface (Fig 8.2a). The whole

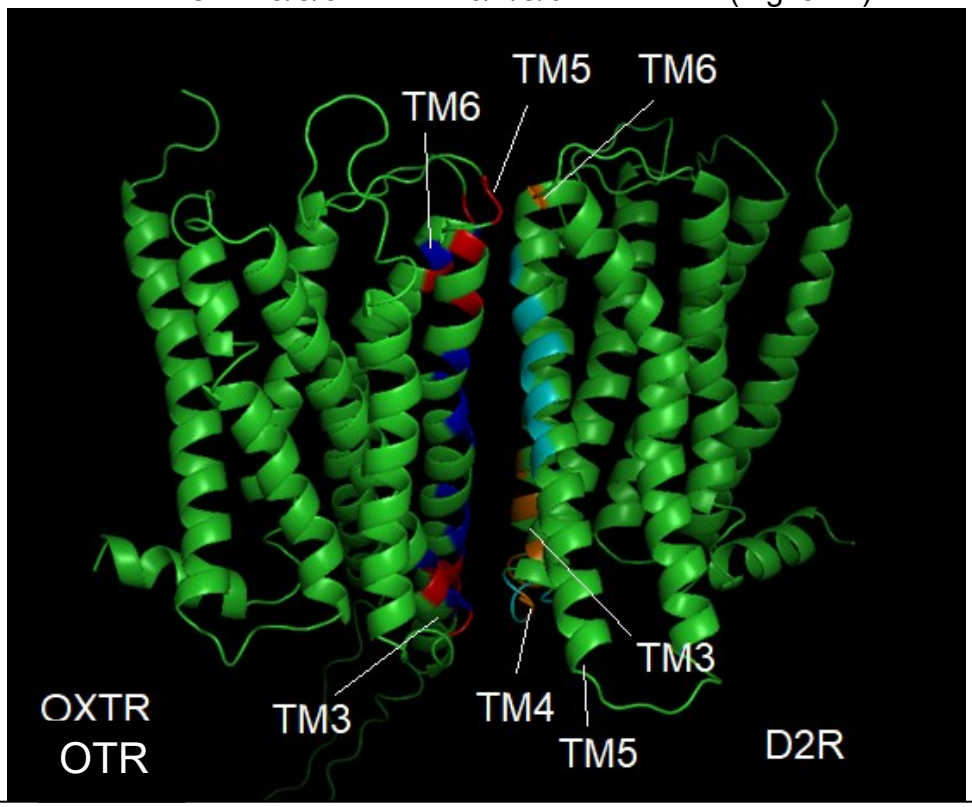


Fig. 8.2a shows the result of the OXR4/5 - D2R4/5 orientation and the transmembrane helices involved in the formation of the resulting “OXR 3/5/6 – D2R 3/4/5/6”-interface. Blue and cyan areas indicate a positive energy shift upon mutation to alanine, and hence, a good fit of the original residue can be presumed. Red and orange areas indicate that mutation to alanine lowers the energy of the complex, questioning the optimal fit for the residue.

resulting best complex for the “OTR 4/5- D2R 4/5” orientation will be referred to as “best RosettaMP docking result”.

The blue/cyan- colored amino acids are contacts that are beneficial for binding according to the alanine scan, whereas red/orange areas are rated with positive free energy contributions. Table 8.2a summarizes the most important amino acids that are assumed to contribute to the interface. The colors in the table correspond with the residue’s colors in the figures.

	Oxytocin receptor	Dopamine D2 receptor
Positive energy contribution upon mutation	L206, V211, I214, G221, F225, W300	M138, V200A, I203, V204
Negative energy contribution upon mutation	P194, W195, K198, K226	S147, R151
H-Bond	Q232	S147

Table 8.2a. Important residues that are likely to be involved in the formation of the “OXR 3/5/6 – D2R 3/4/5/6”-interface.

Figures 8.2b - 8.2e depict details of the interface amino acids according to the PyRosetta Alanine scan and the hydrogen bond prediction function of Pymol.

Fig. 8.2b shows the top view of the protein complex. The residues shown lie outside the membrane; hence they are exposed to water which is congruent with the polar nature of the residues.

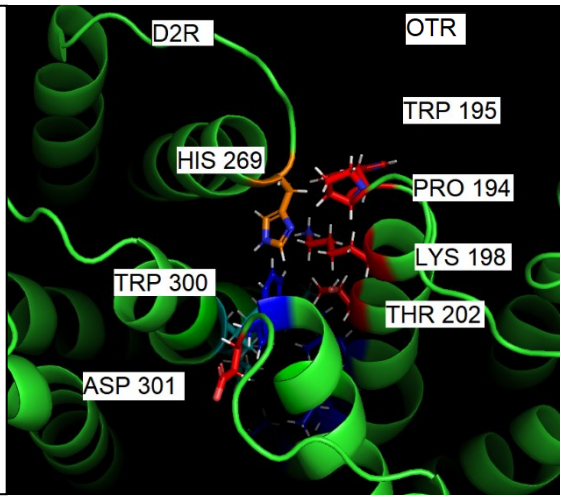
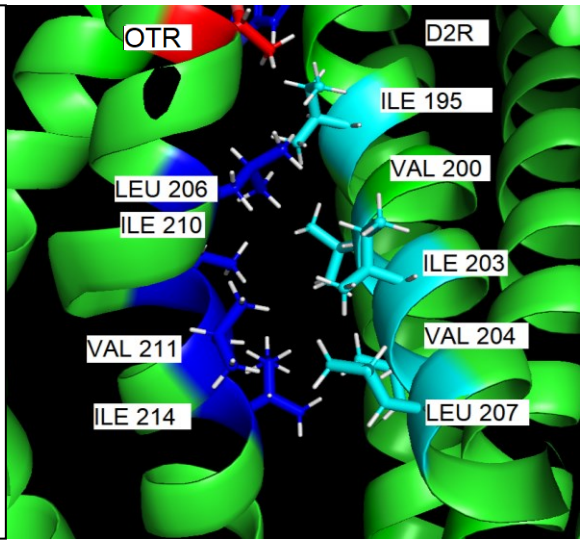


Fig. 8.2c. The middle section of the complex is characterized by lipophilic amino acids in close proximity to each other, and one may assume that shape complementarity is sufficient to allow for intermolecular hydrophobic attraction.



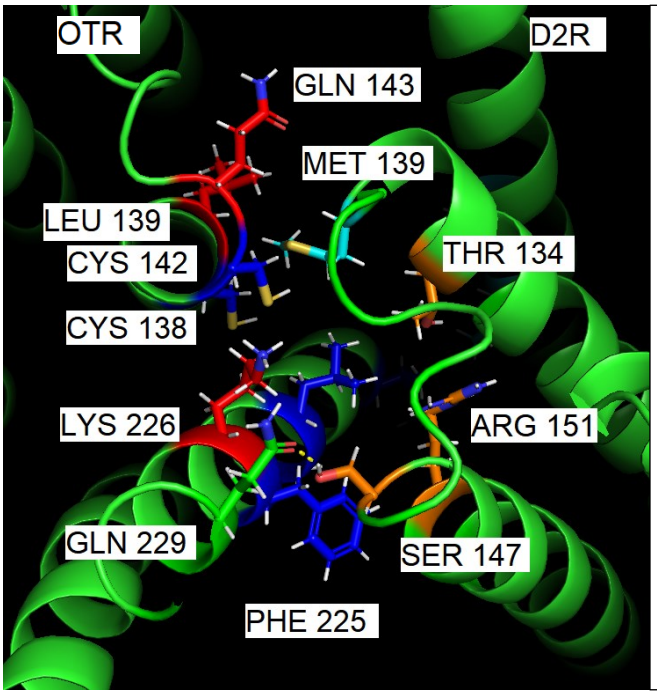
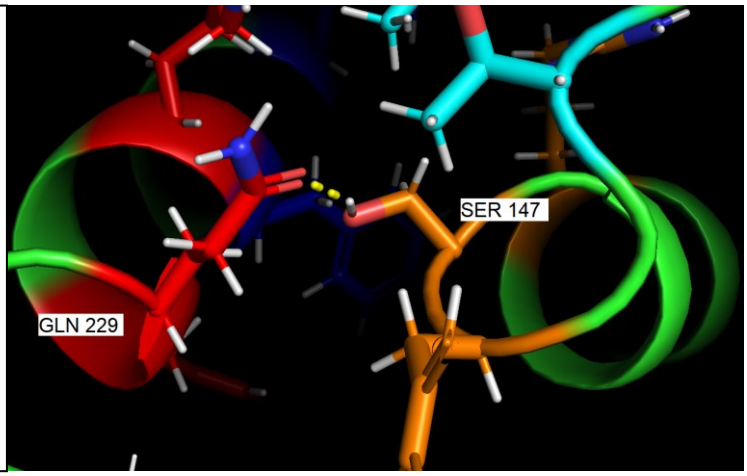


Fig. 8.2d shows the bottom view of the protein-protein interface.

Fig. 8.2e depicts the hydrogen bond between Glutamine 229 of OXR and Serine 147 of D2R.



OTR 4/5 - D2R 1/7 orientation

As shown in table 8.2a, the interface of this is complex comparable to the 4/5 - 4/5 docking result in terms of I_Sc, total energy, and weighted minimum score. Visual inspection reveals that there is only little interaction in the middle section between helices, so in terms of shape complementarity this structure cannot compete with the 4/5 - 4/5 interface where all sections are in contact. 35 amino acids that were involved in interactions were found via Alanine Scan (figures 8.2f-8.2j). As before, polar and ionizable amino acids tend to be rated with more positive free energy values than their alanine mutations.

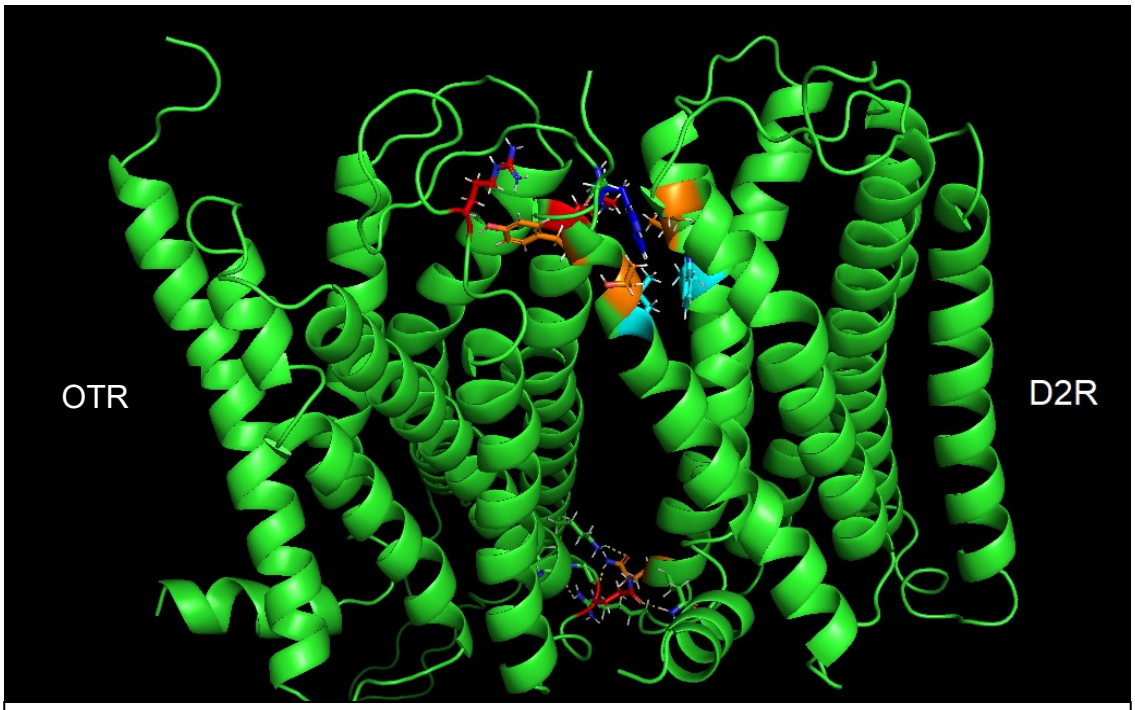


Figure 8.2f: The OXR 4/5 - D2R 1/7 orientation produced contacts at the top and the bottom of the complex. The membrane core region lacks contacts.

