University of Brighton

BY392 Biology Project

Dissertation

2022 - 2023

In silico study of Sertraline complexes with Danio rerio (drSERTaa) by free binding energy & RSMD of protein-ligand complex, a molecular dynamics analysis

By

Stefan Ritchie
19841251

Supervisor

Dr. Nicolas Stewart

Word Count: 10,125

LSP Statement

Steph has diagnoses of dyspraxia, dyslexia, autism spectrum condition, ADHD, anxiety, depression, and pernicious anaemia. These are fluctuating conditions that can impact on attendance.

Adjusted deadlines

Modification of deadlines for assignments and projects, of 14 calendar days. This adjustment to deadlines is to take into account the impact of your disability and can be used to help manage your workload if needed.

I confirm that I have a Learning Support Plan for which includes adjustment deadlines as recommended by the Disability & Dyslexia Team, and agreed by the school. I understand the deadline for my assessment has been adjusted (as per the required School protocol) and that this should be taken into consideration when my assessment is marked/ graded.

Alternative to presentations

Alternative arrangements for assignments requiring a presentation will be discussed with you. This may include alternatives such as presenting to one or two members of staff, video recording a presentation or preparing a presentation for someone else to give.

This is to ensure you meet the learning outcomes of the course while managing the impact of your disability.

Spelling and grammar

Due to the impact of your disability, you are more likely to make errors in spelling and grammar than other students. This will be taken into consideration as far as possible in marking and provided it does not conflict with an essential course requirement, the work should not be significantly downgraded. It is your responsibility to ensure the following wording is included on your assignment when it's submitted. I confirm that I have a Learning Support Plan for 'spelling and grammar' as recommended by the Disability & Dyslexia Team, and agreed by the school. I understand that this should be taken into consideration when my assessment is marked/ graded.

Summary

This research article investigates the bonding strength of the Selective Serotonin Reuptake Inhibitor (SSRI) Sertraline Hydrochloride and its associated target serotonin transporter in Zebrafish *Danio rerio* sodium-dependent serotonin transporter (drSERTaa) to determine if *Danio rerio* the bonding strength between the protein and ligand are higher than in humans. Using molecular dynamics, which is a simulation that uses physics and chemistry-based equations to simulate the evolution of particles within a system at discreet time steps, and artificial intelligence (AI) to predict the 3-dimensional shape of protein structures that haven't been modelled before. The relevance of such a study allows us to measure the bonding strength and extrapolate whether the required concentration of SSRIs to have an effect is lower than that of the human counterpart. This study is also a proof of concept to demonstrate the capability of AI generated protein structures and their practicality in the field of bioinformatics. To demonstrate this, 3D models of the Sertraline molecule were generated and placed within their optimum binding position within a model of drSERTaa derived from a homologue of a human equivalent, then superimposed with an AI generated model from AlphaFold to determine the optimum shape. The study found that the Zebrafish drSERTaa protein bound to the sertraline ligand produced a stronger bond strength than the human protein equivalent, but the amount of shift the sertraline molecule experienced when bound to the serotonin receptor in Zebrafish was greater, which indicated that the bound while stronger, was also more unstable. The impact of such a finding could be that it provides a framework for determining the concentration required to constitute a danger to Zebrafish populations, and further investigations of concentrations of SSRIs in wastewater should be evaluated to determine if there is a risk to aquatic organisms near wastewater outlets.

Abstract

Introduction

Molecular dynamics (MD) has been extensively applied within the pharmaceutical industry. Current processes, workflows and methodologies involved in the field of Molecular Dynamics were investigated in this review, to identify current best practices and efficient workflows to maintain the highest level of accuracy and reproducibility for future academic studies. MD has been successful in identifying numerous chemical compounds that show promise as potential drug candidates (Jorgensen, 2004). MD uses the underlying principles of Newtonian physics to simulate trajectories of molecules in set time steps. Forcefields, a term used to describe a computational method of estimating forces of atoms acting upon each other (Mayo, Olafson et al., 2002) Simulating binding free energy with numerous psychoactive drugs, identifying variations in formations of enzyme-substrate complexes, and linking them to their differences in their results. Such as Zebrafish and humans.

Method

Far less is known about the effects of these SSRIs on aquatic organisms. To address this challenge, an analysis of binding free energy and Root mean square deviation (RMSD) within a protein-ligand complex of Sertraline and aquatic organism *Danio* rerio Sodium-Dependant serotonin transporter (drSERTaa) protein generated using experimental protein structures developed by DeepMind's AlphaFold to form a homology model.

A 3D protein homology model of *Danio rerio* sodium-dependent serotonin transporter (drSERTaa) was generated using the genetic sequence of *Danio rerio* (UniProtKB Q1WGB5) and human sodium-dependent serotonin transporter (hSERT) genetic sequence (UniProtKB P31645) using SWISS-MODEL, a protein structure homology-modelling server.

Results

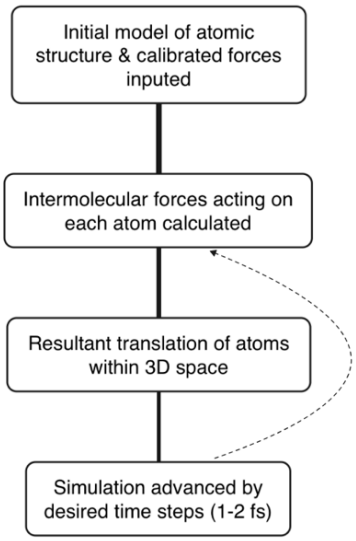
The homology model for drSERTaa achieved a sequence identity of 70.83%, exceeding Bacterial leucine transporter (LeuT) at 20-25% sequence identity commonly used as a simplified model for simulations involving serotonin transporters. A QMEANDisCo Global score of 0.74±0.05 was achieved, signifying a high-quality model. Using the best scoring conformational binding pose of Sertraline, the binding free energy within a drSERTaa-Sertraline protein-ligand complex achieved was of -11.5559 kcal/mol, achieving a lower score than that of Serotonin-hSERT protein-ligand complex at -8.9 kcal/mol. This indicates that Sertraline forms a stronger bond between drSERTaa-Sertraline complex than that of the human equivalent hSERT-Sertraline complex. The Root mean squared deviation (RMSD) of sertraline within a drSERTaa-Sertraline complex was higher, ranging from 1.5-2.4Å. Signifying a lower ligand binding stability in drSERTaa-Sertraline complex. Binding affinity in drSERTaa-Sertraline was -11.5559 kcal/mol, lower than the human equivalent, both these data points signify a potential lower active dose for Sertraline to have a therapeutic effect on *Danio rerio.*

Discussion

The study faced limitations in computational performance available which would have allowed for a greater simulation time frame. This could have potentially allowed for tertiary structure changes in the protein to be observed after the formation of the protein-ligand complex. Future areas of studies involving surveying current filtration techniques employed at water management plants and their effectiveness at removing SSRIs from wastewater. The same 3 functional binding regions on sertraline were observed in the formation of the drSERTaa-Sertraline complex as we see in hSERT-Sertraline complexes, this highlights a toxicological risk for *Danio rerio* as it suggests potential toxicity risks.

Introduction

Prescriptions for psychopharmaceutical drugs have increased within the UK over the past few years, with statistics show the NHS prescribed a total of 70.8 million antidepressants in 2018 alone (Iacobucci, 2019). An increase of 6.15% to 20.8 million antidepressant prescriptions in the third quarter of 2021 from the same quarter in the previous year (NHSBSA, 2021). SSRIs have been demonstrated to be an effective treatment for numerous psychological conditions. A meta-analysis containing 16,056 participants over 57 trials examined the efficacy of SSRIs for Generalised Anxiety Disorder (GAD) found that there was a consensus of an improvement during the acute treatment phase (Jakubovski, Johnson et al., 2019). The applicability of SSRIs is not just limited to psychological conditions. Premenstrual syndrome, chronic pain conditions such as fibromyalgia & chronic headaches along with irritable bowel syndrome (IBS) due to gut-brain axis interactions (Carabotti, Scirocco et al., 2015), are some of the conditions that SSRIs are approved for. The large numbers of prescriptions for SSRIs, along with the high variance of conditions they are used for has led to concerns over potential negative outcomes from environmental and health agencies. Contamination of aquatic environments has become a cause of concern and is now listed as a Contaminate of Emerging Concern (CEC) (Mole and Brooks, 2019). SSRIs have been measured at detectable concentrations in aquatic waterways, with one study performed in Fourmile Creek, USA showing concentrations of antidepressants higher within fish neural tissue than surrounding water ways. Detectable levels were found up to 8.4km from wastewater discharge points (Schultz, Furlong et al., 2010). Due to the nature of this study, it would be beneficial to minimise any harm or tests that could be carried out on live organisms. One of the ways to do this is through an *in-silico* study, which involves computational simulations to model biological interactions. One of the ways to do this is through molecular dynamics (MD). Molecular dynamics was first introduced by Alder & Mainwright's simulation of hard sphere interactions in a closed system (Alder and Wainwright, 1957). The discovery was very limited in its scope What is paved the way for future developments that more accurately represented real world observations. It is a method for analysing the physical movements of atoms and molecules within a system. Notable mentions are the first realistic system of liquid water (Stillinger and Rahman, 1974) which was ground-breaking as it demonstrated the ability to simulate real world physical fluidic environments in a scientifically accurate way. Some of the benefits are ability to save on raw materials and the ability to easily scale studies through parallel processing of numerous variants of structures in rapid succession. For example HOOMD-blue 1.0 was able to scale throughput 12.5x through the use of Graphical Processing Units (GPUs) (Glaser, Nguyen et al., 2015). This has developed into large-scale supercomputers able to simulate extremely high throughputs up to 150,000x in High-performance Compute (HPC) applications such as the Anton 2 supercomputer (Shaw, Grossman et al., 2014). Performing HPC computations for time periods of milliseconds (ms). In some cases, this machine can compute full biological systems including significant numbers of biochemical processes that would not be possible in a lab. MD provides us the ability to rapidly test millions of drug variations to identify viable drug structures and identify interaction points of biological pathogens complexes. SARS-CoV-2 is a recent example of this. Small molecule inhibitors demonstrated how resourceful MD can be, 300 peptide-like structures were screened to measure their effectiveness as a main protease inhibitor for SARS-CoV-2 (Pant, Singh et al., 2021) with 60 viable structures were identified during the study, a process that would have taken months of repetitive testing and be subject to human error. Another study by Selvaraj's identified the N7-MTase enzyme protein located at its c-terminal, and its role in RNA capping during the emigration of viral RNA from immune cells, which plays a vital role in preventing rapid RNA degradation (Selvaraj, Dinesh et al., 2021). This was significant as it identified N7-MTase's as potential targets for inhibition. MD has been one of the most ground-breaking advancements in modern medicine, allowing us to perform complex simulations covering molecular



interactions, protein folding, drug discovery and more. Molecular dynamics works by performing inter & intramolecular interactions at discreet timesteps and translating calculated processes into motions & energy levels of individual molecules. An adapted version of Kai Nordlund's simplified diagram explaining this process (Durrant and McCammon, 2011) *Fig.1*. In his model, you can see how initial vector coordinates are used to quantitatively measure molecular interactions, which in turn are used to produce vector translations within 3D space.

Figure 1: simplified schematic showing the steps involved in a molecular dynamic's simulation is performed.