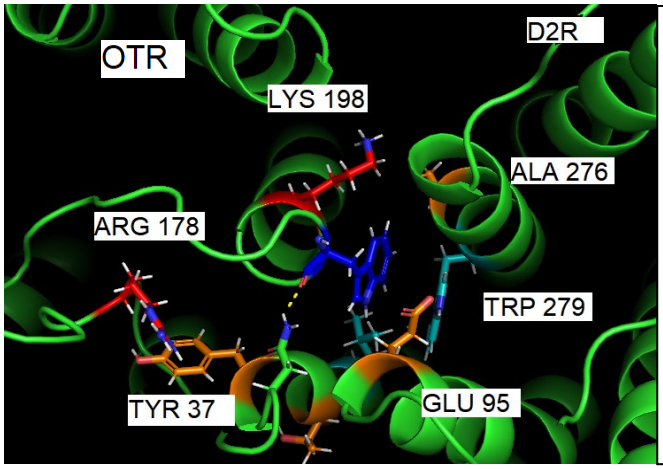


Fig 8.2g

Top view of the 4/5 -1/7 orientation docking result

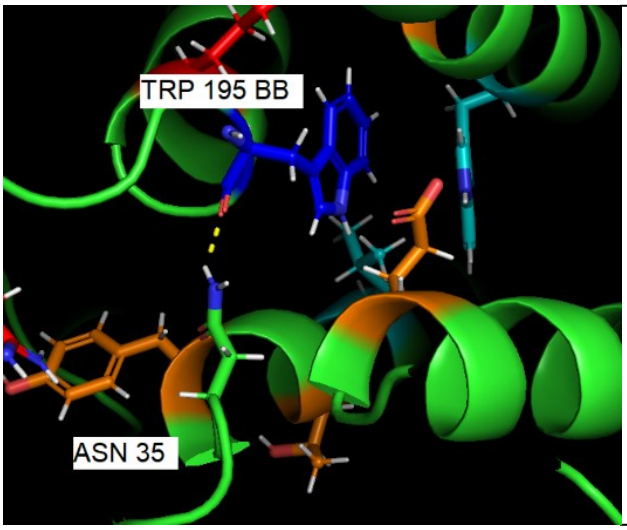
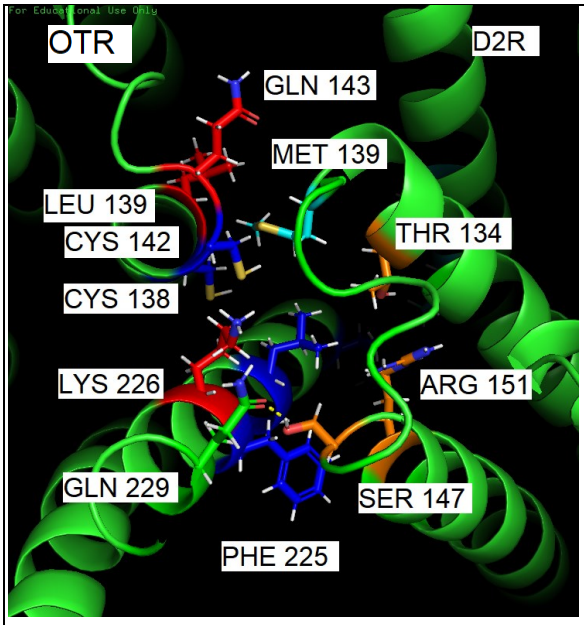


Fig 8.2h: At the top, a hydrogen bond between OXR tryptophane 195 (backbone) and asparagine 35 was found by Pymol.

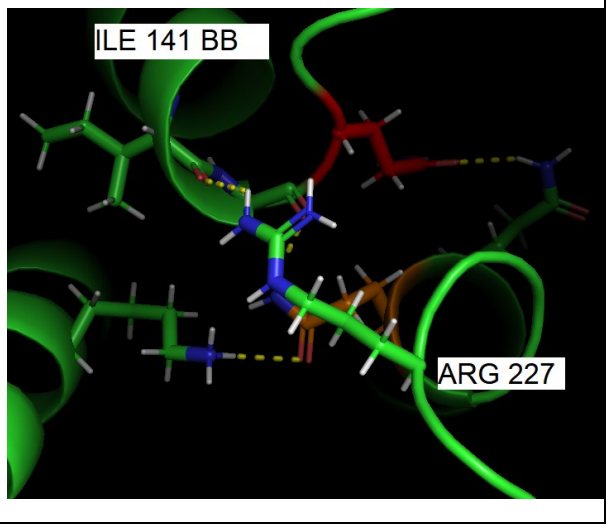
.



Interactions at the bottom

Fig 8.2i: Looking at the hydrogen bonds detected by Pymol one can see that there is a hydrogen bond network at the bottom formed by the backbone oxygen of cysteine 142 and lysine 243 of OTR which are both linked to glutamine 231 of D2R. Glutamine 232 of D2R and glutamine 143 of OTR also form an H-bond.

Fig. 8.2j shows a hydrogen interaction between the backbone of OTR-isoleucine 141 and arginine 227. This view is just below Fig 8.2i.



OTR 1/2 - D2R 4/5 orientation

As shown in figure 8.2k, the best result of this orientation offers at least some contact of helices in the membrane core region, three hydrogen bonds at the bottom and - according to the Alanine Scan analysis - mostly hydrophobic interaction, even at the top which is exposed to water. 32 interacting amino acids were found.

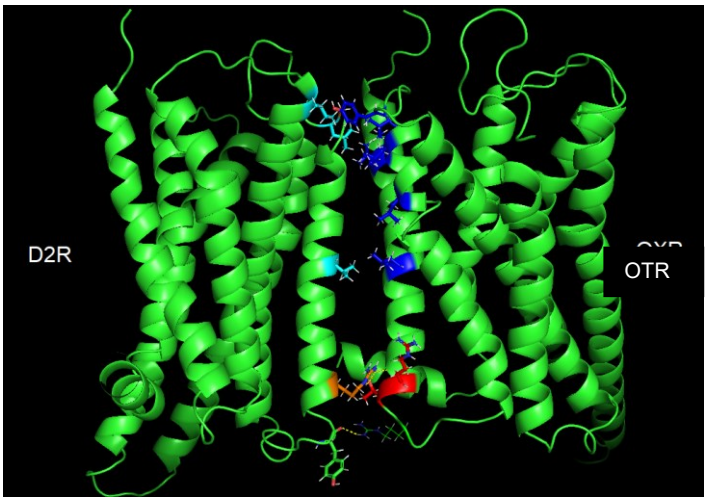


Fig. 8.2k:

The OXR 1/2 – D2R 4/5 interface in total perspective

OTR 1/2 - D2R 1/7 orientation

The best result contains one hydrogen bond at the top and four at the bottom, involving three arginine residues, two of which interact with glutamate residues, which might result in a salt bridge. Interestingly, the amphiphilic helix 8 of both proteins are orientated right angled to each other, with the C-terminal leucine 345 of OTR near leucine 304 of D2R. According to the Alanine scan, there are 38 interacting amino acids in total. A total perspective is provided in figure 8.2l.

Fig.8.2l:

The interface generated by OXR 1/2 – D2R1/7 orientation in total perspective

8.3 Complex structure selection

The aim of the selection was to find the most promising docking result among all orientations for further investigation. Each of the four orientations produced 1000 results. A comparison of features can be found in table 8.3a.

The “4/5-4/5” orientation produced the most low energy structures with an I_Sc below -75. It also generated the structure with the lowest I_Sc of all 4000 poses and the Alanine Scan detected by far the most interacting amino acids. Also, the best structure of this orientation was the only one among all top scorers with contacts at the top, middle and bottom section of the interface. Interestingly, there was only one interchain hydrogen bond detected by Pymol, in contrast to other top scoring results, which achieved 3 to 5 hydrogen bond interactions.

Intermolecular hydrogen bonding is weighted relatively high in the RosettaMP protocol scoring function; hence, all other energy terms have to compensate even more to achieve such a I_Sc score. This led to the assumption that some additional hydrogen bonds might appear by further minimization or the following step of a molecular dynamics simulation. The best result of the “45-45” orientation comprising an OTR TM 3/5/6 - D2R TM 3/4/5/6 interface was selected for further testing because of the highest number of interacting amino acids.

Orientation / **Criterion**	**OTR 4/5 D2R 4/5**	**OTR 1/2 D2R 4/5**	**OTR 1/2 D2R 1/7**	**OTR 4/5 D2R 1/7**
Alanine Scan-detected interface amino acid count	48	32	38	35
Interchain hydrogen bond count	1	3	5	5

Table 8.3a: Comparison of data gained via PyRosetta Alanine Scan and Pymol hydrogen bond detection.

9. MOLECULAR DYNAMICS WITH DOCKED LINKER

The goal of this step is to examine the stability of the docked protein complexes that are connected by a heterobivalent linker molecule by running a molecular dynamics simulation. Stability can be measured in an overall RMSD in relation to the initial coordinates of the complex. Also, the interactions at the endpoint of the MD simulations are analyzed by performing an Alanine Scan, additionally extended by other mutation scans (arginine, aspartate, and serine instead of alanine).

The linker consists of an oxytocin receptor agonist and a dopamine D2 receptor agonist that are connected with an aliphatic linker bridge by covalent bonds.

9.1 Docking ligands with LigandScout

First, known agonists of proteins were chosen: dOTK for the oxytocin receptor (Figure 9.1a), and a 2-N,N-dipropyl-aminoindane based ligand (Figure 9.1b). Then, models of them were built in LigandScout's molecule 2D editor (Wolber & Langer, 2005). They were consecutively docked to their receptor by LigandScout using the implemented AutoDock Vina algorithm. The most important property of the docking results, besides an acceptable score, was the "cytosolic" orientation of the residues for linker attachment. This was necessary to avoid steric clashing in the next step.

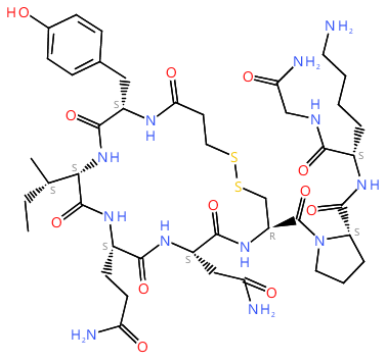


Figure 9.1a: dOTK

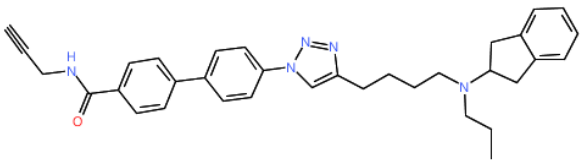


Figure 9.1b: D2 agonist

9.2 Linking the ligands with Maestro

In the real world, the linker molecule would be synthesized by amide bond formation between dOTK and the carboxyl group of the aliphatic linker (Figure 9.2a) and between the terminal C-C triple bond of the D2 agonist molecule by “click chemistry”, resulting in a pentacyclic azide ring (Figure 9.2b, Kolb et al. 2001). The Maestro software was used to model this reaction products, connecting the two docked ligands with an aliphatic bridge (Schrödinger, 2021).

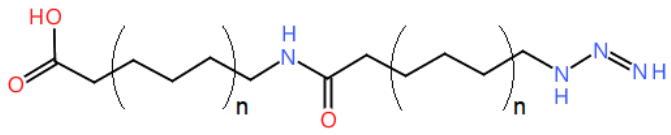


Figure 9.2a:

n-aliphatic linker bridge

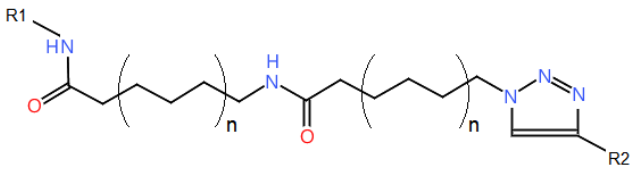


Figure 9.2b:

Connected linker,
R1 = dOTK, R2 = dopamine D2 agonist

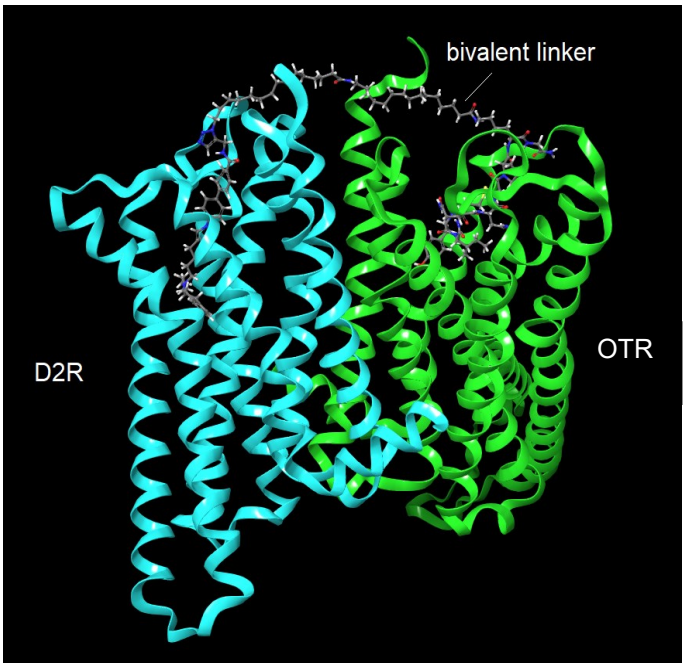


Figure 9.2c:

The Memdock best result for the D2R-OTR complex with both ligands docked and linked.

For the best docking result of RosettaMP dock, the linker with a length of n=3 was built. In case of the best Memdock docking result, n=4 was needed to overcome the larger distance between the binding sites (figure 9.2c).

9.3 Building a membrane with CHARMM-GUI

CHARMM (Chemistry at HARvard Macromolecular Mechanics) is a molecular simulation program applicable to many particle systems. It targets biological systems such as proteins, small molecule ligands or nucleic acids in solution, crystals and membrane environments. It includes computational tools for sampling, free energy estimation, minimization, dynamics, modeling, and analysis. There are a variety of energy functions to choose from, for example mixed quantum mechanical-molecular force fields or classical all-atom energy potentials (Brooks et al., 2009).

CHARMM-GUI is a web-based graphical user interface for CHARMM, designed for the generation of various CHARMM input files, e.g., a membrane-embedded system via Membrane Builder. Its protocol consists of six steps: reading the protein structure, protein orientation, system size determination, generation of components (water, ions, lipid bilayer), assembly and equilibration (Jo et al. 2007 & 2008).

CHARMM-GUI's Membrane builder was used to set up a plasma membrane-like lipid bilayer, water, and solved ions for each docking method with and without the docked ligand, respectively. The protein dimers were left in their orientation from the docking process, since both docking algorithms should account for the correct membrane orientation of the docked model. The composition of the lipid bilayer was adapted from figure 2 of the paper *"Membrane Lipids: Where They Are and How They Behave"* by Van Meer et al., 2008. Table 9.3a shows the numbers of molecules that were inserted. Potassium and chloride ions were added to reach a concentration of 0.15M KCl, which mimics physiological salt concentration. The water layer was generated with a 60 Angstrom height on both sides to cover all regions of the proteins. The system was assembled (Figure 9.3a), and input files were generated for "openmm"-applications.

Molecule	Upper Leaflet	Lower Leaflet
Cholesterol	106-108	105
POPC (Phosphatidylcholine)	45	42
POPE (Phosphatidylethanolamine)	25	24
POPS (Phosphatidylserine)	10	10
POPI (Phosphatidylinositol)	4	4
PSM (Sphingolipids)	26	25

Table 9.3a: the lipid bilayer was created based on these numbers of lipids.

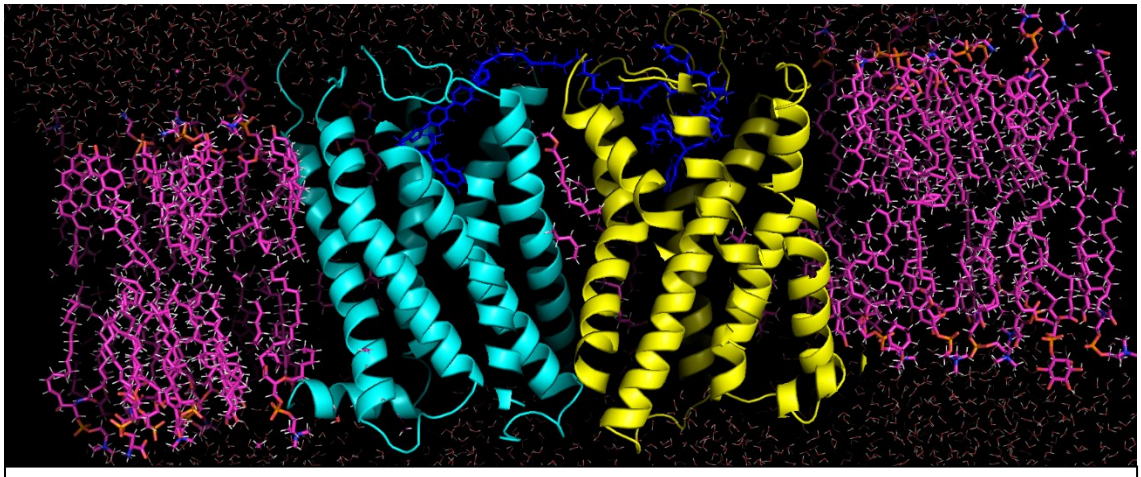


Fig. 9.3a Cross-section through the membrane system built by CHARMM-GUI
The lipid bilayer molecules are presented in magenta, D2R in cyan, OTR in yellow and the bivalent ligand in blue. Water can be seen below and on top of the membrane bilayer.

9.4 Molecular dynamics simulation

A molecular dynamics (MD) simulation is an *in silico* method that captures the behavior of biomolecules such as proteins in full atomic detail and a fine temporal resolution; it predicts the movement of every atom in a molecular system by modeling the physics of interatomic interactions. MD simulations can help to understand functional mechanisms of protein folding, conformational change or ligand binding. Moreover, the response to perturbations such as mutation, modification or protonation of biomolecules can be observed (Hollingsworth and Dror, 2018).

The OpenMM molecular dynamics simulation toolkit was used to perform the MD in the CHARMM force field (Eastman et al., 2017, Brooks et al., 2009). For each system, a 50 ns simulation was carried out in 5 million steps, with coordinates written to a frame every 5000 steps. The last of the resulting 1000 frames was captured and converted to a pdb file for further analysis.

10. ANALYSIS OF MD RESULTS

10.1 Surface area assessment

The surface areas of the last frames of the MDs were measured and then compared to the docking results before MD via Pymol's "get surface area" function. After removing membrane, water and ligand molecules, leaving only the proteins, the accessible surface area (ASA) of the complex and of each individual partner was measured. Then, D2R was translated by 100 Angstrom, separating the complex and the whole process was repeated. The size of the interfaces was calculated as buried surface area, by subtracting the accessible surface area of the of the bound complex from sum of the unbound subunits (Chothia and Janin, 1975). The difference between the BSA before and after MD was also included.

The results are shown in Table 10.1a. The Memdock interface right after docking extends to 3479 Å², whereas the Rosetta interface reaches only 2567 Å². After MD, the interface of the Rosetta MP docking result lost ~15% of its initial size

(397Å²), while the Memdock result showed the opposite behavior: it gained 703 Å² upon MD simulation, which makes a surplus of ~20% of BSA. After MD, the Memdock interface has nearly twice the buried surface area than its Rosetta MPDock counterpart.

Surface area table		**ASA before MD** (Å²)		**BSA** (Å²)	**ASA after MD** (Å²)		**BSA** (Å²)	**ΔBSA** (Å²)	**ΔASA bound** (Å²)
		bound	unbound		bound	unbound			
Memdock result	OTR	17890	19685		17851	20003			
	D2R	13700	15376		14145	16166			
	All	31582	35056	**3479**	31987	36164	**4182**	**+703**	**+405**
Rosetta MPDock result	OTR	18473	19741		20054	21132			
	D2R	14069	15360		14866	15950			
	all	32534	35095	**2567**	34914	37078	**2168**	**-399**	**+2380**

Table 10.1a provides an overview for the complexes' surface areas and their change before and after the MD simulation. It is clear that the buried surface area (BSA) of the Memdock interface, which is already larger before MD, even increases, while BSA of the Rosetta MPDock interface decreases with the MD.

10.2 RMSD analysis method

The visual molecular dynamics program (VMD) was used to analyze the results of the MD, particularly the RMSD analysis tool, which enables one to measure the spatial root mean square deviation of the structure over time, in this case in relation to starting coordinates (Humphrey et al., 1996).

The RMSD of the backbone of the protein complex without linker, lipids, and water was measured over all 1000 frames in relation to the first frame, which represents the whole simulation time of 50 ns. Also, via the RMSD analysis alignment option, all complex coordinates were aligned to the position of frame 0 (the first frame), allowing to take the movement of the whole complex in the membrane out of the calculation, and instead elucidate internal structural changes of the backbone.

10.3 Rosetta MPDock result MD analysis

The most obvious changes happen at the beginning, where the whole system is adapting to the CHARMM force field, which can be considered or used as an energy minimization (Figure 10.3a) The average RMSD reaches 3.429 +/- 0.545 Angstrom. Following “jumps” in RMSD value may occur due to the high mobility of ICL3 of the oxytocin receptor (Figure 10.3b). It should be mentioned that 50 ns of simulation time was probably not enough to reach the global minimum of this conformation.

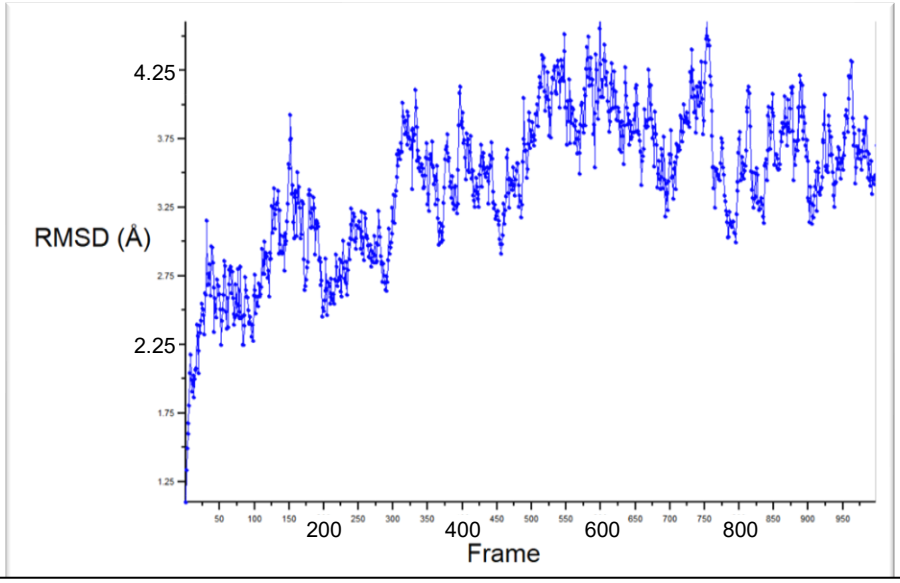


Fig. 10.3a: The RMSD analysis tool plot shows rapid change in the beginning of the MD of the best RosettaMP result, followed by moderate fluctuations.

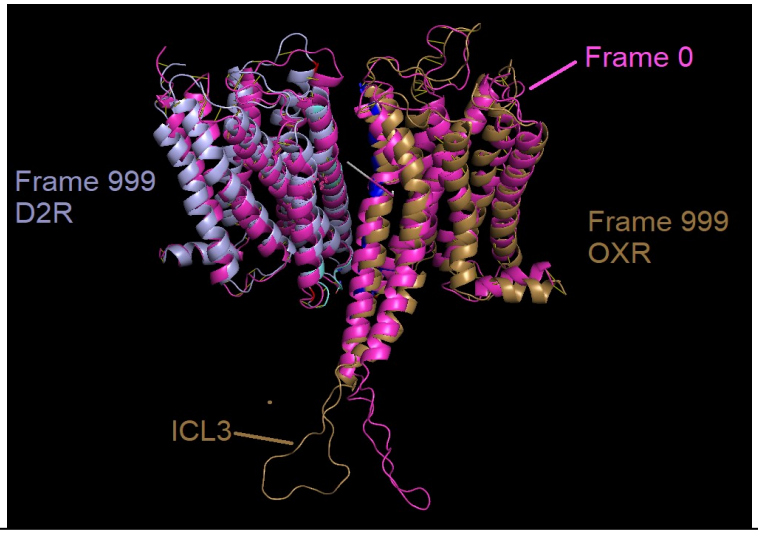


Fig. 10.3b. As can be seen, flexible loops underwent more conformational change than the transmembrane alpha helices, shifting RMSD towards higher values.

10.4 Memdock result MD analysis

The average RMSD of the backbone in relation to frame 0 is 4.755 +/- 0.872. Figure 10.4a shows that the most significant changes take place in the first quarter of the simulation, which is about 12.5 nanoseconds, then the protein remains quite stable. As with the Rosetta MPDock result, 50 ns are probably not enough simulation time to reach the global energy minimum. A shift of TM4 during MD was observed (Fig. 10.4b).

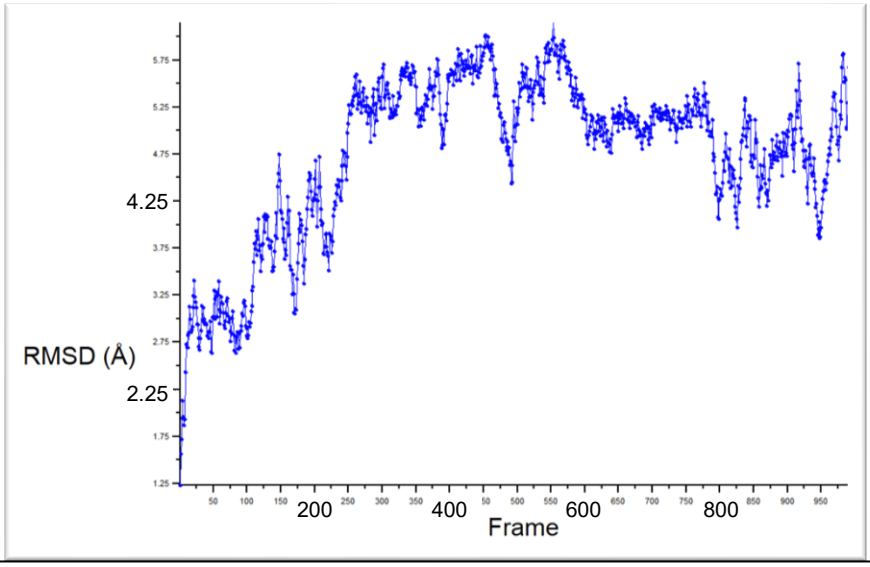


Figure 10.4a. The best Memdock result's RMSD plot over time is shown. Compared to the RosettaMP result, there is more overall conformational change in the first 250 frames, but less fluctuation is indicated afterwards.