MD allows us to identify the crucial residues involved in the binding free energy conformation between proteins, enabling the development of vaccines from these findings. This is often conducted in conjunction with molecular docking, a process that utilizes the starting coordinates of recognized protein-ligand complexes to assess ligand interactions and ensure the complex's activity is both desirable and within accepted boundaries. Implementing these procedures in a conventional manner would have required significant time and substantial financial resources. However, MD significantly reduces both the time and the financial costs involved. At present, research in this field is constrained due to the intricacy of the process. Consequently, this literature review will concentrate on our current comprehension and research regarding the utilization of MD, specifically for discerning specific residue-protein complexes and their variation between humans and certain marine organisms. Protein structures are sorted using the Research Collaboratory for Structural Bioinformatics PDB (RCSB PDB) databank (Berman, Westbrook et al., 2000), an online database supported by the Federal Drug Association (FDA). This database categorizes various forms and isotypes/isomers of distinct protein molecular structures, spanning from intricate DNA variants derived from various microorganisms to different chiral protein structures located within intermembrane domains. An area of potential study is developing the foundations for future research concerning psychopharmacological drugs and their impact on marine organisms. Studies have shown that SSRIs can have a range of effects on fish, including changes in behaviour, growth (de Farias, Oliveira et al., 2020), reproduction, and neuroendocrine function (Lister, Regan et al., 2009). Some studies have found that low levels of SSRIs in the environment (ng/L range) can affect fish behaviour and physiology, while other studies have found no effects at these levels. It is also worth noting that the effects of SSRIs on fish may not be the same as their effects on humans, as different species can have different responses to the same drug, due to varying metabolism processes, excretion pathways & immune responses.

Using MD, the effects of Selective Serotonin Reuptake Inhibitors (SSRIs) using molecular dynamics to simulate the transfer of the psychopharmacological drugs across membrane bilayers using the Human Serotonin reuptake inhibitor (hSERT) membrane channel protein homologue as a baseline model for investigation. The hSERT model was produced using a protein homologue derived from *Drosophila melanogaster* flies. *Drosophila melanogaster* dopamine transporter (dDAT) (Penmatsa, Wang et al., 2013). Two potential model templates were available in this test, a new model Drosophila melanogaster dopamine transporter (dDAT) and Bacterial Leucine Transporter (LeuT). dDAT was chosen due to its higher sequence identity (53%) than LeuT (23%) and thus was used for the first time in the creation of a hSERT homology model (Xue, Wang et al., 2016). The research underscored alterations in the binding free energy between the substrate AChR and different drugs. Five new hotspot residues (Ala169, Ala173, Thr439, Gly442, and Leu443) were identified as shared binding site influencers for the four tested SSRIs. The investigation determined that the disparity in binding free energy between various SSRIs was moderate, with the active binding sites persisting uniformly across the tested SSRIs.

Another study involving the use of the bacterial homologue LeuT conducted by Weiwei Xue studied the inhibitory effects of FDA approved selective serotonin reuptake inhibitors (SSRI) and how they interact with their conformational binding sites within the Bacterial Leucine Transporter (LeuT) membrane protein structures, by measuring free binding energy levels using MD. The main protein used in the study was RCSB PDB (Protein Data Bank) 4M48 as listed under the database. Results showed that pre-residue binding of 245 residues were successful. Binding mode interactions between SSRI and 11 hotspot residues in hSERT. This study provides a good baseline for possible residues available for future analysis. The findings of this study align with similar findings under Zhou's paper on antidepressant specificity of Serotonin transporters among three LeuT-SSRI protein-ligand complexes. The binding energy potentials of protein substrate complex formations between co-crystallised LeuT and various SSRI residues, along with enantiomers R-Fluoxetine & S-Fluoxetine. All SSRI's bound to LeuT at the same site. This study further verifies the validity of LeuT-SSRI complex studies on binding free energy experiments. Which may be used to formulate a study on free-energy binding of SSRI residuals to LeuT variants from differing organisms. The investigation of enantiomers for Fluoxetine is an important aspect to take note of, as S-Fluoxetine is more potent than R-Fluoxetine (Gram, 1994). This work is of significance as it lays the groundwork for MD studies relating to LeuT-SSRI integrations, proving the validity of results true to real-world observations. Which leads us on to our next study investigating Sertraline's effect on marine invertebrates during early life development. This study indicates potential areas for improvement in future studies. The use of LeuT can come into question for future MD simulations as new technologies allowing for more protein structures to be identified. Protein models representative of human Sodium-dependent serotonin transporters (SERT) have been generated under DeepMind's AlphaFold program. However, the study identified both hydrophobic and polar amino acid residues between SSRI's and LeuT which are identical to that of SSRI-hSERT complexes (Xue, Wang et al., 2016).

On reviewing this research, the application of membrane-bound channel proteins could present difficulties for 'individually focused' studies. Limited availability of HPC resources mean that only one SSRI can be used for a study, as time restrictions and processing capabilities pose limitations that need to be factored in. Consequently, the simulation of a membrane bilayer may pose as a challenge as MD simulations rely heavily on a few key major models. The protein, the biomolecular membrane the protein is situated in (if applicable), the water model, force field & minimisation model. These all heavily impact the outcome, and thus is important to select the correct one before carrying out a simulation. Limitations regarding membrane lipid bi-layers within a simulation can be mitigated with the use of an implicit membrane model which simulates the general properties of a phospholipid

bilayer thus increasing computational performance, but at the expense of simulation accuracy (Lindahl and Sansom, 2008). Limitations of water model accuracy to performance cost can me managed through the choice of an appropriate water model (Further detail in Table 4). In terms of protein models, various variants of AChR, LeuT and 5-hydroxytryptamine (5-HT) are generally used and should be considered.

A study by Canesi confirmed the major role Serotonin (5-HT) plays a wide variety of roles in bivalves. Some of the listed roles included gametogenesis, gill ciliary beating and heart functionality (Shi, Han et al., 2020). This was analogous for a wide range of bivalve species. Investigations in contaminates of emerging concerns (CECs) which include pharmaceutical drugs demonstrated measurable effects of CECs on serotonergic system. Which is particularly applicable as SSRIs are one of the highest detected pharmaceutical drugs in wastewater. Detectable levels were recorded to have increased concentrations in comparison to the surrounding environment (Canesi, Miglioli et al., 2022). One of the potential causes could be bioaccumulation. However, a recent 2017 study found that within a laboratory environment a 3-level aquatic food chains involving *Acer plantanoides, Asellus aquaticus, Notonecta glauca* & *Pungitius pungitius* showed no increase in SSRI concentrations of sertraline and fluoxetine due to bioaccumulation. Mean sertraline bio accumulation factors (BAF) were 2200 L/kg, 360 L/kg, 26 L/kg, and 49 L/kg respectively, and mean fluoxetine BAFs 300 L/kg, 110 L/kg, 11 L/kg, and 41 L/kg respectively (Bostrom, Ugge et al., 2017). The recorded BAFs did however fit within the recorded Bioconcentration factors (BCF) for the organisms, but the range was significantly high. Some of the reported reasons for this are differences in biotransformation and metabolism of the selected organisms in the study, that may have a greater ability to filter out SSRIS, and thus opposes the argument that bioaccumulation may be a significant cause of concern for SSRIs. The experiment was conducted within a controlled laboratory environment and doesn't fully represent the habitat the organisms are native to.

Rilei Yu's study into MD simulations of competitive agonist dihydro-B-erythroidine (DHβE) replacing nicotine. Nicotinic acetylcholine receptors (nAChR) are an abundant receptor in the human brain is associated with numerous CNS disorders. Binding to human nicotinic heteromeric acetylcholine receptor (a4β2 nAChR) was analysed and the mechanism of conformational transition from a desensitised to a closed resting state was studied to build a model for the membrane bound protein during competitive inhibition. The study found a stable structure for nAChR at both its closed resting states when suspended in a membrane Bilayer and identified L264 as an influencing factor for the slow component of desensitisation. The study found numerous physical changes of the DHβE bound receptor during its transition from active to desensitised state. Poor lining α-helix was observed to be tilting an average of 2.9° from parallel to its central axis. In the crystal structure of a4β2 nAChR was observed to be 8.4° in its desensitised state. a4β2 nAChR from different origins did not affect the tilt angle. The main constraints of the transmembrane pore were α4 L257 (9') & α4 L264 (16'). The component of desensitisation was also identified to be L257. The significance this is that a potential mechanism of inhibition was identified. AChR receptors have shown to be a potential candidate for this MD study and poses as an acceptable model for other CNS related experimentation (Yu, Tae et al., 2019).

While studies involving the use of LeuT for *in-silico* studies of SSRI interactions does have benefits in simplicity, reduced computational performance cost and high levels of supporting clinical evidence., issues remain regarding the applicability of such a model for studies involving clinical effects of drugs as recent studies have highlighted the high genetic sequence difference between LeuT and human serotonin receptor and other equivalent mammalian counterparts as stated above.

Comprehensive physiological review of SSRI Sertraline on premature marine invertebrates (Estevez-Calvar, Canesi et al., 2017). The study found levels of Fluoxetine, sertraline, and their metabolites norfluoxetine and desmethylsertraline respectively within brain matter tissues at significant level. Fluoxetine levels values of mean ± standard deviation: fluoxetine, 1.58 ± 0.74 ng/g; norfluoxetine, 8.86 ± 5.9 ng/g; sertraline, 4.27 ±1.4 ng/g; desmethylsertraline, 15.6 ± 14.3 ng/g). N. Kreke's study on this indicates some flaws in the analytical process currently employed to monitor environmental changes in SSRI concentrations. Tests of toxicity occurred at both 24 hours and 48 hours significant effects of these tests were only registered at 100 mg/L with mortality rates of 24 h: p = 0.047, U = 2; 48 h: p = 0.014, U = 0). Results at 1000 mg/L. Measures of behavioural changes such as swimming speed were affected by sertraline concentrations. Swimming speed was found to increase in small bursts after 48 hours of exposure at 1 µg/L (p = 0.014, U = 0). Is that sertraline in certain marine organisms. This is of significance as it highlights the importance of us managing the release of SSRIs into waste systems as humans may be significantly lower for marine organisms.

A study on how Bivalves were affected by Fluoxetine sourced from waste pollution provided some key insight on fluoxetine concentration differences in aquatic environments from various countries. All regions had detectable levels of Fluoxetine within their surface waters. Canada 46 $ng\ L^{-1}$, Croatia 66 $ng\ L^{-1}$, Spain 18-66 $ng\ L^{-1}$& 100 $ng\ L^{-1}$, USA 2.6, 12 & 111 $ng\ L^{-1}$. Thus, promoting the notion that SSRI concentrations are present and detectable in wild aquatic environments and highlighting the need for further studies on potential impacts this may be having in wild populations.

Biomarker response evaluation was performed on mussel populations regarding exposure to Fluoxetine concentrations of 75 $ng\ L^{-1}$, which noticeably caused to caused tissue damage to the gills. (Brooks, Turner et al., 2003).

Zebrafish have been shown to express various forms of 5-hydroxytryptamine receptor(5-htr) (Norton, Folchert et al., 2008). In regards for the specific receptor 5-htr1a found in humans, the genetic equivalent serotonin receptor hSERT's (slc6a4a & slc6a4b) show 66-69% and 75% amino acid sequence homology. Htr1aa & htr1ab show 69% and 76% identical to human htr1a equivalents, respectively, which is much higher than the LeuT and dDAT bacterial homologues. An analysis of the prevalence of antidepressants of surface water was analysed from around the world where Zebrafish populations have been located highlight Zebrafish as a potential candidate for an *in-silico* study. The top three highest concentrations were India, Canada & Brazil with concentrations of 40500ng L^{-1}, 410 ng L^{-1} & 202 ng L^{-1} respectively. Other notable mentions are United States of America (USA) at 142 ng L^{-1} and the United Kingdom at 40ng L^{-1}d (Gould, Winter et al., 2021).