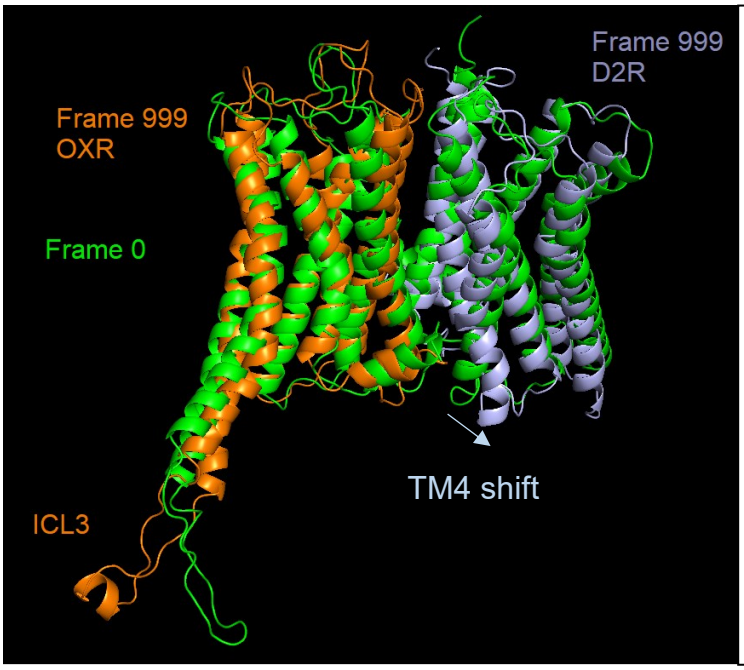


Fig. 10.4b: s before, the ICL3 and all other flexible loops of OXTR is very mobile in comparison to the alpha helices. Also, a shift of the lower part of transmembrane helix 4 was observed.

10.5 Alanine Scan of the Rosetta MPDock interface after MD

The last frame of each molecular dynamics simulations was edited for PyRosetta Alanine Scan: heterobivalent ligand, membrane and water molecules were removed, leaving only the proteins for analysis. This also applies to the scan of the Memdock docked complex after MD.

Table 10.5a and 10.5b show an overview of all amino acids of the interface according to Alanine Scan. The numbers are energy scores: the greater the value, the more positive the energy shift upon mutation becomes, which is indicated by blue (OTR) and green (D2R) bars. Red bars indicate a negative energy shift upon mutation.

	Alanine	Asparagine	Arginine	Serine
C138	0.5153	2.4625	3.7682	1.2818
L139	1.0069	1.7984	0.8071	1.2538
C142	-0.2745	4.7157	21.1483	1.1523
Q143	-0.2478	0.0256	0.7070	0.2776
P144	-0.9503	-0.4599	-0.3518	-0.4384
W195	1.5150	2.1409	2.0582	1.9674
K198	-0.0726	0.5124	-0.1892	-0.0299
T202	0.7737	1.8383	0.9999	1.2039
L206	2.4213	3.3909	3.0872	2.5999
I210	0.4973	0.5514	0.2303	0.4796
V211	0.5253	0.9390	0.0635	0.5737
A218	0.0000	0.4105	-0.2361	0.1174
L222	1.3916	1.1779	2.0242	1.0527
F225	0.0526	2.9244	3.4820	0.8519
K226	-2.4516	-0.5168	-0.7238	-1.4955
W228	-0.2922	1.8628	3.1109	0.9119
Q229	0.4777	0.6015	0.4290	-0.0873
R232	-0.9636	2.0876	-0.0072	0.9024
W300	2.0803	2.3355	2.0627	2.0258

Table 10.5a: Oxytocin receptor mutation map (OTR TM3/5/6- D2R TM 3/4/5/6) This map indicates that a mutation of cysteine 142 to arginine or asparagine has potential to interfere with complex formation. Leucine 206, leucine 222, tryptophan 300 and tryptophan 195 also seem to play important roles at the interface of the oxytocin receptor. The higher the energy value, the more positive the energy of the complex becomes and the more destabilizing the residue gets upon mutation.

	Alanine	Asparagine	Arginine	Serine
I130	0.0503	0.0719	-0.1723	0.0119
T134	-0.0231	0.1974	0.9507	-0.3890
M138	1.2457	3.6760	9.0878	2.2425
P139	0.5671	0.4279	0.5330	0.6963
M140	-0.2801	1.3808	0.3394	0.3752
L141	0.3882	3.1829	7.8270	1.5860
T144	-0.5620	0.0858	-0.1994	-0.5048
Y146	-1.0498	1.2609	0.5429	0.1625
S147	-0.5267	1.3865	7.3716	-0.6650
R150	-0.1794	1.0597	0.0000	0.8264
A188	0.0000	0.9990	0.8283	0.4261
V191	-0.5984	1.6855	2.4204	0.8436
Y192	0.5416	1.5236	0.7541	1.3491
I195	1.2212	1.8208	2.1960	1.3751
V200	1.1472	2.2935	1.0070	1.2553
I203	0.2578	0.6940	0.3412	0.6369
V204	0.3472	0.4233	0.0022	0.2222
H264	-0.4888	-0.4596	-0.7937	-0.4597

Table 10.5b: **Dopamine D2 Receptor mutation map (OTR TM3/5/6- D2R TM 3/4/5/6)** When mutated, methionine 138, leucine 141 and serine 147 show the largest shifts of Alanine scan's energy values upon mutation. Also, tyrosine 192, valine 195 and valine 200 are likely to be important for a possible interface generated by Rosetta MPDock.

According to the Alanine Scan, the overall number of amino acids involved in interface formation dropped from 48 before MD to 37 afterwards.

10.6 Alanine scan of the “Memdock interface” after MD

Alanine scan revealed 79 amino acids participating in the interface formation, as shown in tables 10.6a and 10.6b. OTR residues 83 – 101 as well as D2R 74 – 93 have strictly hydrophobic sidechains and they are positioned in the core of the complex, as they are part of TM2. According to the scoring function, their mutation results in significant loss of stability. Glycine 102 is located in a loop, and its mutation to arginine causes astronomical changes of energy.

	Alanine	Asparagine	Arginine	Serine
E36	-3.1975	0.1319	-0.6543	-1.3434
A37	0.0000	0.5292	0.0484	0.3338
A39	0.0000	0.8282	1.4669	-0.0096
R40	-0.2555	0.3572	0.0000	0.1957
V43	0.4071	0.5042	2.3395	0.3590
L50	1.5609	2.1710	1.6684	1.8103
L54	1.6584	1.6753	1.7629	1.6719
A58	0.0000	0.2728	0.8661	0.0367
L61	-0.3262	1.8922	2.7420	-0.0442
L62	0.5006	1.4786	1.6063	-0.3453
R65	-3.1552	0.2379	0.1367	-1.1036
T66	0.8768	1.7551	5.3668	0.9084
K79	0.0729	0.0950	0.6586	0.3429
H80	-0.7071	-0.3358	1.5950	-0.3328
S82	0.0000	0.0000	0.0788	0.0000
I83	1.2180	3.4089	2.7122	2.0005
L86	0.4706	0.5324	2.3797	0.4512
V87	1.5801	2.4455	1.4124	1.8923
V90	1.1909	3.1230	5.9951	1.6604
F91	0.3924	0.4224	0.3807	0.3920
V93	0.6488	2.9203	8.4704	2.0362
L94	1.6422	2.8048	4.0113	2.3014
L97	1.5612	3.1349	3.0628	1.8994
L98	1.0874	1.4574	1.9682	0.9984
I101	1.2714	3.9566	7.6360	2.4602
T102	-1.9279	1.5728	1.3040	-0.4615
F103	-0.2635	0.1734	-0.2892	-0.0392
F105	1.5849	1.9076	4.9188	1.6395
Y106	0.4767	4.1358	3.6154	2.8980
G107	-0.6411	13.3603	138.0973	1.7563
P108	-2.5403	1.0619	4.9884	-1.7346
L111	0.2517	0.4024	0.5742	0.3216
R154	-3.2883	-1.1880	0.0000	-3.0595
L155	0.0093	0.0084	0.0060	0.0093
V157	-0.0128	1.2047	2.7057	-0.2227
L158	-1.8696	1.3216	0.7551	-0.1873
W161	-0.2206	-0.1034	-0.0028	-0.0080
F185	0.0193	0.0193	0.0193	0.0193
F344	0.0207	0.0008	0.0101	-0.0184
L345	-0.2330	0.9039	0.0787	0.0936

Table 10.6a: Oxytocin receptor mutation map / OTR 3/5/6 - D2R 3/4/5/6 interface:
The residues 83 – 101 of TM2 are located at the center of the complex and their mutation leads to positive energy shifts. This indicates that the wildtype residues fit into their environment, which can be interpreted as a hint for good shape complementarity, especially when compared with mutations to Alanine. Glycine 107 is located at ECL1 and may lead to complex destabilization upon mutation.

	Alanine	Asparagine	Arginine	Serine
P32	-0.2159	0.4973	0.4039	0.0689
H33	-0.0855	0.0017	-0.0018	0.0016
Y34	0.4987	2.3122	0.8975	1.0708
N35	0.0556	0.0558	0.0557	0.0556
A38	-0.8575	5.9649	11.4321	2.2706
T39	0.0841	-0.2629	0.0013	-0.2708
T42	0.0258	1.2349	0.8753	0.5290
L43	0.0005	0.0005	0.0002	0.0005
A46	0.2862	0.0751	-0.0820	0.0623
V49	0.3292	0.5692	0.6054	0.3248
V53	1.6987	2.3385	1.0557	2.1890
C56	-0.4060	1.2584	3.8972	0.5574
M57	0.5216	2.0027	1.0759	1.0604
S60	-1.4148	3.6434	17.2983	0.0000
Q66	-1.0057	-0.2446	1.3668	-0.9616
Y71	-0.7790	1.1495	1.8992	0.1215
V74	-0.0848	0.4969	2.5480	0.1208
V78	1.2306	1.9919	1.6003	1.4001
L81	0.5438	3.5750	3.2508	1.0929
L82	1.1189	0.6041	2.4194	1.3034
A84	0.0000	0.0468	-0.0218	-0.0192
T85	1.2162	4.5640	20.8021	2.5708
L86	1.3124	2.0565	2.6439	1.5522
P89	0.1732	3.4122	2.7244	1.1263
W90	0.5949	1.3987	1.8968	0.2695
V92	0.5802	0.5309	0.8997	0.8574
Y93	1.3241	1.7115	3.0891	1.7694
F102	-0.3329	0.1349	0.6154	0.0538
S103	-0.9551	1.3575	0.5929	0.0000
I105	1.0523	2.3113	1.6766	1.7247
H106	-1.8553	-0.7297	5.1750	-1.6255
I109	1.6479	1.1089	1.5995	1.0210
K149	-0.0017	-0.0009	0.0252	-0.0015
T153	-0.9604	1.4276	-0.3541	0.2047
I156	0.8018	1.6317	2.9408	1.0881
S157	-0.2030	0.1953	0.2627	0.0000
K305	-1.3608	-0.9175	0.7152	-1.2074
L307	-0.5119	-0.5011	-0.5175	-0.5003
H308	-2.8099	3.3251	6.2216	0.0311

Table 10.6b: Dopamine D2 receptor mutation map / OTR 1/2/3/-D2R 1/2/3 interface

Similar to OTR, TM2 is contributing to the interface with its hydrophobic residues 78 – 93. Alanine 38, serine 60 and tyrosine 85 give high changes in energy score upon mutation.

10.7 Pymol visual inspection: Rosetta MPDock interface

The last frame of the MD was visualized with Pymol. As before, the most important residues found by the Alanine scan where colored accord to their energy shift upon mutation.The interface comprises reciprocal TM 5/6 interactions of both partners in the upper region (Fig 10.7b and 10.7c), TM5-TM5 contacts in the middle section (Fig 10.7d) and TM3-TM3 interactions at the bottom, which also involve D2R's TM4 and intracellular loop 2 as well as OTRs TM5 (Fig 10.7f and Fig 10.7g). Pymol found one hydrogen between glutamine 143 of OTR and the backbone of D2R methionine 140 (Fig 10.7e). Due to small conformational changes of the harboring transmembrane helices, the hydrogen bond between OTR glutamine 229 and D2R serine 147 that was observed before MD was not there anymore after MD. Also, the density of hydrophobic residues in the center of the membrane dropped, resulting in a total of 37 interactions, as discussed before (overview in figure 10.7a).