Table 1: Distribution of serotonergic receptor (5-HTT) subtypes in *Danio rerio* (Zebrafish)

Organ	Subtype
CNS	1aa, 1ab, 1bd, 2c
Gonad/Ovaries	2c
Muscle	2c
Skin	2c
Unknown	2a, 7

CNS, Central Nervous System

Ryan Wong's study on gene expression in Zebrafish (n=30) when exposed to 100µg/L S-Fluoxetine their behavioural analysis of Zebrafish identified an increased presence in the top half the tank ($t = -6.49$, $p = 2 \times 10^{-6}$) compared to control. Which suggested that fluoxetine reduced stress and anxiety related behaviours. No significant difference was identified between R and S-Fluoxetine ($t = 1.47$, $p = 0.16$) (Wong, Oxendine et al., 2013)

Table 2: Comparison of Molecular Dynamics Simulation Software Packages

Name	License	Integration
NAMD (Phillips, Braun et al., 2005)	Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)	GPU Compute NVIDIA CUDA Acceleration Parallel processing
GROMACS (Van Der Spoel, Lindahl et al., 2005)	GNU Lesser General Public License (LGPL), version 2.1	GPU Compute Full CUDA Acceleration
CHARMM (Brooks, Brooks et al., 2009b)	Proprietary	GPU Compute
CHARMM-GUI (Wu, Cheng et al., 2014)	Proprietary	Web-based platform
AMBER20 (Case, Aktulga et al., 2021)	Proprietary	Streamlined setup process CUDA SDK GPU Accelerated free energy binding
Ambertools22 (Case, Aktulga et al., 2022)	GNU General Public License (GPL).	GPU Acceleration
VMD (Humphrey, Dalke et al., 1996)	Proprietary; UIUC Open-Source is still looking to study today license	GUI based workflow

Nanoscale Molecular Dynamics (NAMD); Groningen Machine for Chemical Simulations (GROMACS); Chemistry at Harvard Molecular Mechanics (CHARMM); Assisted Model Building with Energy Refinement (AMBER); Visual Molecular Dynamics (VMD); Graphical Processing Unit (GPU); Compute Unified Device Architecture (CUDA); Software Development Kit (SDK).

Table 3: MD Force Field (FF) Comparison

Name	Advantages	Disadvantages
CHARMM27 (Brooks, Brooks et al., 2009a)	Improved reliability Greater compatibility with existing programs Most widely used. Greater integration with 3rd party software packages	Greater performance overhead. Inaccuracies in reproducing hydrogen bonds and ionic interactions Depreciated water model
CHARMM36 (Huang and MacKerell, 2013)	Better management of disordered proteins Improved lipid parameters More up to date water model Greater integration with hardware acceleration	Reduced library of remodelled residues Higher computational cost than CHARMM27
GROMOS96 (Scott, Hünenberger et al., 1999)	Optimised for glycoprotein confirmational shapes	Outdated model Emulation for ARM based
OPLS-AA (Abdel-Azeim, 2020)	Optimised for high concentration electrolyte solutions	Fail to reproduce properties of long alkanes accurately (Siu, Pluhackova et al., 2012)
AMBER94 (Cornell, Cieplak et al., 2002)	Ideal water model angle of 109.47	Outdated forcefield model

Key: CHARMM, Chemistry at Harvard Macromolecular Mechanics; AMBER, Assisted Model Building and Energy Refinement

Water model (WM) comparison.

Liquid water is the most important solvent in nature providing the environment for biological reactions and interactions to occur. Therefore, it is important that we model water within a biological environment as accurately as possible. Water models play a critical role in accurately simulating biochemical systems and choosing the water model appropriate for the task is significant as it can have a drastic effect on the results of a simulation. Water models

Table 4: Comparative analysis of available MD Water Models

Name	Pros	Cons
SPC (Mark and Nilsson, 2001)	Greater accuracy than TIP3P	Reduced temperature scaling 3-point model Assumes inaccurate bond angle in water molecules
SPC/E (Mark and Nilsson, 2001)	Greater density and diffusion constant than SPC Greater accuracy at describing self-polarisation effect (Schmidt, Roberts et al., 2007) Highest generated binding energies (Nguyen, Viet et al., 2014) better density and diffusion constant than SPC	Less stability in Mean square Displacement over time Greater time required for equilibrium of a system to be reached 3-point model Only suitable for studying the properties of bulk water, and it is not able to accurately reproduce the properties of water at interfaces or in confined spaces.
TIP3P (Jorgensen, Chandrasekhar et al., 1983)	High computational throughput Most compatible water model with other MD programs Most widely used water model Extensively tested Simplified implementation process	Lower accuracy than a 4-point water model QM effects are not accounted for Inaccuracies in modelling
TIP4P (Jorgensen and Madura, 2006)	Greater generalised accuracy Higher accuracy than TIP3P at high pressure and low temperature extremes Reduced probability of false geometries in both water & gas phases (Kiss and Baranyai, 2011) 4-point water model Commonly used water model for simulations requiring high accuracy water models Greatest accuracy in hybrid *ab initio* QM/MM systems (Shaw, Woods et al., 2009)	Reduced experimental data available O-H bond angles are rigid, preventing bond flexibility Implementation process is more complicated than TIP3P More computationally intensive than 3-point water models
TIP5P (Mahoney and Jorgensen, 2000)	Highest accuracy for a large range of properties Average density error from -37.5℃ to 62.5℃ at 1 atm of 0.0006 g cm^{-3}	Greatest computational cost Low accuracy in hybrid *ab initio* QM/MM systems (Shaw, Woods et al., 2009)

Key: SPC, Single Point Charge; SPC/E, Single Point Charge Extended; TIP3P, Transferable intermolecular potential 3-point; TIP4P, Transferable intermolecular potential 4-point; TIP5P, Transferable intermolecular potential 5-point; QM, Quantum Mechanics; QM/MM, Quantum Mechanics/ Molecular Mechanics; g cm^{-3}, Grams per centimetre cubed.

Homology modelling

Homology modelling is a technique used in computational biology to predict the three-dimensional structure of a protein based on its amino acid sequence. It assumes that proteins with similar amino acid sequences will have similar structures. To create a homology model, the amino acid sequence of the protein of interest is compared to a database of known protein structures. If a similar protein is found, the structure of the known protein can be used as a template to generate a model of the protein of interest. Homology modelling is often used when the structure of a protein is unknown and cannot be determined experimentally, such as through X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. It can also be used to refine the structure of a protein that has been determined experimentally, by using the experimental structure as a template to refine the model. Homology modelling is a valuable tool for understanding the function of proteins and how they interact with other molecules. It can also be used to design new drugs or to predict the effects of mutations on protein structure and function. Some examples of homology models commonly used in MD are LeuT and LeuBAT.

Some examples of software are SWISS-MODEL (Waterhouse, Bertoni et al., 2018). A web-based platform for protein structure prediction. It is a tool that uses homology modelling to predict the three-dimensional structure of a protein based on its amino acid sequence. To use SWISS-MODEL, users can enter the amino acid sequence of the protein they are interested in, and the platform will search a database of known protein structures to find a template similar to the protein of interest. The template protein's structure is then used as a model to generate predictions of the structure of the protein of interest. SWISS-MODEL is a widely used resource for protein structure prediction and is particularly useful for predicting the structure of proteins for which there is no experimental data available. By using a widely used platform, it ensures standardization is maintained, increased likelihood of producing accurate results due to thorough testing, and reproducibility of generated outcomes. The software is licensed under the CC BY-SA 4.0 Creative Commons Attribution-ShareAlike 4.0 International License, meaning you can freely copy and redistribute models in any medium or format, including building and transforming upon prior models for any purpose. Using a web-based platform reduces performance limitations of local hardware. Running processes on non HPC platforms can be very time consuming and increases chances of incorrect structure generation when systems are not configured properly.

Binding profile visualisation tools

Molecular visualization software PoseView provided as part of the proteins.plus package, which allows for the 3D representation and manipulation of protein structures. One example of this is PoseView, a molecular visualization software designed specifically for analysing protein-ligand interactions. Some negatives are the limited complexity of the program. It does not provide the ability to create complex analysis tools or any form of editing ability of the generated visualisation. Poseview is also a web-based tool which requires access to the internet, however this allows for cross-compatibility with of the platforms.

LeView tool designed to analyse ligand efficiency of drugs with their target proteins. It allows more complex calculations between various ligand efficiency-based metrics such as the proprietary Ligand efficiency (LE), ligand lipophilicity efficiency (LLE) and ligand dependent of the lipophilicity (LELP). It provides a user-friendly Graphical user interface (GUI), however due to the use case for protein-ligand complex interaction visualisation software in this study, these features would not be required as similar ligand cluster ranking is provided by platforms such as SWISS-MODELS Qualitative Model Energy Analysis for Discrepancies and Optimization (QMEANDisCo) & Dockthor's virtual screening of top-energy binding modes (Santos, Guedes et al., 2020).

Future developments

AlphaFold is an artificial intelligence system developed by DeepMind to generate 3D molecular structures of proteins and other biological molecules using predictive algorithms. It has achieved over 200 million database entries in partnership with EMBL's European Bioinformatics institute (EMBL-EBI). With human proteomes and 47 other key organisms. Some notable examples are the structural predictions of SARS-CoV-2 spike proteins (Jumper, Tunyasuvunakool et al., 2020). AlphaFold can be a useful tool for further studies. Some of the protein structures generated by this study have more accurate human equivalents than traditional methods, such as UniProt P31645. This is an ai generated approximation of Sodium-dependent serotonin transporter of humans (hSERT). Which inherently provides greater relevance than bacterial homologues like LeuT, a common bacterial homologue model derived from *Aquifex aeolicus* used in antidepressant studies. In the latest edition of the biennial Critical Assessment of protein Structure Prediction (CASP) competition, AlphaFold achieved the highest overall accuracy of any participating system, outperforming all other methods in predicting the structure of proteins with medium and high resolution. This demonstrates the impressive accuracy of AlphaFold and its potential to aid in a wide range of research applications. However, it is worth noting that protein structure predictions are a complex and challenging task, and there is still room for improvement in the accuracy of AlphaFold and other methods.

Advancements in the field of MD have been made down to the quantum scale, this branch of MD is called Quantum molecular dynamics (QMD). To improve the accuracy of molecular systems, simulations are treated as quantum mechanical systems, meaning their behaviour is described using the principles of quantum mechanics. Some of the benefits of this method are increased small scale simulation accuracy, and greater accuracy in high energy systems (Pan, Van et al., 2022). The predictive power of QMD is greater than that of MD, allowing for the simulation for much finer and more intricate interactions between molecules occurring withing the picosecond (ps) and femtosecond (fs) timeframes. QMD also has greater accuracy in high count systems due to its ability to simulate wave function behaviours of large volumes of particles. These benefits do come with a cost. QMD simulations are on average are 100x more computationally expensive than traditional MD.

Great progression has been made in simulating SSRI protein-ligand interactions with various transporters. However, many of these structures use bacterial derivates of SSRI target proteins, such as LeuT and LeuBAT. There are benefits to this such as reduced performance overhead, greater simplicity in simulating interactions. This comes at a cost of accuracy and validity of simulation results due to low sequence identity and variations in system properties that don't represent the human homologue. There is potential for further investigational studies on aquatic organism constituents of neurotransmitter related membrane channel proteins and how they interact with psychopharmaceutic medications such as SSRI's. The new SERT models generated by DeepMind, a program by AlphaFold, allow for significant numbers of structures to be simulated as a large catalogue of over 220 million protein structures is now available (Jumper, Evans et al., 2021).

CHARMM36 has been chosen as the forcefield model of choice. The performance cost is not significant enough to warrant not using it. It is a more up to date model than CHARMM27 and incorporates parameters derived from quantum mechanical calculations to better simulate real world interactions, as evident in improvements in protein structure backbone accuracy (Huang and MacKerell, 2013). The water model TIP3P will be used during solvation due to the in-depth documentation and resources available to support this water model. This water model is also the current industry standard, and to maintain the comparability of this study with others and high accuracy. MD preparation will be done using CHARMM-GUI saving on both time and reducing chances of human error. Using CHARMM-GUI also improves repeatability due to its accessibility and ease of use.