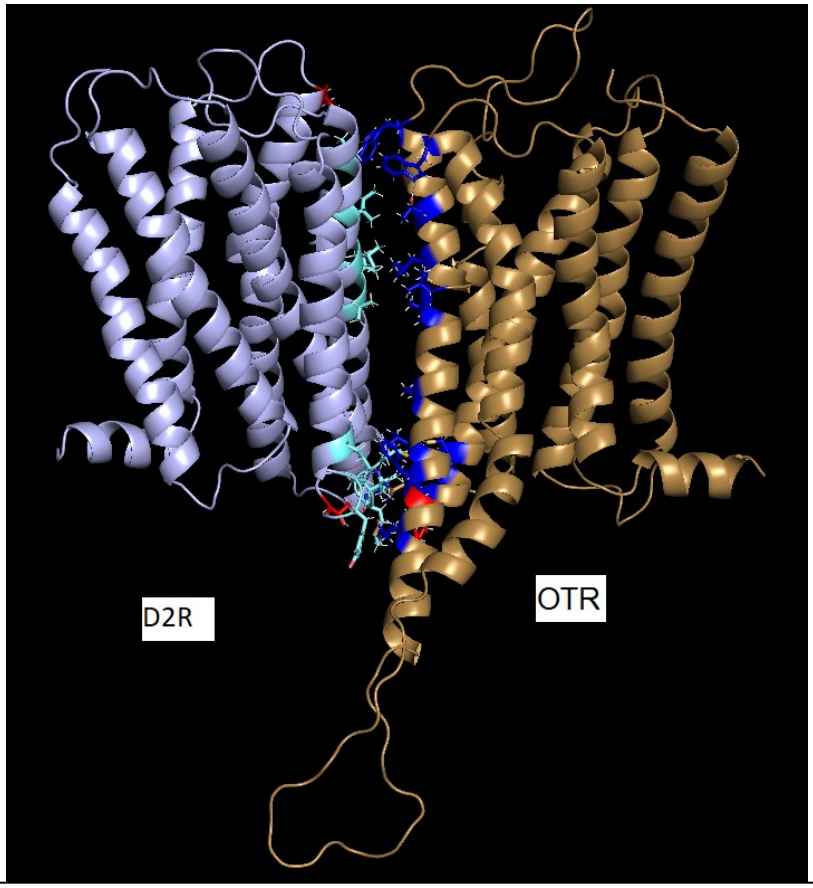


Fig. 10.7a: Total view of the Rosetta MPDock interface after MD.
Some amino acid contacts were lost according to Alanine Scan, especially in the mid-section of the protein.

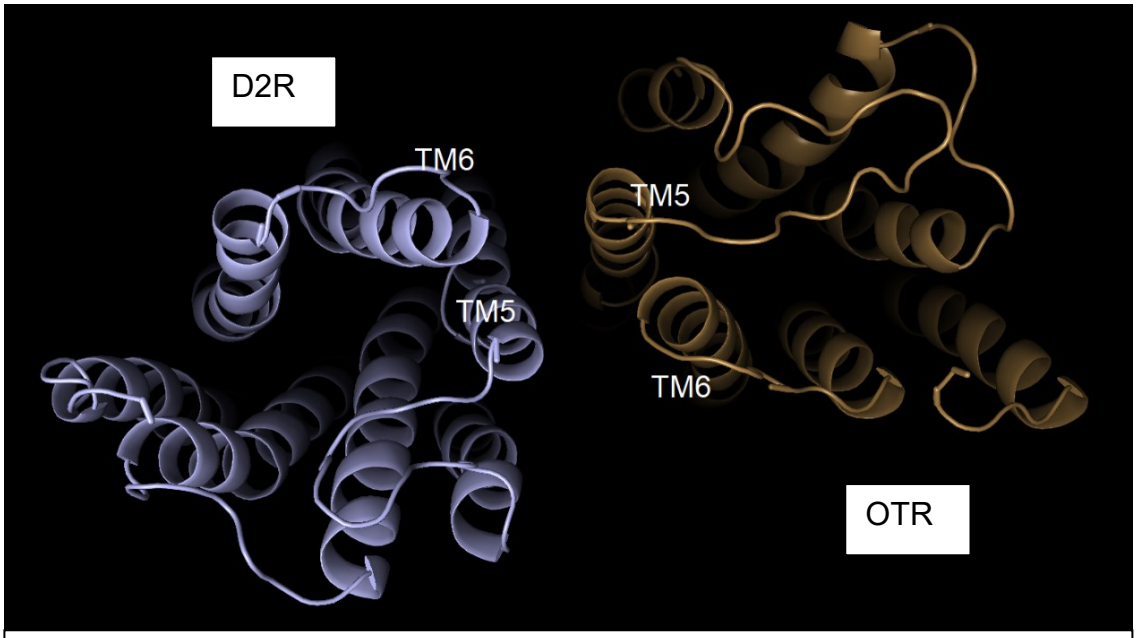


Fig. 10.7b: Top view of the Rosetta MPDock interface after MD
It involves residues located on the top of TM5 and TM6 of both receptors.

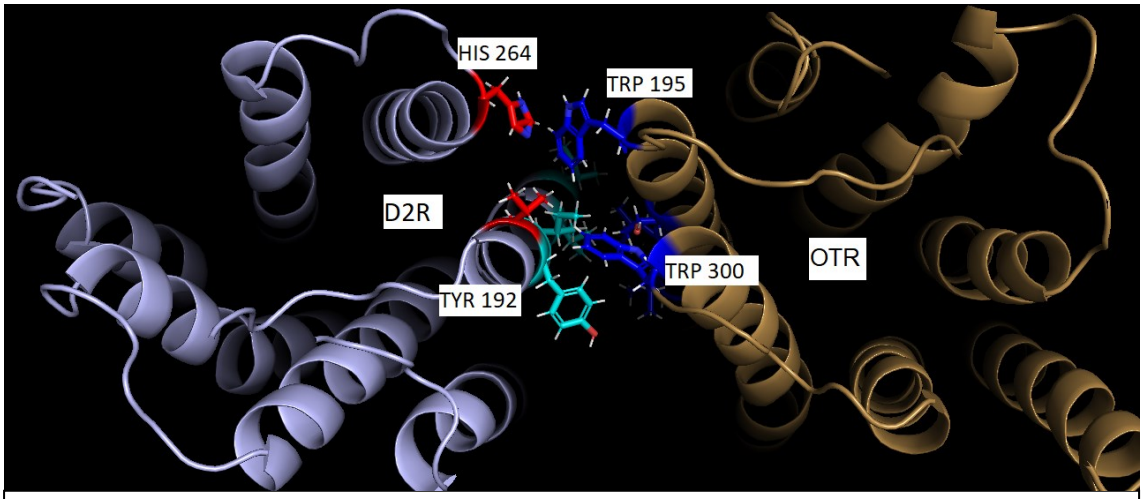


Fig 10.7c: Involved residues of the top section.
D2R histidine 264 (TM6) is close to tryptophane 195 of OTR (TM5), also tyrosine 192 of D2R (TM5) is in proximity of OTR tryptophane 300 (TM6).

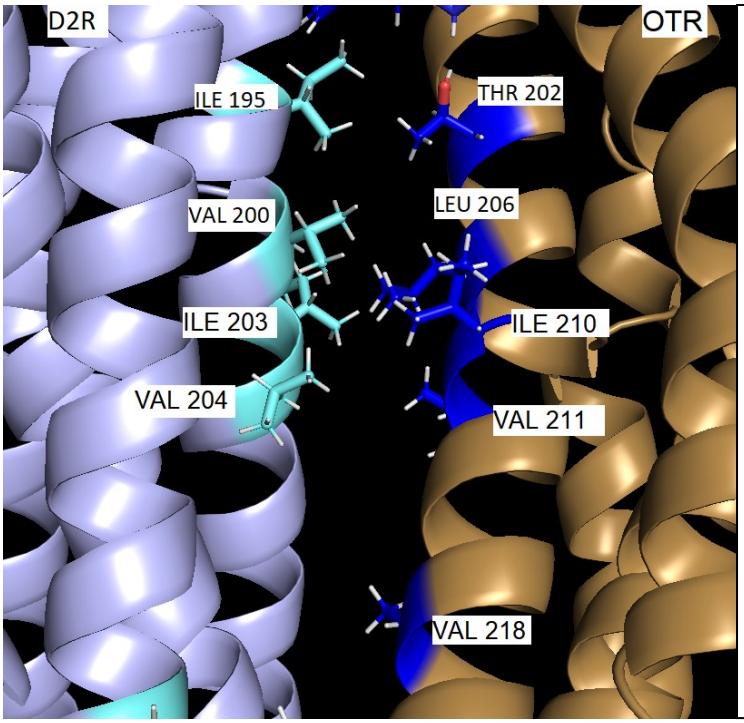


Fig 10.7d: In the middle section some hydrophobic residues of both TM5 segments engage in hydrophobic interaction, forming the core of the interface. Compared to before the MD, some contacts got lost in the lower middle section.

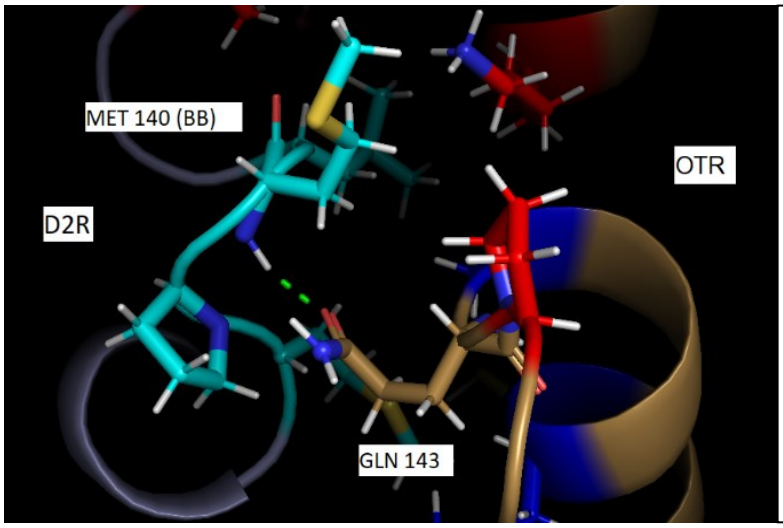


Fig 10.7e:
H-Bond at the bottom, formed between glutamine 143 of OTR and the backbone of methionine 140.

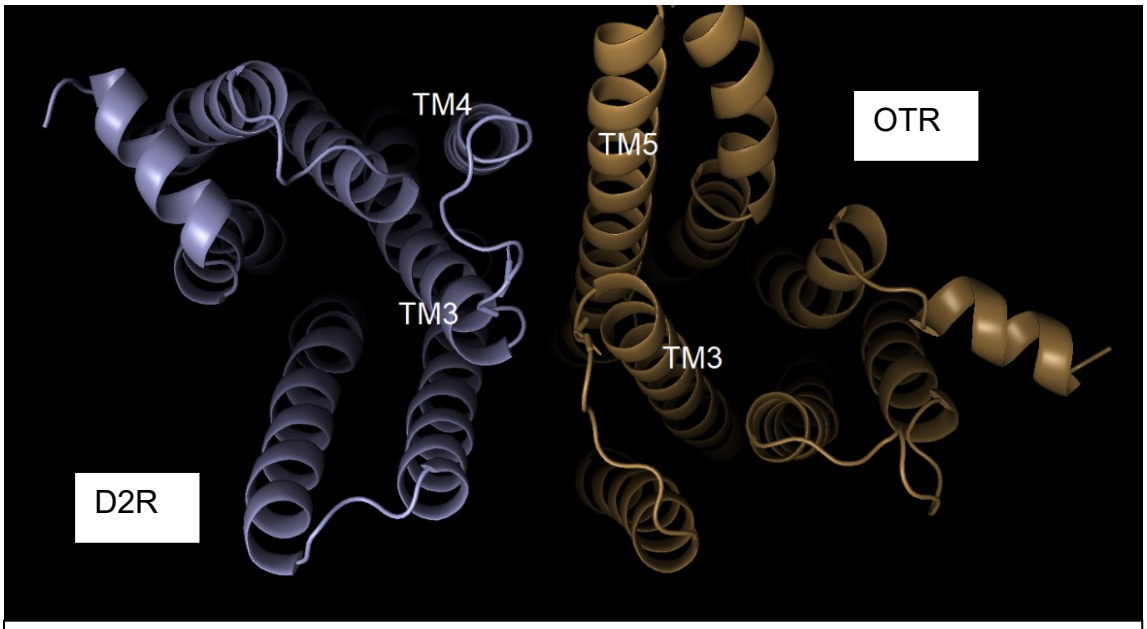


Fig. 10.7f: Bottom view of the Rosetta MPDock result after MD. Inter chain contact takes place between OTR TM3/TM5 and D2R's TM4 / intracellular loop 2 (ICL2).