Due to the current nature of the study and current limitations in computational performance, the decision has been made to study the binding free energy potential of SSRI sertraline with variants of sodium-dependent serotonin transporter (SERT) like protein structures among *Danio rerio*. This will be significant in understanding the risks involved with wastewater contaminated with SSRIs. By learning how Sertraline binds to serotonin transporters in zebrafish, we can develop a better understanding on the effective dose and allowing us to make more informed decisions on determining levels/concentrations that pose a risk to zebrafish populations around the globe. Little is known about the effects of these SSRIs on aquatic organisms, especially at a biomechanical level. To address this challenge, an analysis of binding free energy (ΔG) and Root mean square deviation (RMSD) within a protein-ligand complex of Sertraline and Sodium-Dependant serotonin transporter (SERT) proteins derived from fish will be performed using experimental protein structures developed by DeepMind as a validation method to confirm the 3D topology of generated homologues.

Molecular dynamics analysis has provided insight into the effects of selective serotonin reuptake inhibitors (SSRIs) on zebrafish. SSRIs are commonly used in the treatment of mental health disorders, such as depression and anxiety, but their use can also lead to side effects in humans. By analysing the molecular changes in the zebrafish that result from the SSRI, researchers can better understand the potential effects of SSRIs on humans, and potentially develop new, more effective treatments. In this essay, I will discuss the molecular dynamics analysis of SSRI effects on zebrafish, and how it helps inform our understanding of the use of SSRIs in medicine. As of recent research, it has been found that SSRI's influence zebrafish gene expression and behaviour (Chen et al., 2018). According to an experiment conducted by Chen et al., it was found that the expression of the serotonin transporter gene, which is involved in the regulation of serotonin levels, were upregulated in the test fish. It was also determined that it created a change in their behaviour, specifically higher locomotive activity and higher levels of anxiety were displayed as they attempted to flee from unfamiliar. The study suggests that modulating serotonin levels can have a statistically significant impact on expression of specific genes and the social behaviour of *Danio rerio*.

SSRIs have been studied for their effects on gene expression and behaviour in aquatic species, such as zebrafish. A study conducted in 2018 identified that the administration of SSRIs to zebrafish had a "significantly different effect on their social behaviour and gene expression that may serve as biomarkers of future onset of depression or anxiety", along with a "change in gene expression" in response to SSRI treatment. Two genetic pathways were identified to show significant upregulation. This change in gene expression could partly explain the behavioural effects previously observed in adult zebrafish following exposure to SSRIs. Ultimately, this research enhances our understanding of how SSRIs effect zebrafish behaviour, as well as providing potential biomarkers that can be further researched as tools in the diagnosis and treatment of depression and other mental illnesses.

Molecular dynamics analysis has been used to investigate the effects of Selective Serotonin Reuptake Inhibitors (SSRI) on zebrafish. The study found that there were numerous significant changes in the behaviour and biochemistry of the fish when exposed to the SSRIs. The study also found key differences in the behaviour and biochemistry of the fish which correlate with aggressive behaviour. This research highlights the importance of understanding the potential effects of pharmaceutical drugs on aquatic animals and emphasizes the need for further research and study into their effects.

Prescriptions for antidepressants have increased significantly within the last 30 years, providing medical aid for those who struggle with various psychological conditions. Consequently, we have seen an increase in Selective serotonin reuptake inhibitors (SSRI) released into wastewater. Cumulative concentration increases have been observed occurring within aquatic water systems, potentially posing a risk to aquatic wildlife.

Over the past few years, an increase in the use of neurological medications has increased drastically for the treatment of various neurological conditions such as generalised anxiety disorder, depression, bipolar disorder etc. This increase in the use of psychopharmacological drugs has resulted in detectable levels of SSRIs in wastewater. Concern over the effect of increasing concentrations of these drugs on the livelihood of aquatic organisms has become a concern for the scientific community and thus, investigatory actions have been taken to investigate and form solutions to this issue. This study aims to investigate the effect of SSRIs on aquatic organisms through an in-silico study on the associated ion gated channel sodium-dependent serotonin transporter (SERT). SERT is a commonly found neurotransmitter in vertebrate central nervous systems (CNS) and in mammals. SERT functions as a monoamine transporter protein, facilitating the transport of Serotonin from the synaptic cleft to the presynaptic neuron (Mortensen, Kristensen et al., 1999).

While crystallographic studies demonstrate the importance role plays in ligand binding steps, the cost and extensive labour required to generate data has led many to seek computational developments that allow parallelisation and automation in discovering protein ligand binding interactions. The utilization of Newtonian physics is pivotal in extrapolating the approximate results of molecular interactions to facilitate the streamlining and acceleration of the process.

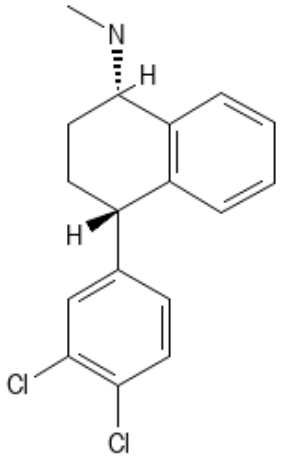


Figure 2. The chemical structure of the Sertraline compound: CHEMBL809.

The hypothesis is that higher binding potentials of SSRIs within aquatic organism protein homologues will be observed and may suggest lower concentrations of pharmacological drugs would be required to have a therapeutic effect. To investigate this, the study will perform a computational analysis of Root mean square deviation (RMSD) which looks at the overall deviation in starting position of the chosen molecules and binding free energy potential ΔG° (kcal/mol) of antidepressant Sertraline with the associated target membrane protein drSERTaa, to providing the foundations for quantitative analysis on the difference in effect SSRI's have on aquatic organisms than their human counterparts.

Methodology

The description of time evolution of particles in MD simulation relies on Newton's laws of motion. The fundamental equations are Newtons Second law which declares the acceleration of an object is determined by the net force acting on the object and its mass. Potential energy is the energy of a system based on its location and configuration. Hamilton's equations of motion equation use to describe the evolution of a systems position and momentum over time. The Verlet Algorithm is a second order mathematical method used to integrate Hamilton's equations of motion.

Preparation

The genetic sequence of *Danio rerio* was retrieved from UniProtKB (Q1WGB5) as a FASTA sequence, which is a text-based format representing amino-acid or nucleotide sequences as individually lettered codes (Pearson, 1994). Protein processing was performed using CHARMM-GUI (Lee, Cheng et al., 2016). CHARMM-GUI will reduce preparation time by a significant margin. Time reduction is estimated to be around 120 hours.

Structure manipulation

Missing residues for the protein are modelled using CHARMM-GUI's PDB readers structure manipulator process. DNA structural abnormalities were fixed using the GalaxyFill algorithm (Coutsias, Seok et al., 2004). Missing residue topology and bond interaction information was generated using Ligand Reader & modeler (Kim, Lee et al., 2017). The topology generation tool created two extra-long pair hydrogens hydrogen bonded to atoms Cl1 & Cl2, named LPH. This is incompatible with NAMD2, so this was swapped out for placeholder hydrogen molecules with the topology and parameter information of the lone pair hydrogens.

Homology model

Generated using SWISS-MODEL. An online cloud-based platform for the generation of new 3D protein structures using a homologous reference protein that already has a generated 3D structure. A 3D mol2 model representation of Sertraline was sourced from ZINC database (ZINC 1853550) (Irwin, Tang et al., 2020) providing topology and coordinate information for residue modelling. The mol2 file was then loaded into ChimeraX (Pettersen, Goddard et al., 2021) and converted to pdbqt format in preparation for merging ligand and protein files.

Energy minimisation

Energy minimisation is performed using CHARMM-GUI, time steps are 2 ns in length with 1000 minimisation processes. CGenFF3.0.1 topology and parameter files were used to perform basic energy minimisation of ligand molecules prior to NAMD minimisation step. The VMD Extension NAMD-GUI was used to create a NAMD run script involving a 20,000-step minimisation phase & a 10,000 step molecular dynamics phase for the protein and ligand individually. The protein in the minimised structure should be isolated, and the VMD plugin psfgen is used to generate protein structure files (psf) for the target protein. A custom tcl script which reads pdb co-ordinate files & psf structure files of both the protein and ligand to produce combined protein-ligand psf and pdb files should be used. Minimisation was performed in NAMD for 20,000 timesteps due to limitations with the allowed maximum speed of molecules within the system. The presence of lone pair hydrogens did cause some difficulty but reducing time steps from the initial 2 femtoseconds down to 0.5 femtoseconds solved this issue as molecules moved an effective 4 times less per bond/force calculations with molecular docking involving. To produce initial poses required for the starting state of the simulation, the ligands were docked with the modelled hSERT using Glide. Docking poses for sertraline were identified using SwissDock (Grosdidier, Zoete et al., 2011) running the program EADock (Grosdidier, Zoete et al., 2007). In order to optimise processing time, models homologically similar to hSERT are used to maintain

accuracy within an acceptable timeframe. An implicit POPC lipid bilayer was used to anchor the protein to its membrane pose orientation, exposing appropriate segments as intracellular and extracellular and minimising computational cost. Using a TIP3P water model, the protein-ligand complex was hydrated with a margin of 5 Å Providing a balance of performance and thermodynamic predictive accuracy. A topology file based on CHARMM version toppar_c36_jul22 (Huang and MacKerell, 2013). specifically edited for use with NAMD was used during the processing and preparation stages.

Topology & parameterization

Topology information is provided by a topology file which converts names of residuals from a database into complete PSF structure files. Ligand parameterization was performed using the CHARMM General Force Field (CGenFF) web server hosted by MacKerell lab in collaboration with SilcsBio, LLC (Vanommeslaeghe, Hatcher et al., 2010). CHARMM forcefield topology Top_all36 (Best, Zhu et al., 2012) was used to convert residue names to associated PSF structures. Lipid (Klauda, Venable et al., 2010). Initial Energy minimisation of protein structure will be performed using CHARMM-GUI before running the MD simulation in NAMD to reduce chances of catastrophic disassembly of the protein-ligand structure.

MD Simulation

MD simulation was performed on Visual Molecular Dynamics (VMD) v. 1.9.4a53 (64-bit Intel x86_64). Using the QwikMD (Ribeiro, Bernardi et al., 2016) module running NAMD v. 2.14 (Win64-CUDA). This version contains NIVDIA Compute Unified Data Algorithm (CUDA) Acceleration for improved ns/day performance (Phillips, Hardy et al., 2020). Using CHARMM Force fields V4.6 par_all36_prot & par_all36_lipid for proteins and lipids, respectively. Ion parameters for TIP3P water model represented by CHARMM forcefields toppar_water_ions.str in a CUBIC crystalline structure. NVT ensemble, which represents the use of N= fixed numbers of atoms, V= fixed volume & T= Fixed temperature was used for input generation equilibration. NPT (N= fixed number of atoms, P= fixed pressure, T= fixed temperature) ensemble was used for Input generation at a temperature of 297 K. Molecular bonds containing hydrogen atoms were constrained with the SHAKE algorithm (Andersen, 1983) constraining two O-H bonds and the Angle between H-O-H bonds with timesteps at 1 fs. Periodic boundary conditions (PBCs) were used to simulate an infinite system, allowing for the maintenance of pressure, temperature and volume while still allowing for fluctuations and shifts in these values to occur. Direct space interaction calculations with long range electrostatic boundary condition cut-off at 5Å allowed for biochemical interactions to be accurately simulated while preventing premature crashing due to Lone Pair Hydrogen (LPH) molecules LP1 & LP2, electrostatically coupled with the chlorine atoms on Sertraline hydrochloride.

VMD plugin QwikMD allows for fast preparation of protein structures for simulation using NAMD backend MD processing. This will reduce preparation time and technical debugging time significantly. Both an implicit and Explicit water model can be used for free energy binding calculations. Implicit water models save on computational time by simulating the general properties of water without simulating individual water molecules.

Protein-ligand system

A recently derived protein structure formed using cryogenic electron microscopy (cryo-EM) crystalline structure at a resolution of 3.30Å (PDB code 6vrh) (Coleman, Navratna et al., 2020) was used as the hSERT homology model for the generation of the target protein drSERTaa. The homology model torsion bond torsion angles were validated using a Ramachandran plot. Named after G. N. Ramachandran in 1963. It functions by representing the distribution of phi (ϕ) and psi (ψ) bond angles for each amino acid residue in a protein in the form of a scatter plot (Zhou, O'Hern et al., 2011). Helping

us identify energetically favourable and unfavourable conformations of amino acids around conformational spaces.

QMEANDisCo Global measures distance distributions between homologous protein structures using neural networks to combine the accuracy of consensus methods with the broad appositeness of single model approaches. QMEANDisCo is considered one of the top performing methods for model quality estimation as reported in CAMEO data sets scoring 0.94 (Studer, Rempfer et al., 2020) .

Binding profile

Using PoseView (Stierand, Maass et al., 2006) provided on Proteins.Plus. A simplified 2D representation. PoseView is open source, making it accessible to researchers and allowing for community contributions to its development.

Energy calculations

$G_{bind} \approx -k_B T \ln[\Delta\omega/8\pi^2] - k_B T \ln[C^\circ \Delta V] + \Delta G_{LP}$

Binding free energy ($\Delta G_{bind} = \Delta G_{water} - \Delta G_{protein}$) refers to the difference in energy between the bound and unbound states of the ligand (Cournia, Allen et al., 2017). Binding free energy ($DG_{MM/GBSA}$) of Sertraline-drSERTaa complex was calculated using a single trajectory-based method. The Binding Free Energy Estimation (BFEE) plugin on VMD was used to perform these calculations. Another important thermodynamic function is Gibbs free energy ($\Delta G = \Delta H - T\Delta S$), which defines the spontaneity of a system. A negative value indicates a system that is more spontaneous, and a positive value represents a less spontaneous system (Wilhelm, 2021).

Ramachandran plot

Using SWISS-MODEL's structure assessment tool, a Ramachandran plot assesses the quality of the homology model and identify favourable conformations. It provides insight into the stability of a proteins secondary structure.

Results

Homology model

Homologue drSERTaa (UniProtKB Q1WGB5) generated through SWISS-MODEL using a hSERT template (PDB 6vrh). Model was cross verified with the predictive model from AlphaFold (AF-Q1WGB5-F1) genetic sequence ID, with only evidence at transcription level.

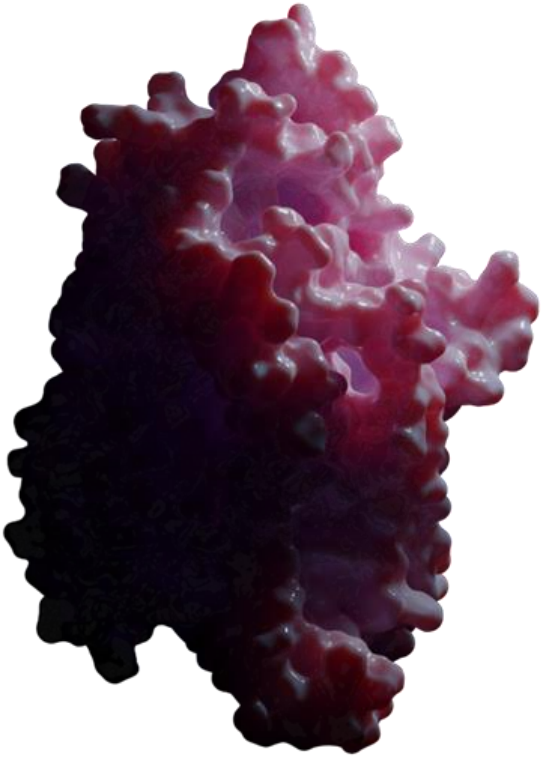


Figure 3. 3D Representation of transmembrane protein drSERTaa. *Visualisation showing protein surface produced in Blender 3D cycles render engine OptiX compute shaders.*

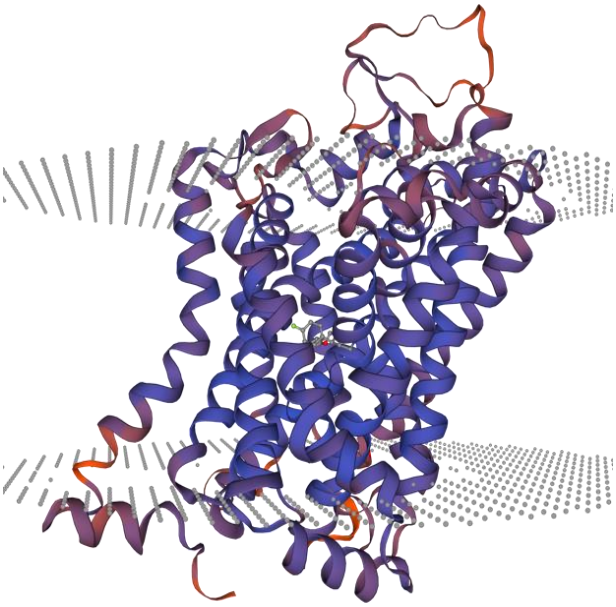


Figure 4. 3D representation of drSERTaa-Sertraline complex with its appropriate transmembrane orientation, after conformational binding. *Membrane represented as grey planes. Protein colour coded based on alignment homologue (drSERTaa) conformation with template (6vrh.1) with blue being high confidence and red meaning low confidence. Ligand Sertraline is represented in grey within the centre of the homologue, with green representing negatively charged Chlorine atoms and red representing positively charged nitrogen atom.*

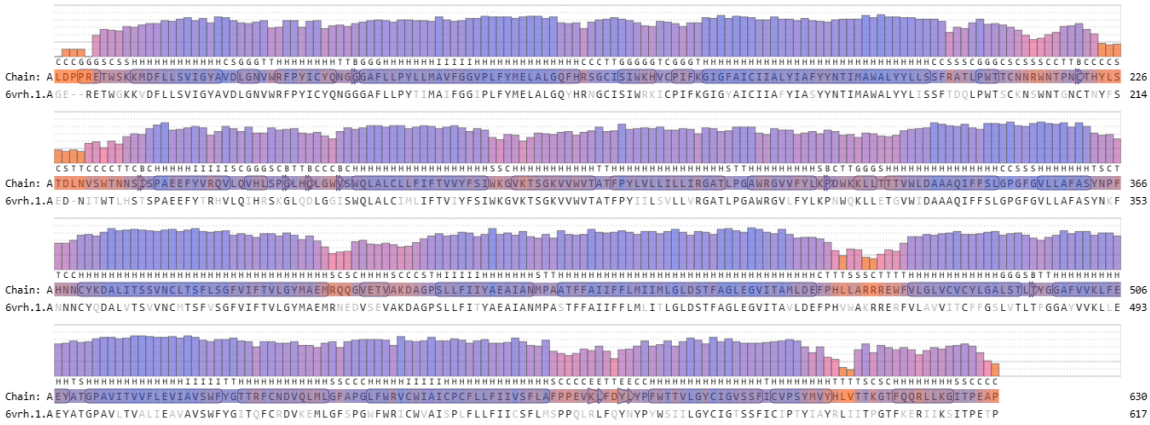


Figure 5. Alignment chart for drSERTaa homologue in comparison to 6vrh.1 template. *Rows represent residues in the sequence of proteins. Residues are aligned between proteins with single letters representing single-letter amino acids in the FASTA sequence. Top row (Chain A) represents homologue & bottom row (6vrh.1) represents template. Residues conserved between structures are boldened due to their significant roles in protein functionality. Amino acid indels represented as arrows due to sequence shift.*

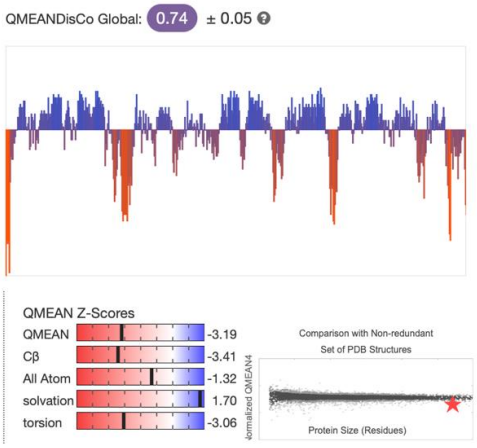


Figure 6. QMEANDisCo Global and Z-score representing model quality in comparison to human homologue.

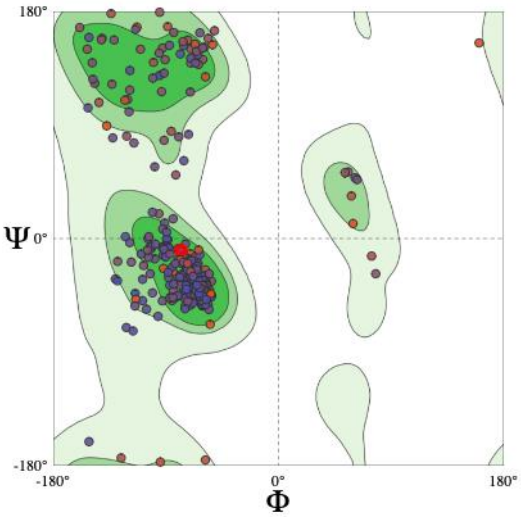


Figure 7. Ramachandran plot representing spread of energetically favourable Phi(ϕ) & Psi(ψ) dihedral angle distribution in drSERTaa homologue.

The homology model 'Q1WGB5_DANRE Q1WGB5 Transporter' was generated in SWISS-MODEL on 9/2/23 using 6vrh. 1 (PDB code 6VRH) (Coleman, Navratna et al., 2020). A structure titled 'Sodium-dependent serotonin transporter Cryo-EM structure of the wild-type human serotonin transporter complexed with paroxetine and 8B6 Fab' as a template due to it receiving the highest Global Model Quality Estimation (GMQE) score of 0.76 and highest Identity score of 70.83. And UniProtKB (The UniProt, 2017) Q1WGB5 as my target. Sequence identity was 70.83% in respect to the human homologue UniProtKB P28223, which demonstrates a high level of variation in genetic sequence. QMEANDisCo Global (Studer, Rempfer et al., 2020) scored 0.74 ± 0.05.

Protein-ligand binding clusters

34 ligand binding cluster groups were generated, with a total of 256 potential binding modes identified between them. Average binding free energy across 256 binding profiles was -7.3531 kcal/mol. The average binding free energy of the best scoring cluster was -11.5559 kcal/mol.