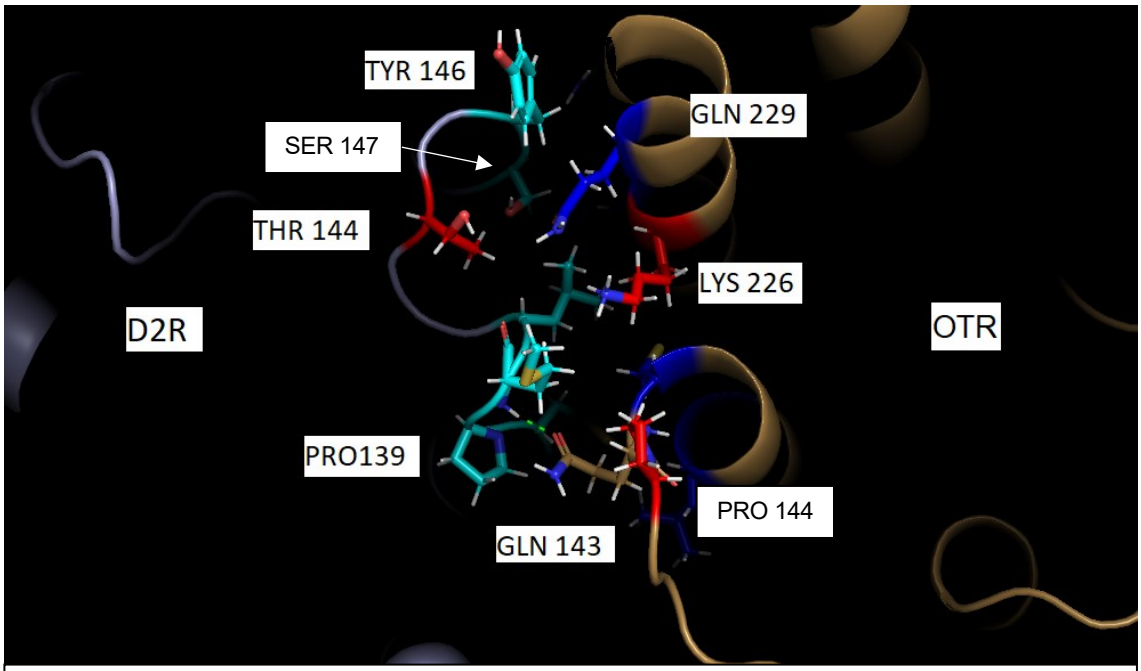


Fig. 10.7g: Residues on the bottom of the interface, including glutamine 229 of OTR and serine 147, which interacted via H-bond before MD.

10.8 Pymol visual inspection: Memdock interface

As before, Pymol was used to visually prepare the features of the Memdock interface after molecular dynamics simulation (overview in figure 10.8a). Pymol found a total of 9 hydrogen bonds, scattered across the whole structure (two at the top, two at the bottom, three at helix 8 of D2R and two at the helix 8 of OTR). 79 interacting residues were previously detected by PyRosetta Alanine scan, the most important of which were, again, colored correspondingly. At the extracellular upper region, TM2/TM2 interactions take place, flanked by TM1 and TM3 of both partners in close proximity (figure 10.8b). In the middle section, TM1 and TM2 of both D2R and OTR form the hydrophobic core of the complex (figure 10.8c). At the bottom, both helices 8 have hydrogen bonds with the opposite TM4 (figure 10.8d), while in the center TM1 of D2R is connected to OTR TM1 and TM2 by hydrogen bonds (figure 10.8.e).

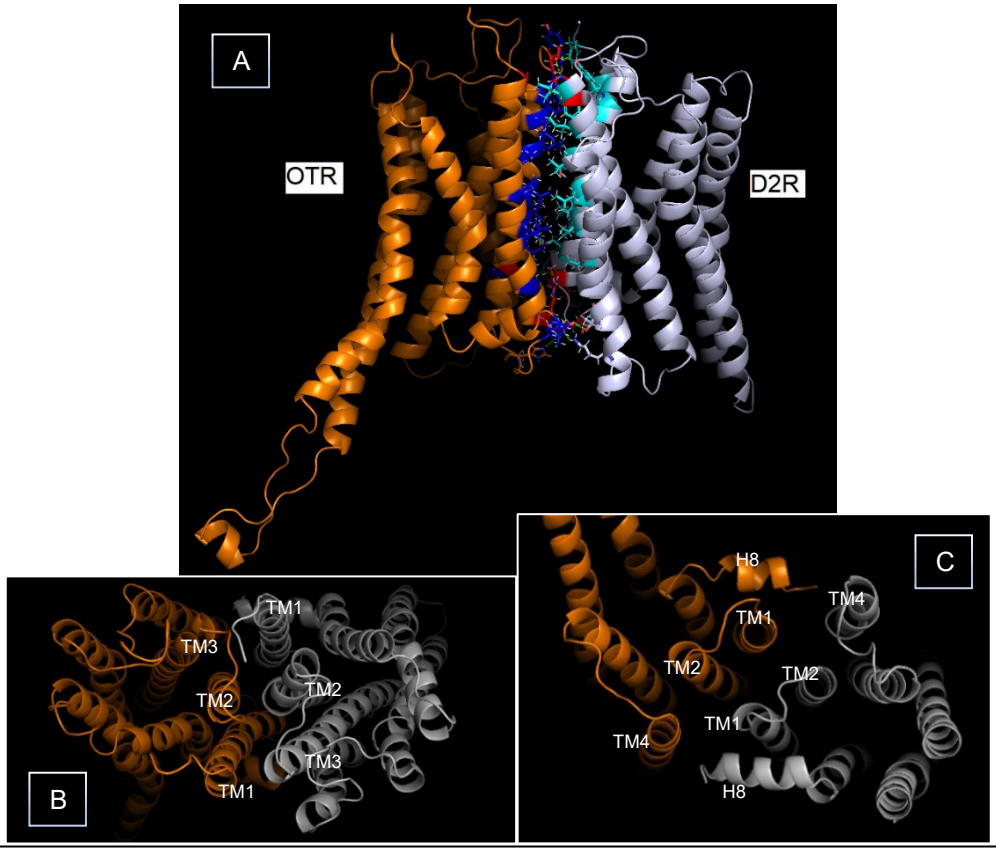


Fig. 10.8a: A: total view of the Memdock result after MD. As shown in **B**, at the extracellular top of the complex, interactions between TM2 of both OTR and D2R take place, flanked by TM1-TM3 interactions. **C**, from the bottom view, additionally shows the vicinity of both Helix 8-domains with the opposite's TM3 at the intracellular side of the structure.

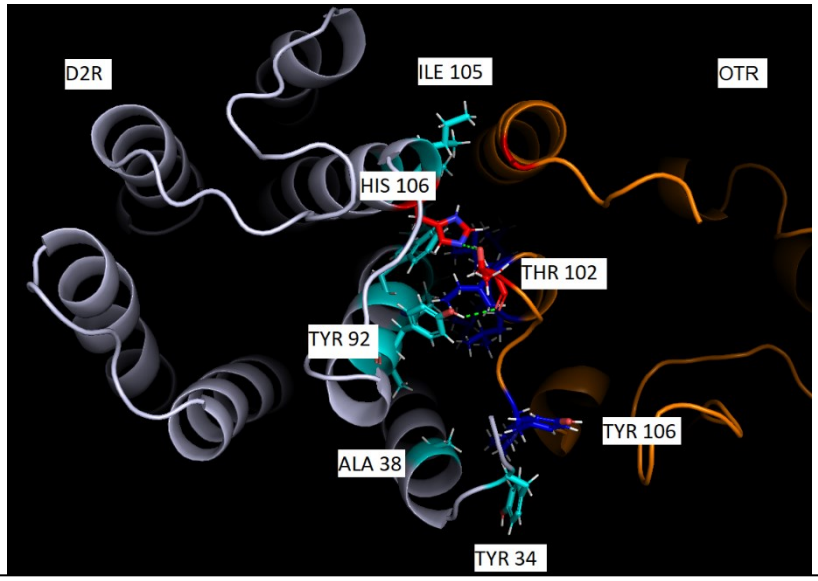


Fig. 10.8b: top view of the Memdock result with interacting amino acids. OXTR's Threonine 102 is involved in two intermolecular H-bonds: with D2R tyrosine 92 (via backbone) and also with histidine 106. D2R Tyrosine 34, part of the N-terminal domain, is in proximity of extracellular loop 1 (ECL1) of OXR, where tyrosine 106 resides.

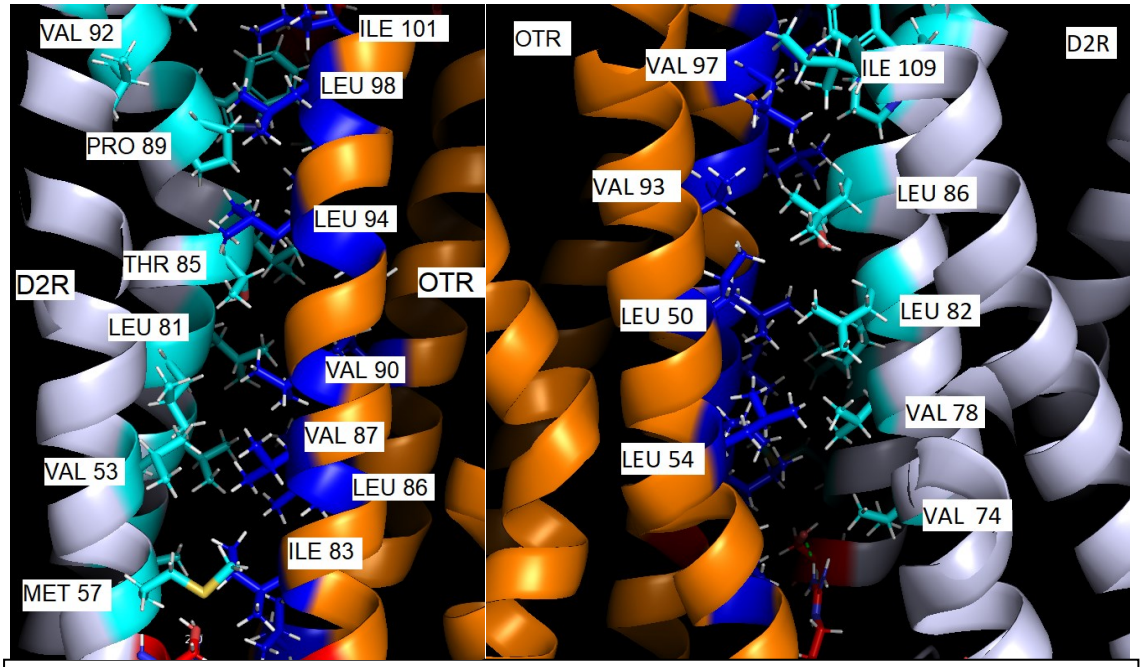


Fig. 10.8c: At the inner region of the complex, a lot of hydrophobic interaction is going on between D2R's TM1/TM2 and OTR's TM1/TM2 (in all possible combinations), but also isoleucine 109 of D2R-TM3 facing TM1 of OTR.

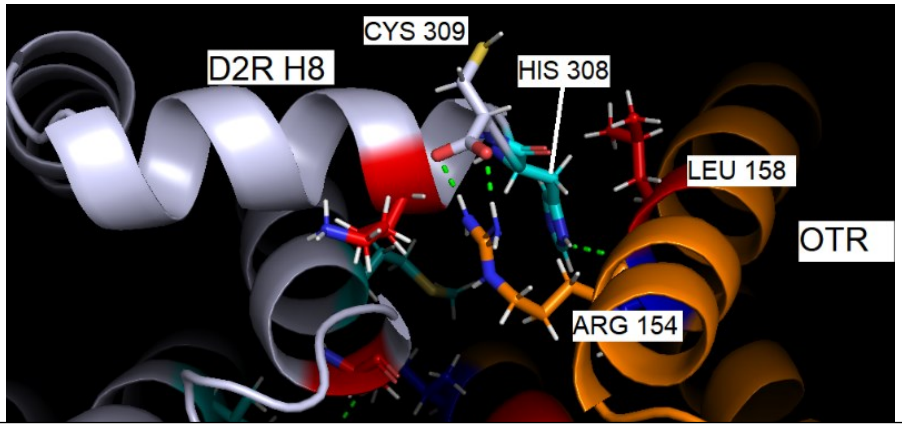


Fig 10.8d: A network of three hydrogen bonds was observed at arginine 154 of OTR; interacting with histidine 308 D2R (via backbone of arginine 154) and cysteine 309 (C-terminus). Such interactions, including two hydrogen bonds, can also be found at Helix 8 of OTR (not shown).

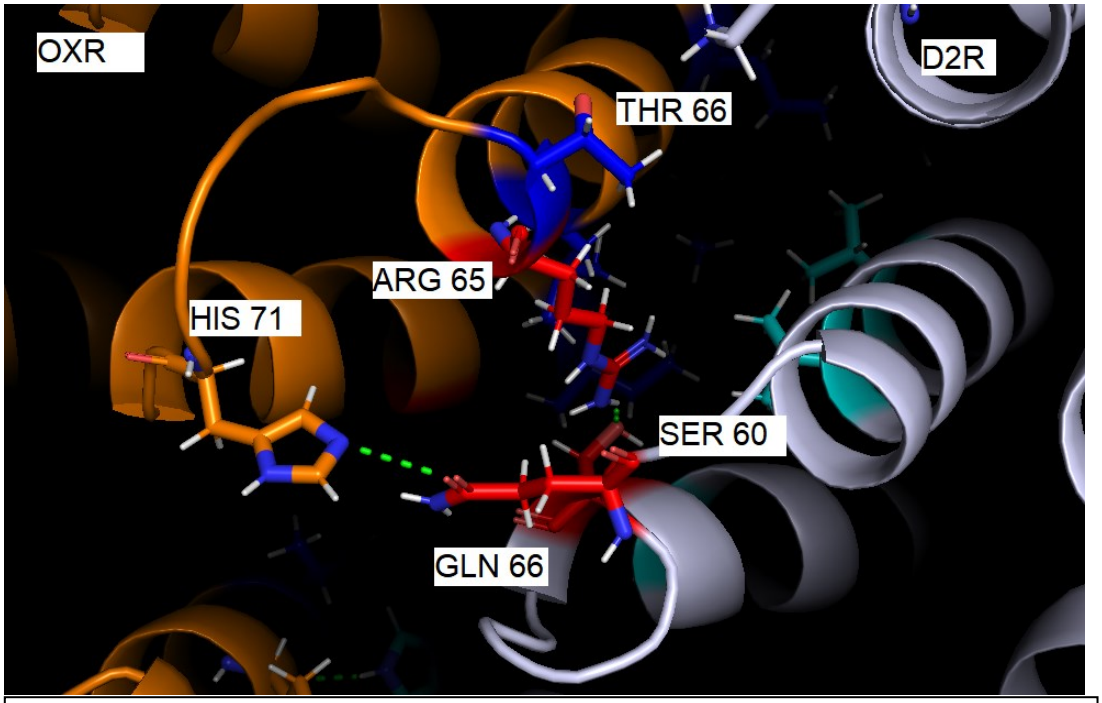


Fig 10.8e: At the bottom center, two hydrogen bonds are formed between OTR arginine 65 with D2R serine 60 and OTR histidine 71 with D2R glutamine 66.