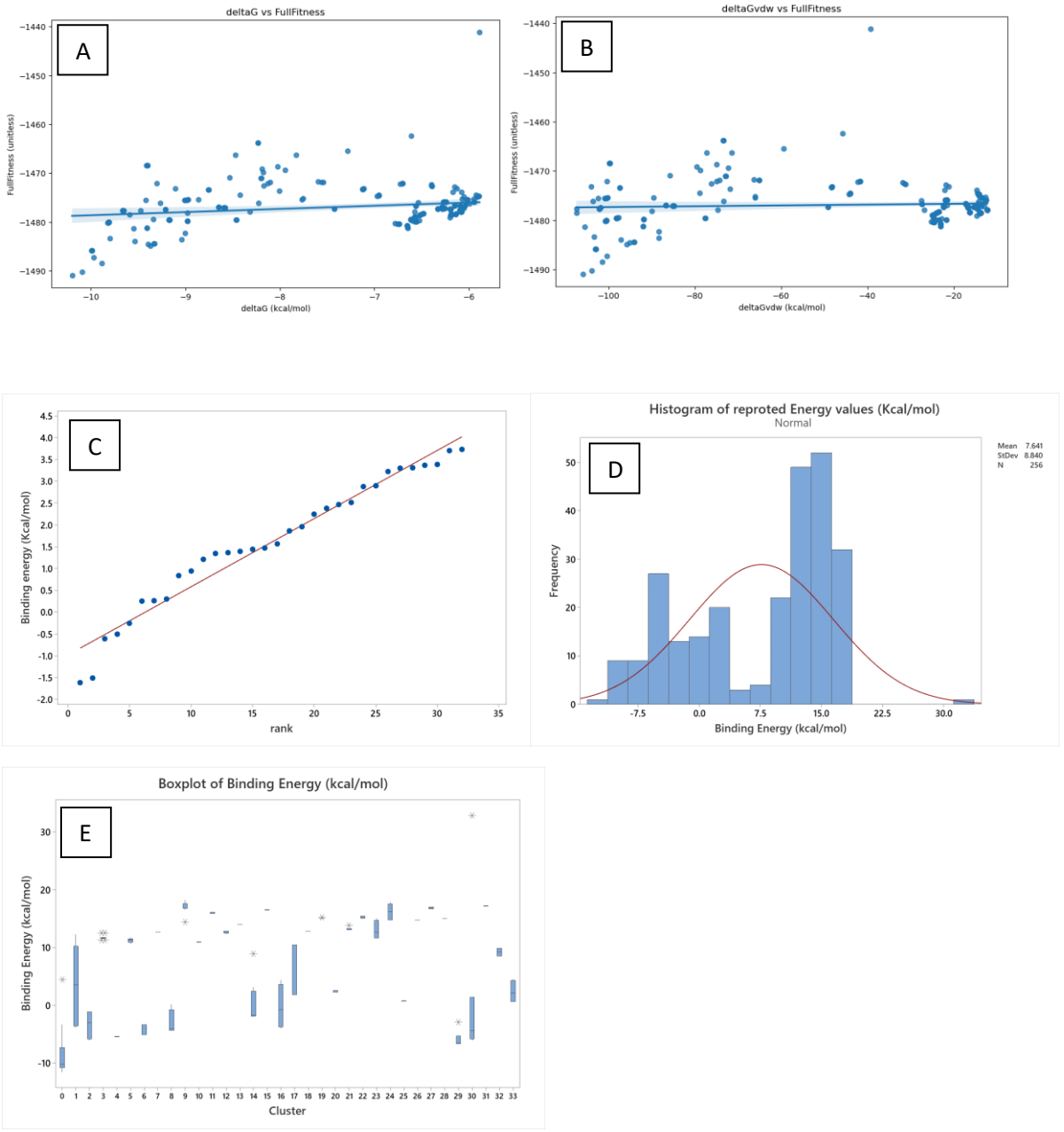


Figure 8. [A] Scatter plot shows the relationship between the binding energy (ΔG) and the full fitness score. The trendline indicates the general direction of the relationship between these two variables. Data shown is in kcal/mol for ΔG and unitless for FullFitness. **[B] Scatter plot showing the relationship between the van der Waals energy (ΔGvdw) and the full fitness score.** The trendline indicates the general direction of the relationship between these two variables. Data shown is in kcal/mol for ΔGvdw and unitless for FullFitness. **[C] Scatterplot representing change in Binding energy (kcal/mol) as cluster rank increases (Lower is better).** Binding energy has a strong correlation with cluster rank indicating the quality of the binding cluster is strongly correlated to the bending energy of the ligand (Sertraline) to the protein homologue (drSERTaa). **[D] Histogram representing the distribution of Full Fitness (Kcal/mol) of ligand binding poses. [E] Boxplot representing the distribution of Binding energies (kcal/mol) within the cluster ranks.** No correlation is found between the clusters and binding energies (kcal/mol)

Protein-ligand binding profile

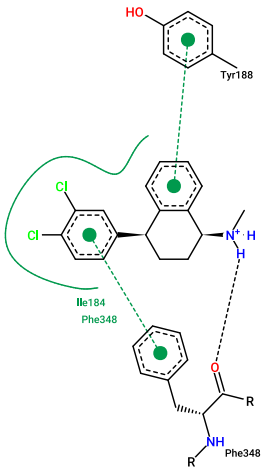


***Figure 9. Binding modes for drSERTaa-Sertraline complex in PoseView**. Green lines represent hydrogen bonding between chlorine groups expressed on Sertraline and amino acids Ile184 & Phe348 of the protein. Green dots connected via dashed lines represent π-π stacking. Salt bridge hydrogen bonds between Phe348 amide group and Sertraline represented by grey dashed line.*

Visualisation of the binding interactions between the drSERTaa-Sertraline protein-ligand complex is represented in figure 6 and generated using Poseview in proteins.plus (Stierand, Maass et al., 2006). The binding mode is represented by mainly hydrophilic, hydrophobic & electrostatic interactions among 3 key chemical groups. R1-3, respectively. Hotspot residues of drSERTaa were Tyr188 in R1, nitrate oxygen group interaction in R2 then Ile184 & Phe348 in R3. Two π-π stacking interactions characterised as attractive noncovalent orbital interactions between the pi bonds of aromatic rings that are parallel to one another and was observed in R2 and R3. R3 also had hydrophobic interactions with the major interactions between Ile184 & Phe348. Salt bridge hydrogen bonds were observed in R1 with Phe348 amide group. In comparison to the human homologue, the same three functional binding regions on Sertraline were identified (Xue, Wang et al., 2016). Significantly more hydrophilic interactions were identified between groups R1-3 for hSERT. In R1 Asp98, R2 Ala169, Ile172, Phe341, Tyr176, R3 Tyr95, Ile172, Ala173, Ser438, Thr439, Gly442, Leu443. Two π-π interactions in groups R2 and R3 were observed in hSERT, identical to the *Danio rerio* homologue drSERTaa.

NAMD minimisation

Minimisation was performed on drSERTaa with 5000 timesteps at a step level of 0.5 femtoseconds. Temperature of the model was set to 300 k for 0.104 nanoseconds then 297K for 10 nanoseconds. Docking was performed on SwissDock with the following parameters.

Root Mean squared deviation (RMSD) Calculations

RMSD was performed using VMD's RMSD visualiser tool. The RMSD was calculated on the Sertraline molecule (resname LIG) using the backbone modifier to negate hydrogen atoms from the calculation. The calculations showed Sertraline had a higher RMSD when aligned to the protein drSERTaa than the human homologue hSERT.

The identification of the inhibitory mechanisms from Sertraline was performed by integrating multiple computational methods. A reported template hSERT SMTL ID: 6vrh.1 from 2020 was successfully adopted to generate the homology model of drSERTaa, and the binding mode shared by hSERT & drSERTaa was identified by hierarchically clustering per-residue binding free energies of 245 residues.

The identified binding mode was defined by 11 hot spot residue interactions (Tyr95, Asp98, Ala169, Ile172, Ala173, Tyr176, Phe341, Ser438, Thr439, Gly442 and Leu443) in hSERT. Although these residue interactions differed to the drSERTaa homologue, three major regional groups where identified, along with 2 of the same π-π interaction in regions R2 and R3. 5 hot spot residues (Ala169, Ala173, Thr439, Gly442 and Leu443) found in this study have not yet been identified as common determinants of all 4 studied SSRIs in binding hSERT. In the last section of this study, changes in SSRIs' binding induced by mutation on hot spot residues were further explored, and 3 mechanisms underlining their drug sensitivity were summarized. The binding mode identified in this study provided significant insights into the inhibitory mechanism of approved SSRIs, which could be utilized as a useful framework for assessing and discovering novel lead scaffolds. The higher RMSD between Sertraline-drSERTaa complex indicates that Sertraline may have a harder time reaching the active site of serotonin receptors in *Danio rerio*. The average score of -11.5559 kcal/mol.

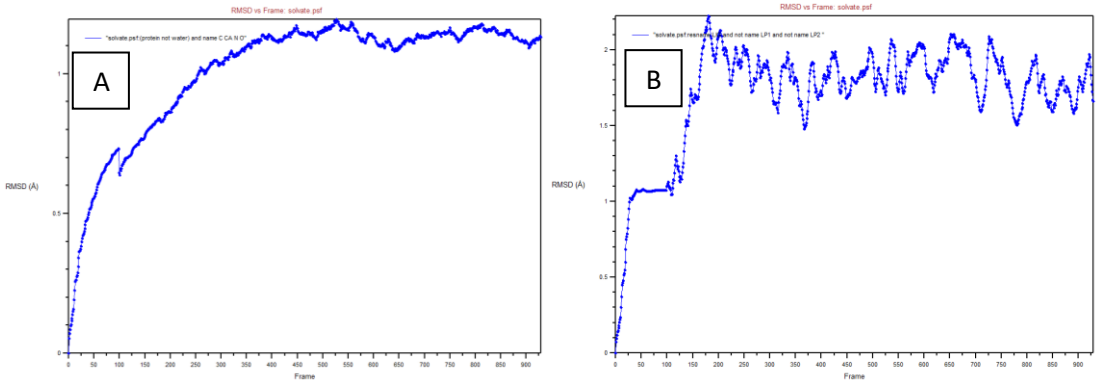


Figure 10. [A] Root Mean Square Deviation (RMSD) of drSERTaa-Sertraline complex over 930 nanoseconds (ns)*. After 100 ns minimisation stage, a large drop can be seen due to minimisation stage being too short. Slight fluctuations in protein structure are observed, signifying that the drSERTaa homologue is stable.* ***[B] RMSD of Sertraline in complex with drSERTaa.*** *High fluctuations ranging from 1.5 to 2.4 observed for ligand Sertraline.*

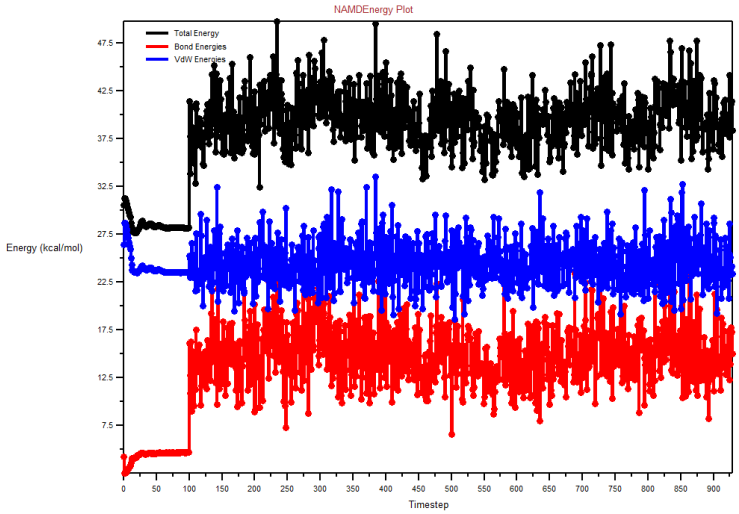


Figure 11. Binding energy (kcal/mol) of Sertraline in complex with drSERTaa over 930ns. *First 100 nanoseconds are low in energy due to 100 nanosecond minimisation stage. And a stable spread of energy level fluctuations of the ligand from 100-930ns, ranging from 7.5-22.5 kcal/mol.*

Discussion

The identification of the inhibitory mechanisms from Sertraline was performed by integrating multiple computational methods. A reported template hSERT SMTL ID: 6vrh.1 from 2020 was successfully adopted to generate the homology model of drSERTaa, and the binding mode shared by hSERT & drSERTaa was identified by hierarchically clustering per-residue binding free energies of 245 residues. The identified binding mode was defined by 11 hot spot residue interactions (Tyr95, Asp98, Ala169, Ile172, Ala173, Tyr176, Phe341, Ser438, Thr439, Gly442 and Leu443) in hSERT. Although these residue interactions differed to the drSERTaa homologue (, three major regional groups where identified, along with 2 of the same π-π interaction in regions R2 and R3.

The higher RMSD ranging between 1.5-2.5Å Sertraline-drSERTaa complex indicates that Sertraline bond with drSERTaa active site of serotonin receptors in *Danio rerio* is less stable than hSERT equivalent. The mean binding free energy score of the best ranking ligand binding pose cluster -11.5559 kcal/mol which is lower than recorded in hSERT-Sertraline complexes derived from humans (–8.9 kcal/mol). This signifies that sertraline more readily binds to the *danio rerio* sodium-dependent serotonin receptors.

It's important to note that there are limitations in terms of computational performance meaning the study could only be run for 930 nanoseconds. Using HPC would allow for a significant improvement in the number of time steps the simulation could be run for. Which could highlight different conformational bonding positions due to changes in tertiary structure of the target protein which tend to take more time to express. With all *in silico* studies it is good practice to have physical tests to validate findings. However, a question of ethics should be raised, as it is in best practice not to harm any organisms. This highlights some of the benefits of doing an *in-silico* study as you're able to calculate biological processes without endangering any organisms.

Further areas of study include the effects of other SSRIs on drSERTaa, along with the effects of SSRIs on other aquatic organisms. be done to investigate concentrations of SSRIs or other drugs released from wastewater discharge points, along with investigations as to whether current filtration techniques in the UK are effective against filtering out SSRIs.

Conflicts of interest

It is important to note that there are no conflicts of interest that could potentially affect the research or the reported results of this study. There are no financial or personal relationships that could influence the research or the results, and there are no funding sources or affiliations that could create potential conflicts of interest. Therefore, readers can trust that the research and the results reported are unbiased and objective.

Bibliography

Abdel-Azeim, S. 2020. Revisiting OPLS-AA Force Field for the Simulation of Anionic Surfactants in Concentrated Electrolyte Solutions. *J Chem Theory Comput* 16(2) 1136-1145.

Alder, B.J. and Wainwright, T.E. 1957. Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* 27(5) 1208-1209.

Andersen, H.C. 1983. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics* 52(1) 24-34.

Berman, H.M. et al. 2000. The Protein Data Bank. *Nucleic Acids Res* 28(1) 235-242.

Best, R.B. et al. 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* 8(9) 3257-3273.

Bostrom, M.L., Ugge, G., Jonsson, J.A. and Berglund, O. 2017. Bioaccumulation and trophodynamics of the antidepressants sertraline and fluoxetine in laboratory-constructed, 3-level aquatic food chains. *Environ Toxicol Chem* 36(4) 1029-1037.

Brooks, B.R. et al. 2009a. CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10) 1545-1614.

Brooks, B.R. et al. 2009b. CHARMM: The biomolecular simulation program. *Journal of computational chemistry* 30(10) 1545-1614.

Brooks, B.W. et al. 2003. Waterborne and sediment toxicity of fluoxetine to select organisms. *Chemosphere* 52(1) 135-142.

Canesi, L., Miglioli, A., Balbi, T. and Fabbri, E. 2022. Physiological Roles of Serotonin in Bivalves: Possible Interference by Environmental Chemicals Resulting in Neuroendocrine Disruption. *Front Endocrinol (Lausanne)* 13 792589.

Carabotti, M., Scirocco, A., Maselli, M.A. and Severi, C. 2015. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol* 28(2) 203-209.

Case, D.A. et al. 2022. *Amber 2022*: University of California, San Francisco.

Case, D.A. et al. 2021. *Amber 2021*: University of California, San Francisco.

Coleman, J.A. et al. 2020. Chemical and structural investigation of the paroxetine-human serotonin transporter complex. *Elife* 9.

Cornell, W.D. et al. 2002. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* 117(19) 5179-5197.

Cournia, Z., Allen, B. and Sherman, W. 2017. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J Chem Inf Model* 57(12) 2911-2937.

Coutsias, E.A., Seok, C., Jacobson, M.P. and Dill, K.A. 2004. A kinematic view of loop closure. *J Comput Chem* 25(4) 510-528.

de Farias, N.O. et al. 2020. Fluoxetine chronic exposure affects growth, behavior and tissue structure of zebrafish. *Comp Biochem Physiol C Toxicol Pharmacol* 237 108836.

Durrant, J.D. and McCammon, J.A. 2011. Molecular dynamics simulations and drug discovery. *BMC Biol* 9(1) 71.

Estevez-Calvar, N. et al. 2017. Adverse effects of the SSRI antidepressant sertraline on early life stages of marine invertebrates. *Mar Environ Res* 128 88-97.

Glaser, J. et al. 2015. Strong scaling of general-purpose molecular dynamics simulations on GPUs. *Computer Physics Communications* 192 97-107.

Gould, S.L., Winter, M.J., Norton, W.H.J. and Tyler, C.R. 2021. The potential for adverse effects in fish exposed to antidepressants in the aquatic environment. *Environ Sci Technol* 55(24) 16299-16312.

Gram, L. 1994. Fluoxetine. *N Engl J Med* 331(20) 1354-1361.

Grosdidier, A., Zoete, V. and Michielin, O. 2007. EADock: docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins* 67(4) 1010-1025.

Grosdidier, A., Zoete, V. and Michielin, O. 2011. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res* 39(Web Server issue) W270-277.

Huang, J. and MacKerell, A.D., Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* 34(25) 2135-2145.
Humphrey, W., Dalke, A. and Schulten, K. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14(1) 33-38.
Iacobucci, G. 2019. NHS prescribed record number of antidepressants last year. *BMJ* 364 l1508.
Irwin, J.J. et al. 2020. ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *J Chem Inf Model* 60(12) 6065-6073.
Jakubovski, E. et al. 2019. Systematic review and meta-analysis: Dose-response curve of SSRIs and SNRIs in anxiety disorders. *Depress Anxiety* 36(3) 198-212.
Jorgensen, W.L. 2004. The many roles of computation in drug discovery. *Science* 303(5665) 1813-1818.
Jorgensen, W.L. et al. 1983. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* 79(2) 926-935.
Jorgensen, W.L. and Madura, J.D. 2006. Temperature and size dependence for Monte Carlo simulations of TIP4P water. *Molecular Physics* 56(6) 1381-1392.
Jumper, J. et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873) 583-589.
Jumper, J., Tunyasuvunakool, K., Kohli, P. and Hassabis, D. 2020. *Computational Predictions of Protein Structures Associated with COVID-19* [*online*].
Kim, S. et al. 2017. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *J Comput Chem* 38(21) 1879-1886.
Kiss, P.T. and Baranyai, A. 2011. Sources of the deficiencies in the popular SPC/E and TIP3P models of water. *J Chem Phys* 134(5) 054106.
Klauda, J.B. et al. 2010. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114(23) 7830-7843.
Lee, J. et al. 2016. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput* 12(1) 405-413.
Lindahl, E. and Sansom, M.S. 2008. Membrane proteins: molecular dynamics simulations. *Curr Opin Struct Biol* 18(4) 425-431.
Lister, A., Regan, C., Van Zwol, J. and Van Der Kraak, G. 2009. Inhibition of egg production in zebrafish by fluoxetine and municipal effluents: a mechanistic evaluation. *Aquat Toxicol* 95(4) 320-329.
Mahoney, M.W. and Jorgensen, W.L. 2000. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics* 112(20) 8910-8922.
Mark, P. and Nilsson, L. 2001. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A* 105(43) 9954-9960.
Mayo, S.L., Olafson, B.D. and Goddard, W.A. 2002. DREIDING: a generic force field for molecular simulations. *The Journal of Physical Chemistry* 94(26) 8897-8909.
Mole, R.A. and Brooks, B.W. 2019. Global scanning of selective serotonin reuptake inhibitors: occurrence, wastewater treatment and hazards in aquatic systems. *Environ Pollut* 250 1019-1031.
Mortensen, O.V., Kristensen, A.S., Rudnick, G. and Wiborg, O. 1999. Molecular cloning, expression and characterization of a bovine serotonin transporter. *Brain Res Mol Brain Res* 71(1) 120-126.
Nguyen, T.T., Viet, M.H. and Li, M.S. 2014. Effects of water models on binding affinity: evidence from all-atom simulation of binding of tamiflu to A/H5N1 neuraminidase. *ScientificWorldJournal* 2014 536084.
NHSBSA. 2021. *Medicines Used in Mental Health – England – Quarterly Summary Statistics July to September 2021* [*online*]: NHSBSA.

Norton, W.H., Folchert, A. and Bally-Cuif, L. 2008. Comparative analysis of serotonin receptor (HTR1A/HTR1B families) and transporter (slc6a4a/b) gene expression in the zebrafish brain. *J Comp Neurol* 511(4) 521-542.
Pan, X. et al. 2022. Accelerating Ab Initio Quantum Mechanical and Molecular Mechanical (QM/MM) Molecular Dynamics Simulations with Multiple Time Step Integration and a Recalibrated Semiempirical QM/MM Hamiltonian. *J Phys Chem B* 126(23) 4226-4235.
Pant, S. et al. 2021. Peptide-like and small-molecule inhibitors against Covid-19. *J Biomol Struct Dyn* 39(8) 2904-2913.
Pearson, W.R. 1994. Using the FASTA Program to Search Protein and DNA Sequence Databases. *In* A. M. Griffin and H. G. Griffin eds. *Computer Analysis of Sequence Data: Part I*. Totowa, NJ, Humana Press. pp. 307-331.
Penmatsa, A., Wang, K.H. and Gouaux, E. 2013. X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* 503(7474) 85-90.
Pettersen, E.F. et al. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* 30(1) 70-82.
Phillips, J.C. et al. 2005. Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16) 1781-1802.
Phillips, J.C. et al. 2020. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics* 153(4) 044130-044130.
Ribeiro, J.V. et al. 2016. QwikMD - Integrative Molecular Dynamics Toolkit for Novices and Experts. *Sci Rep* 6(1) 26536.
Santos, K.B., Guedes, I.A., Karl, A.L.M. and Dardenne, L.E. 2020. Highly Flexible Ligand Docking: Benchmarking of the DockThor Program on the LEADS-PEP Protein-Peptide Data Set. *J Chem Inf Model* 60(2) 667-683.
Schmidt, J.R. et al. 2007. Are water simulation models consistent with steady-state and ultrafast vibrational spectroscopy experiments? *Chemical Physics* 341(1-3) 143-157.
Schultz, M.M. et al. 2010. Antidepressant pharmaceuticals in two U.S. effluent-impacted streams: occurrence and fate in water and sediment, and selective uptake in fish neural tissue. *Environ Sci Technol* 44(6) 1918-1925.
Scott, W.R.P. et al. 1999. The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A* 103(19) 3596-3607.
Selvaraj, C. et al. 2021. Structure-based virtual screening and molecular dynamics simulation of SARS-CoV-2 Guanine-N7 methyltransferase (nsp14) for identifying antiviral inhibitors against COVID-19. *J Biomol Struct Dyn* 39(13) 4582-4593.
Shaw, D.E. et al. 2014. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis.
Shaw, K.E., Woods, C.J. and Mulholland, A.J. 2009. Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. *The Journal of Physical Chemistry Letters* 1(1) 219-223.
Shi, W. et al. 2020. Immunotoxicities of microplastics and sertraline, alone and in combination, to a bivalve species: size-dependent interaction and potential toxication mechanism. *J Hazard Mater* 396 122603.
Siu, S.W., Pluhackova, K. and Bockmann, R.A. 2012. Optimization of the OPLS-AA Force Field for Long Hydrocarbons. *J Chem Theory Comput* 8(4) 1459-1470.
Stierand, K., Maass, P.C. and Rarey, M. 2006. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. *Bioinformatics* 22(14) 1710-1716.
Stillinger, F.H. and Rahman, A. 1974. Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics* 60(4) 1545-1557.
Studer, G. et al. 2020. QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics* 36(6) 1765-1771.