11. DISCUSSION

The Rosetta MPDock docking program, followed by bivalent linker docking and molecular dynamics simulation, predicted a structure with an interface involving a TM5/TM6 – TM5/TM6 interaction at the top, TM5-TM5 in the middle and TM3/4 of D2R having contact with OTR TM3/5. The MD simulation led to decrease in the number of interface amino acids (37 instead of 48) and to decrease of buried surface area by ~400Å² (from 2567 to 2168Å²), implying that the complex was not as stable as, for instance, the Memdock result. The whole complex gained 2380Å² in surface area, making conformational changes very probable. After docking there was one H-bond that was lost, but also one that was won upon MD. On the other hand, although not so stable, the OTR TM3/5/6 – D2R TM 3/4/5/6 interface predicted by Rosetta MPDock showed excess of total surface area change upon MD. This raises the question if this can be considered a conformational change that exposes more residues to the environment, including potential ligands. This could in turn can be interpreted as allosteric effect, on the other hand it could also mean that the complex is instable.

The best docking result of the Memdock server, with docked bivalent linker and after MD, harbored an interface with the symmetrical involvement of TM1/TM2/TM3 at the top of the structure, multiple interactions of TM1/TM2 in the hydrophobic core and at the bottom, and additionally reciprocal helix8 interactions with TM4. The BSA was increased by ~700Å² from 3479 to 4182 Å², meaning that the MD had a stabilizing effect on the conformation of the complex. The overall accessible surface area increased by only by about ~400 Å² throughout the MD. The TM1/2/3/4 interface predicted by Memdock impresses with the number of participating residues and H-bonds, moreover, its sheer size and geometrical reciprocity. 79 amino acids were found to be part of the interface, involving 9 H-bonds. The symmetrical arrangement with its huge interface would rather fit an obligate homodimer assembly than a reversible heterodimer complex: Homodimers tend to depend mainly on hydrophobic interactions, and they have larger, often symmetrical interfaces (Jones and Thornton, 1996). The structural homology of D2R and OTR may enable this “intimate embrace” of

confused receptors that behave like homodimers. The problem with that hypothesis is that this would undermine specificity of complex formation and the evolutionary purpose of the heterodimer, provided that the homodimers share the same interface with the heterocomplexes. Fortunately, this can to a good extent be ruled out, since the structure of homodimers of D2R and OTR have been predicted (Kaczor et al., 2016; Busnelli et al., 2016), and the interfaces that were found do differ from the predictions of this work. OTR homodimers observed by Busnelli et al. also comprised a symmetrical interface including TM1, TM2 and helices 8. However, helices 8 have direct contact at the bottom center and are arranged in an antiparallel fashion, which clearly differs from the OTR-D2R interface generated with Memdock, where helices 8 are located at the outside of the complex. D2R homodimers, according to Kaczor et al., form a TM4/5 -TM1/7 interface, which is also not in conflict with interfaces predicted neither by Rosetta MPDock nor by Memdock.

The results of both docking methods are to some extent symmetrical. In fact, symmetrical interfaces have been proposed for many GPCR homodimer and heterodimer complexes such as the dopamine D2 receptor - adenosine A2 receptor complex, comprising reciprocal contacts between TM4 and TM5 (Borroto-Escuela et al., 2018). The phenomenon of allostery in oligomers (with identical or similar monomers) is often associated with symmetry (Goodsell and Olson, 2000). The fact that facilitatory allosteric effects observed in OTR-D2R complexes were also reciprocal allows one to speculate that these effects are made possible by the symmetrical nature of their interface orientation.

The structures of complexes that were found in this docking experiments cannot give definitive answers to the question, how the interface of the OTR-D2R complex actually looks like. The output of the Rosetta MPDock scoring functions did not allow to pick a “winner” among the created start structures that would qualify for being at the global energy minimum. The molecular dynamics simulation afterwards was probably too short to get near the global minima. Memdock, on the other hand, did clearly prefer one orientation, but the size of the interface is on the edge of credibility for a heterodimer complex. Also, it may

be over stabilized to account for conformational changes necessary for reciprocal allosteric behavior.

For further investigation, the most important question, in the humble opinion of the author, is about the flexible regions that were not modeled (ICL3 of D2R, N-termini) and their part in the protein-protein interaction. If they play a role, contact between the helices might not even be necessary. This was observed in heterodimerization studies of dopamine D1 and D2 receptors (O'Dowd et al., 2012). Also, BRET-assays with both D2R and OTR and fragments of the helices proposed to be part of the interface, just as performed by Busnelli et al. for the determination of the OTR-homodimer interface, could be carried out to test if the helix fragments are capable of disrupting the OTR-D2R complex formation. If they do, their participation in the interface might be plausible. Also, the Alanine Scan results may point to worthwhile targets for mutation in advance to BRET experiments: if mutation of residues with high positive energy shift lead to disruption of complex formation, this could indicate correctness of the corresponding interface.

12. ABSTRACTS

English version

Investigation Of The Interface Between Human Oxytocin Receptor And Human Dopamine D2 Receptor

Abstract:

Dopamine D2 receptor (D2R) and oxytocin receptor (OTR) are membrane bound G-protein coupled receptors (GPCRs) that are known be able to form heteromer complexes (Romero-Fernandez et al., 2012). As these complexes can be involved in function and disease of the brain, they are considered viable drug targets (De la Mora et al., 2016, Borroto-Escuela et al., 2017). Since there are crystal structures of the individual partners (Waltenspühl et al., 2020; Yin et al., 2020), but not of the co-crystallized complex, little is known about the three-dimensional structure of the OTR-D2R heteromer. Employing the protein-protein docking technique, this work attempts to elucidate structural information of the OTR-D2R interface. Two docking programs, in particular the Rosetta MPDock docking protocol and the Memdock server, were used to generate models of the complex. The best results were then docked with a bivalent ligand of a defined length. This was followed by a molecular dynamics simulation, determining stability and behavior of the structure in a membrane environment. The complex predicted by Rosetta MPDock was pointing to an OTR transmembrane helix (TM) 3/5/6 – D2R TM 3/4/5/6 interface, featuring 37 interacting amino acids and one hydrogen bond. The result of the Memdock prediction harbored a symmetric interface with reciprocal involvement of TM 1/2/3/4 and helix 8 of both receptors, including 79 participating amino acids and 9 hydrogen bonds. Whether any of the predicted structures actually correspond to the real-world global minimum of the complex remains uncertain, given unmodeled flexible regions and the shortness of the molecular dynamics simulation.

German version

Untersuchung des Interface zwischen dem humanen Oxytocin Rezeptor und dem humanen Dopamin D2 Rezeptor.

Zusammenfassung:

Der Dopamin-D2-Rezeptor (D2R) und der Oxytocin-Rezeptor (OTR) sind membrangebundene G-Protein gekoppelte Rezeptoren, von denen man weiß, dass sie Heteromer-Komplexe bilden können (Romero-Fernandez et al., 2012). Weil diese Komplexe in Funktion und Erkrankungen des Gehirns eine Rolle spielen können, eignen sie sich als Ziele bei der Entwicklung neuer Arzneimittel (De la Mora et al., 2016, Borroto-Escuela et al., 2017). Es gibt zwar Kristallstrukturen der einzelnen Rezeptoren (Waltenspühl et al., 2020; Yin et al., 2020), nicht jedoch im Komplex, daher weiß man wenig über die 3D-Struktur der OTR-D2R Heteromere. Protein-Protein-Docking wurde angewandt, um darüber mehr über das Interface dieses Komplexes in Erfahrung zu bringen. Zwei verschiedene Docking-Programme wurden zur Modellierung verwendet: das Rosetta MPDock-Docking-Protokoll und der Memdock-Server. In die jeweils besten Resultate wurde anschließend ein bivalenter Ligand mit definierter Länge gedockt. Darauf folgte eine Moleküldynamik-Simulation, mit der die Stabilität und das Verhalten der Struktur in einer Membranumgebung bestimmt wurden. Der von Rosetta MPDock vorhergesagte Komplex deutete auf eine Beteiligung von den Transmembranhelices (TM) 3/5/6 von OTR, und TM 3/4/5/6 bei D2R am Interface hin. 37 Aminosäuren und eine Wasserstoffbrücke sind daran beteiligt. Das Resultat von Memdock beherbergte ein symmetrisches Interface unter reziproker Beteiligung von TM 1/2/3/4 und Helix 8. 79 Aminosäuren und 9 Wasserstoffbrücken waren involviert. Ob eine der ermittelten Strukturen tatsächlich dem globalen Minimum des Komplexes entspricht, bleibt aufgrund nicht modellierbarer flexibler Regionen und einer zu kurzen Zeitspanne der Moleküldynamiksimulation ungewiss.

12. REFERENCES

Alford, Rebecca F., et al. “An Integrated Framework Advancing Membrane Protein Modeling and Design.” *PLOS Computational Biology*, vol. 11, no. 9, 2015, doi:10.1371/journal.pcbi.1004398.

Arunan, Elangannan, et al. “Definition of the Hydrogen Bond (IUPAC Recommendations 2011).” *Pure and Applied Chemistry*, vol. 83, no. 8, 2011, pp. 1637–1641., doi:10.1351/pac-rec-10-01-02.

Baskerville, Tracey A., and Alison J. Douglas. “Dopamine and Oxytocin Interactions Underlying Behaviors: Potential Contributions to Behavioral Disorders.” *CNS Neuroscience & Therapeutics*, vol. 16, no. 3, 2010, doi:10.1111/j.1755-5949.2010.00154.x.

Borroto-Escuela, Dasiel O., et al. “Mapping the Interface of a GPCR Dimer: A Structural Model of the A2A Adenosine and D2 Dopamine Receptor Heteromer.” *Frontiers in Pharmacology*, vol. 9, 2018, doi:10.3389/fphar.2018.00829.

Borroto-Escuela, Dasiel O., et al. “Understanding the Role of GPCR Heteroreceptor Complexes in Modulating the Brain Networks in Health and Disease.” *Frontiers in Cellular Neuroscience*, vol. 11, 2017, doi:10.3389/fncel.2017.00037.

Brooks, B. R., et al. “CHARMM: The Biomolecular Simulation Program.” *Journal of Computational Chemistry*, vol. 30, no. 10, 2009, pp. 1545–1614., doi:10.1002/jcc.21287.

Busnelli, Marta, et al. “Design and Characterization of Superpotent Bivalent Ligands Targeting Oxytocin Receptor Dimers via a Channel-Like Structure.” *Journal of Medicinal Chemistry*, vol. 59, no. 15, 2016, pp. 7152–7166., doi:10.1021/acs.jmedchem.6b00564.

Chang, C.-e. A., et al. “Ligand Configurational Entropy and Protein Binding.” *Proceedings of the National Academy of Sciences*, vol. 104, no. 5, 2007, pp. 1534–1539., doi:10.1073/pnas.0610494104.

Chaudhury, Sidhartha, et al. “Benchmarking and Analysis of Protein Docking Performance in Rosetta v3.2.” *PLoS ONE*, vol. 6, no. 8, 2011, doi:10.1371/journal.pone.0022477.

Cho, DI, et al. “The N-Terminal Region of the Dopamine D2 Receptor, a Rhodopsin-like GPCR, Regulates Correct Integration into the Plasma Membrane and Endocytic Routes.” *British Journal of Pharmacology*, vol. 166, no. 2, 2012, pp. 659–675., doi:10.1111/j.1476-5381.2011.01787.x.

Chothia, Cyrus, and Joël Janin. “Principles of Protein–Protein Recognition.” *Nature*, vol. 256, no. 5520, 1975, pp. 705–708., doi:10.1038/256705a0.

Chuang, Gwo-Yu, et al. “DARS (Decoys As the Reference State) Potentials for Protein-Protein Docking.” *Biophysical Journal*, vol. 95, no. 9, 2008, pp. 4217–4227., doi:10.1529/biophysj.108.135814.

De la Mora, Miguel Pérez, et al. “Signaling in Dopamine D2 Receptor-Oxytocin Receptor Heterocomplexes and Its Relevance for the Anxiolytic Effects of Dopamine and Oxytocin Interactions in the Amygdala of the Rat.” *Biochimica Et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1862, no. 11, 2016, pp. 2075–2085., doi:10.1016/j.bbadis.2016.07.004.