The UniProt, C. 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45(D1) D158-D169.
Van Der Spoel, D. et al. 2005. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* 26(16) 1701-1718.
Vanommeslaeghe, K. et al. 2010. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* 31(4) 671-690.
Waterhouse, A. et al. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46(W1) W296-W303.
Wilhelm, E. 2021. CHAPTER 1. Gibbs Energy and Helmholtz Energy: Introduction, Concepts and Selected Applications. *In* E. Wilhelm and T. M. Letcher eds. *Gibbs Energy and Helmholtz Energy*. The Royal Society of Chemistry. pp. 1-120.
Wong, R.Y., Oxendine, S.E. and Godwin, J. 2013. Behavioral and neurogenomic transcriptome changes in wild-derived zebrafish with fluoxetine treatment. *BMC Genomics* 14(1) 348.
Wu, E.L. et al. 2014. CHARMM-GUIMembrane Buildertoward realistic biological membrane simulations. *Journal of Computational Chemistry* 35(27) 1997-2004.
Xue, W. et al. 2016. Identification of the inhibitory mechanism of FDA approved selective serotonin reuptake inhibitors: an insight from molecular dynamics simulation study. *Phys Chem Chem Phys* 18(4) 3260-3271.
Yu, R. et al. 2019. Molecular dynamics simulations of dihydro-beta-erythroidine bound to the human alpha4beta2 nicotinic acetylcholine receptor. *Br J Pharmacol* 176(15) 2750-2763.
Zhou, A.Q., O'Hern, C.S. and Regan, L. 2011. Revisiting the Ramachandran plot from a new angle. *Protein Sci* 20(7) 1166-1171.

Appendix I – Poster presentation

drSERTaa free-binding energy potential with Sertraline

University of Brighton

By: Steph Ritchie Supervisor: Corina Ciocan

School of Applied Sciences, University of Brighton

Scan to view the protein-ligand complex in 3D

SCAN ME

Introduction

Molecular Dynamics simulations are a form of in silico experiment used to study the physical behaviour of molecules at the atomic level. The motion of atoms and molecular structures are simulated in discrete time steps and incorporate fundamental mathematical principles to determine the physical & chemical behaviour of a system.

The ligand Sertraline is a very common Selective serotonin Reuptake inhibitor (SSRI) FDA approved for the treatment of Major depressive disorder, Obsessive-compulsive disorder (OCD) and PTSD and many various other psychological conditions [1].

The protein drSERTaa is a Serotonin Transporter found in Zebra fish (*Danio rerio*) is a membrane protein responsible for the reuptake of Serotonin from the synaptic cleft back into the presynaptic neuron terminating synaptic impulses. Serotonin is responsible for many psychological processes including mood, sleep, digestion and perception of sensory stimuli. Disruptions to Serotonin transporters can lead to dysregulation leading to a multitude of various issues such as feeding behaviour, aggression and locomotion in aquatic organisms [2].

Through the use of AI generated protein structures of *Danio rerio* (Zebrafish) serotonin receptors (drSERTaa) derived from Deepmind's AlphaFold[3] (UniProtKB Q1WGB5), an *In silico* analysis of root mean square deviation (RMSD) and binding affinity (Kcal/mol) was performed between a custom made sertraline-drSERTaa protein-ligand complex to determine if *danio rerio* are more susceptible to SSRIs in waste water and the effective dose is lesser than that of Human serotonin concentration (hSERT).

Aims

This research aims to asses the binding free energy and RMSD of Sertraline and Zebrafish drSERTaa protein-ligand complexes, and the results from such measurements show significance to suggest waste water needs to be treated for SSRIs.

Hypothesis

Binding free energy and RMSD within a protein-ligand complex of Sertraline and Zebrafish drSERTaa is greater than that of *Homo sapien* equivalent hSERT, highlighting the potential need to process waste water for SSRIs due to their lower effective dose.

Discussion

Limitations of *in silico* testing:

- Limited accuracy due to the use of mathematical models which may incorporate simplified methods in order to speed up calculations. Fundamental issues related to floating point calculations are also important to consider in regards to divisions [6].
- Experimental validation may be required in order to confirm findings in *in silico* studies.
- High performance compute (HPC) would be useful to resolve higher datapoints over an extended period of time.

Future studies:

- Further testing of genetic variants of drSERTaa in *Danio rerio* could be useful to determine if certain populations may be more susceptible to the effects of SSRIs in wastewater.

Methods

- Homology model Q1WGB5_DANRE Q1WGB5 was developed using SWISS-MODEL, scoring 70.83% sequence identity and a QMEANDisCo[4] global score of 0.74±0.05.
- Residue modelling was performed using the sdf file for ZINC1853550 from ZINC Database for Sertraline.
- Missing residue parameters were generated using the LigParGen web server.
- Input generation for NAMD was performed using CHARMM-GUIs PDB Reader to fix any errors with simulation input files.
- Energy minimisation was performed in VMD using NAMD at 1000 minimisation steps.
- A 2D Binding profile was produced in Poseview.
- Solvation was performed using the Solvate plugin with a 5 Å boundary and 0.15M NaCl solute concentration.
- Molecular dynamics simulation was performed using NAMD.
- Root mean squared deviation (RMSD) graphs generated using NAMD Energy plot plugin in VMD.

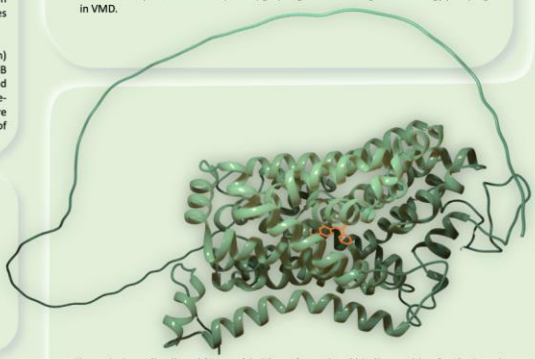


Figure 4. Sertraline ligand (Orange) in it's conformational binding position for *danio rerio* serotonin receptor (drSERTaa). Ligand is shown in orange and Protein is shown in green.

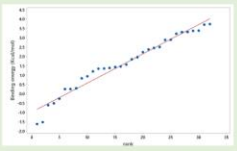


Figure 5. Distribution of sertraline-drSERTaa binding scores as rank increases. Ranked positions of binding poses from 1-32 with 1 being best pose changes and their associated Binding energy (Kcal/mol).

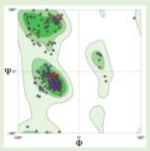


Figure 6. Low deviation in torsional angles of amino acids are within reference ranges as shown in the Ramachandran plot between the homology model for drSERTaa from the reference hSERT model. All values shown are Phi (Φ) and psi(ψ).

Results

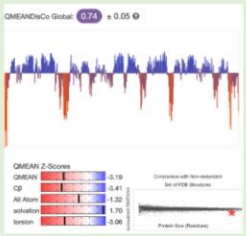


Figure 1. Composite absolute score of generated drSERTaa model derived from Q1WGB5_DANRE in comparison to hSERT homology model. QMEANDisCo represents the generated models structural quality. Generated model has a high protein size, however the QMEANDisCo Global score is within reference range, which indicates the protein structure has high dimensional accuracy.

Table 1: Sertraline ligand binding characteristics

Variable	Mean	SE Mean	StDev	Variance	Skewness	Kurtosis
Binding energy (Kcal/mol)	-1.598	0.264	1.493	2.229	-0.48	-0.51

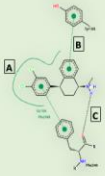


Figure 2. 2D representation of binding mode in Sertraline-drSERTaa complex shows highlights three fundamental binding regions. (A) Green lines represent hydrogen bonding between chlorine groups expressed on Sertraline and amino acids Ile184 & Phe348 of the protein. Green dots connected via dashed lines represent π-π stacking (B). Salt bridge hydrogen bonds between Phe348 amide group and Sertraline (C).

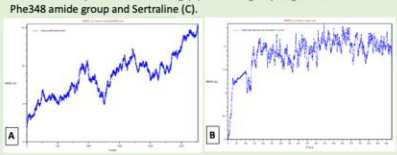


Figure 3. RMSD of Sertraline in complex with drSERTaa. (A) RMSD of drSERTaa protein over 2750 microseconds, showing the deviation from origin over time as protein shifts towards lower energy state. (B) RMSD of Sertraline in complex with drSERTaa isolated backbone over 950 microseconds. Data presented is Root mean squared deviation (RMSD) in ångström units (Å)

Further Research

- Further studies on the long term effect of SSRIs on aquatic organisms need to be studied
- With the advent of AlphaFold, predictive protein structures for aquatic organisms have become widely available. Validating studies/tests should be conducted to verify these findings
- Future SSRIs, what they are
- Studies using SERT homologues of other aquatic organisms
- drSERTaa studies involving other psychopharmacological medicines

Conclusions

- Binding profile for Zebrafish serotonin receptors and SSRIs are identical as shown in figure 2, indicating that similar clinical effects should be observed in Zebrafish when exposed to sertraline.
- Zebrafish serotonin receptors have a high sequence identity (70.83%) to human equivalents (hSERT, SLC6A4). Greater than that of LeuT at 25% [5], a common bacterial homologue derived from *Aquifex aeolicus* used in *in silico* studies of hSERT with clinical drugs.
- Molecular dynamics simulations showed average RMSD higher than human homologue hSERT, indicating increased binding instability among Sertraline-drSERTaa complex, as shown in figure 3.
- Sertraline has a higher binding affinity to drSERTaa (Danio Rerio serotonin receptors) than hSERT (Human Serotonin receptors), as shown in Table 1.
- This study highlights the possibility of lower active doses required for SSRIs such as sertraline to reach an effective dose on *Danio rerio*. Further studies investigating measured concentrations in wastewater close to populations of Zebrafish may prove useful in identifying if any populations are at risk.

Footnote

[1] Singh, H.K. and Saadabadi, A. 2023. Sertraline. StatPearls. Treasure Island (FL), StatPearls Publishing

[2] Norton, W.H., Folchert, A. and Bally-Cuif, L. 2008. Comparative analysis of serotonin receptor (HTR1A/HTR1B families) and transporter (slc6a4a/b) gene expression in the zebrafish brain. J Comp Neurol 511(4) 521-542.
[3] Jumper, J. et al. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596(7873) 583-589.
[4] Studer, G. et al. 2020. QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics* 36(6) 1765-1771.
[5] Zhou, Z. et al. 2009. Antidepressant specificity of serotonin transporter suggested by three LeuT-SSRI structures. Nat Struct Mol Biol 16(6) 652-657.
[6] Panchekha, P., Sanchez-Stern, A., Wilcox, J.R. and Tatlock, Z. 2015. Automatically improving accuracy for floating point expressions. *ACM SIGPLAN Notices 50(6) 1-11.*

QwikMD Gateway for Easy Simulation · EMBL-EBI · RCSB PDB PROTEIN DATA BANK · NIH National Institutes of Health · VMD Visual Molecular Dynamics · NAMD Scalable Molecular Dynamics · DeepMind