De Las Rivas, Javier, and Celia Fontanillo. “Protein–Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks.” *PLoS Computational Biology*, vol. 6, no. 6, 2010, doi:10.1371/journal.pcbi.1000807.

“Die Chemischen Bestandteile Der Zelle.” *Lehrbuch Der Molekularen Zellbiologie*, by Bruce Alberts et al., 2nd ed., Wiley-VCH, 1998, pp. 74–74.

Dolinsky, T. J., et al. “PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations.” *Nucleic Acids Research*, vol. 32, no. Web Server, 2004, doi:10.1093/nar/gkh381.

Eastman, Peter, et al. “OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics.” *PLOS Computational Biology*, vol. 13, no. 7, 2017, doi:10.1371/journal.pcbi.1005659.

Eisenhaber, Frank. “Hydrophobic Regions on Protein Surfaces. Derivation of the Solvation Energy from Their Area Distribution in Crystallographic Protein Structures.” *Protein Science*, vol. 5, no. 8, 1996, pp. 1676–1686., doi:10.1002/pro.5560050821.

Fink, Avner, et al. “Transmembrane Domains Interactions within the Membrane Milieu: Principles, Advances and Challenges.” *Biochimica Et Biophysica Acta (BBA) - Biomembranes*, vol. 1818, no. 4, 2012, pp. 974–983., doi:10.1016/j.bbamem.2011.11.029.

Fuxe, Kjell, et al. “Receptor-Receptor Interactions in Heteroreceptor Complexes: a New Principle in Biology. Focus on Their Role in Learning and Memory.” *Neuroscience Discovery*, vol. 2, no. 1, 2014, p. 6., doi:10.7243/2052-6946-2-6.

Gimpl, Gerald, and Falk Fahrenholz. “The Oxytocin Receptor System: Structure, Function, and Regulation.” *Physiological Reviews*, vol. 81, no. 2, 2001, pp. 629–683., doi:10.1152/physrev.2001.81.2.629.

Goodsell, David S., and Arthur J. Olson. “Structural Symmetry and Protein Function.” *Annual Review of Biophysics and Biomolecular Structure*, vol. 29, no. 1, 2000, pp. 105–153., doi:10.1146/annurev.biophys.29.1.105.

Gray, Jeffrey J., et al. “Protein–Protein Docking with Simultaneous Optimization of Rigid-Body Displacement and Side-Chain Conformations.” *Journal of Molecular Biology*, vol. 331, no. 1, 2003, pp. 281–299., doi:10.1016/s0022-2836(03)00670-3.

Gurevich, Vsevolod V., and Eugenia V. Gurevich. “How and Why Do GPCRs Dimerize?” *Trends in Pharmacological Sciences*, vol. 29, no. 5, 2008, pp. 234–240., doi:10.1016/j.tips.2008.02.004.

Hollingsworth, Scott A., and Ron O. Dror. “Molecular Dynamics Simulation for All.” *Neuron*, vol. 99, no. 6, 2018, pp. 1129–1143., doi:10.1016/j.neuron.2018.08.011.

Humphrey, William, et al. “VMD: Visual Molecular Dynamics.” *Journal of Molecular Graphics*, vol. 14, no. 1, 1996, pp. 33–38., doi:10.1016/0263-7855(96)00018-5.

Hurwitz, Naama, et al. “Memdock: an α-Helical Membrane Protein Docking Algorithm.” *Bioinformatics*, vol. 32, no. 16, 2016, pp. 2444–2450., doi:10.1093/bioinformatics/btw184.

Hübner, Harald, et al. “Structure-Guided Development of Heterodimer-Selective GPCR Ligands.” *Nature Communications*, vol. 7, no. 1, 2016, doi:10.1038/ncomms12298.

Jo, Sunhwan, et al. “Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations.” *PLoS ONE*, vol. 2, no. 9, 2007, doi:10.1371/journal.pone.0000880.

Jo, Sunhwan, et al. “CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM.” *Journal of Computational Chemistry*, vol. 29, no. 11, 2008, pp. 1859–1865., doi:10.1002/jcc.20945.

Jones, S., and J. M. Thornton. “Principles of Protein-Protein Interactions.” *Proceedings of the National Academy of Sciences*, vol. 93, no. 1, 1996, pp. 13–20., doi:10.1073/pnas.93.1.13.

Kaczor, Agnieszka A., et al. “The Dopamine D2 Receptor Dimer and Its Interaction with Homobivalent Antagonists: Homology Modeling, Docking and Molecular Dynamics.” *Journal of Molecular Modeling*, vol. 22, no. 9, 2016, doi:10.1007/s00894-016-3065-2.

Khechinashvili, Nikolay N., et al. “Thermal Stability of Proteins Does Not Correlate with the Energy of Intramolecular Interactions.” *Biochimica Et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1784, no. 11, 2008, pp. 1830–1834., doi:10.1016/j.bbapap.2008.07.008.

Koehbach, Johannes, et al. “Insights into the Molecular Evolution of Oxytocin Receptor Ligand Binding.” *Biochemical Society Transactions*, vol. 41, no. 1, 2013, pp. 197–204., doi:10.1042/bst20120256.

Kolb, Hartmuth C., et al. “Click Chemistry: Diverse Chemical Function from a Few Good Reactions.” *Angewandte Chemie International Edition*, vol. 40, no. 11, 28 May 2001, pp. 2004–2021., doi:https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.

Kortemme, T., and D. Baker. “A Simple Physical Model for Binding Energy Hot Spots in Protein-Protein Complexes.” *Proceedings of the National Academy of Sciences*, vol. 99, no. 22, 2002, pp. 14116–14121., doi:10.1073/pnas.202485799.

Kozakov, Dima, et al. “The ClusPro Web Server for Protein–Protein Docking.” *Nature Protocols*, vol. 12, no. 2, 2017, pp. 255–278., doi:10.1038/nprot.2016.169.

Lengauer, Thomas, and Matthias Rarey. “Computational Methods for Biomolecular Docking.” *Current Opinion in Structural Biology*, vol. 6, no. 3, 1996, pp. 402–406., doi:10.1016/s0959-440x(96)80061-3.

Liu, Shiyong, and Ilya A Vakser. “DECK: Distance and Environment-Dependent, Coarse-Grained, Knowledge-Based Potentials for Protein-Protein Docking.” *BMC Bioinformatics*, vol. 12, no. 1, 2011, doi:10.1186/1471-2105-12-280.

Lomize, Mikhail A., et al. “OPM Database and PPM Web Server: Resources for Positioning of Proteins in Membranes.” *Nucleic Acids Research*, vol. 40, no. D1, 2011, doi:10.1093/nar/gkr703.

Mason, Aaron C., and Jan H. Jensen. “Protein-Protein Binding Is Often Associated with Changes in Protonation State.” *Proteins: Structure, Function, and Bioinformatics*, vol. 71, no. 1, 2008, pp. 81–91., doi:10.1002/prot.21657.

Morris, Garrett M., and Marguerita Lim-Wilby. “Molecular Docking.” *Methods in Molecular Biology*, 2008, pp. 365–382., doi:10.1007/978-1-59745-177-2_19.

O'Toole, Nicholas, and Ilya A. Vakser. “Large-Scale Characteristics of the Energy Landscape in Protein–Protein Interactions.” *Proteins: Structure, Function, and Bioinformatics*, vol. 71, no. 1, 2008, pp. 144–152., doi:10.1002/prot.21665.

O’Dowd, Brian F., et al. “Two Amino Acids in Each of D1 and D2 Dopamine Receptor Cytoplasmic Regions Are Involved in D1–D2 Heteromer Formation.” *Biochemical and Biophysical Research Communications*, vol. 417, no. 1, 2012, pp. 23–28., doi:10.1016/j.bbrc.2011.11.027.

Ritter, Stefanie L., and Randy A. Hall. “Fine-Tuning of GPCR Activity by Receptor-Interacting Proteins.” *Nature Reviews Molecular Cell Biology*, vol. 10, no. 12, 2009, pp. 819–830., doi:10.1038/nrm2803.

Romero-Fernandez, W, et al. “Evidence for the Existence of Dopamine d2-Oxytocin Receptor Heteromers in the Ventral and Dorsal Striatum with Facilitatory Receptor–Receptor Interactions.” *Molecular Psychiatry*, vol. 18, no. 8, 2012, pp. 849–850., doi:10.1038/mp.2012.103.

Ruvinsky, Anatoly M., and Ilya A. Vakser. “Sequence Composition and Environment Effects on Residue Fluctuations in Protein Structures.” *The Journal of Chemical Physics*, vol. 133, no. 15, 2010, p. 155101., doi:10.1063/1.3498743.

Salahpour, Ali, and Marc G. Caron. “Food for Thought: The Physiological Relevance of Ghrelin and Dopamine D2 Receptor Heterodimerization in the Regulation of Appetite.” *Neuron*, vol. 73, no. 2, 2012, pp. 210–211., doi:10.1016/j.neuron.2012.01.004.

Schneidman-Duhovny, D., et al. “PatchDock and SymmDock: Servers for Rigid and Symmetric Docking.” *Nucleic Acids Research*, vol. 33, no. Web Server, 2005, doi:10.1093/nar/gki481.

Schrödinger LLC . “The PyMOL Molecular Graphics System, Version 2.4.” 2020.

Schrödinger LLC. “Maestro”, New York, NY, 2021

Terrillon, Sonia, and Michel Bouvier. “Roles of G-Protein-Coupled Receptor Dimerization.” *EMBO Reports*, vol. 5, no. 1, 2004, pp. 30–34., doi:10.1038/sj.embor.7400052.

Torres, Pedro H., et al. “Key Topics in Molecular Docking for Drug Design.” *International Journal of Molecular Sciences*, vol. 20, no. 18, 2019, p. 4574., doi:10.3390/ijms20184574.

Trott, Oleg, and Arthur J. Olson. “AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading.” *Journal of Computational Chemistry*, 2009, doi:10.1002/jcc.21334.

UniProt Consortium, The. “D(2) Dopamine Receptor.” *UniProt Consortium European Bioinformatics InstituteProtein Information ResourceSIB Swiss Institute of Bioinformatics*, 22 Apr. 2020, www.uniprot.org/uniprot/P14416.

UniProt Consortium, The. “Oxytocin Receptor.” *UniProt ConsortiumEuropean Bioinformatics InstituteProtein Information ResourceSIB Swiss Institute of Bioinformatics*, 22 Apr. 2020, www.uniprot.org/uniprot/P30559.

UniProt Consortium, The. “UniProt: the Universal Protein Knowledgebase.” *Nucleic Acids Research*, vol. 46, no. 5, 2018, pp. 2699–2699., doi:10.1093/nar/gky092.

Vakser, Ilya A. “Protein-Protein Docking: From Interaction to Interactome.” *Biophysical Journal*, vol. 107, no. 8, 2014, pp. 1785–1793., doi:10.1016/j.bpj.2014.08.033.

Van Meer, Gerrit, et al. “Membrane Lipids: Where They Are and How They Behave.” *Nature Reviews Molecular Cell Biology*, vol. 9, no. 2, 2008, pp. 112–124., doi:10.1038/nrm2330.

Veselovsky, A. V., et al. “Protein-Protein Interactions: Mechanisms and Modification by Drugs.” *Journal of Molecular Recognition*, vol. 15, no. 6, 2002, pp. 405–422., doi:10.1002/jmr.597.

Waltenspühl, Yann, et al. “Crystal Structure of the Human Oxytocin Receptor.” *Science Advances*, vol. 6, no. 29, 15 July 2020, doi:10.1126/sciadv.abb5419.

Wang, S., et al. “Structure of the D2 Dopamine Receptor Bound to the Atypical Antipsychotic Drug Risperidone.” *Nature*, no. 555, 2018, pp. 269–273., doi:10.2210/pdb6c38/pdb.

Waterhouse, Andrew, et al. “SWISS-MODEL: Homology Modelling of Protein Structures and Complexes.” *Nucleic Acids Research*, vol. 46, no. W1, 2018, doi:10.1093/nar/gky427.

Wolber, Gerhard, and Thierry Langer. “LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters.” *Journal of Chemical Information and Modeling*, vol. 45, no. 1, 2005, pp. 160–169., doi:10.1021/ci049885e.

Yin, J., et al. “Structure of a D2 Dopamine Receptor-G-Protein Complex in a Lipid Membrane.” *Nature*, vol. 584, 11 June 2020, pp. 125–129., doi:10.2210/pdb6vms/pdb.

York, Nathaniel, et al. “Oxytocin (OXT)-Stimulated Inhibition of Kir7.1 Activity Is through PIP 2 -Dependent Ca 2+ Response of the Oxytocin Receptor in the Retinal Pigment Epithelium in Vitro.” *Cellular Signalling*, vol. 37, 2017, pp. 93–102., doi:10.1016/j.cellsig.2017.06.005.

Zhang, Qian, et al. “Recent Advances in Protein-Protein Docking.” *Current Drug Targets*, vol. 17, no. 14, 2016, pp. 1586–1594., doi:10.2174/1389450117666160112112640.