List of CBSB3 miniprojects in 2024

Ying Chi

Project 1. Novel antibody de-novo design and verification using suitable algorithms

The primary reason for designing new antibodies is to combat a variety of diseases, including cancer, infectious diseases, and autoimmune disorders. Due to the limited range of existing antibody libraries, which cannot cover all pathogens and disease types, there is a need to design new antibodies to enhance recognition and attack capabilities against specific pathogens or diseases. Additionally, designing new antibodies can improve their affinity, specificity, and selectivity, thereby enhancing treatment efficacy and reducing side effects. Moreover, the creation of new antibodies can be used in the development of novel drugs and vaccines to address emerging pathogens and diseases.

As introduced in Feb 26 morning lecture, ProteinGenerator generates sequence-structure pairs with RoseTTAFold. The goals of this project are to tailor and enhance ProteinGenerator for novel antibody de-novo design, and eventually figure out a way to check the effectivity of results using merely computer (no need to do wet-lab experiment). Detailed objectives are as the following:

1. Learn the details of ProteinGenerator.
2. Think about the difference between normal proteins and antibodies.
3. Tailor ProteinGenerator for antibody de-novo design.
4. Using your own method in "3" to design 100 novel antibodies for a protein with this sequence:
   "MGRGLLRGLWPLHIVLWTRIASTIPPHVQKSVNNDMIVTDNNGAVKFPQLCKFC DVRFSTCDNQKSCMSNCSITSICEKPQEVCVAVWRKNDENITLETVCHDPKLPYH DFILEDAASPKCIMKEKKKPGETFFMCSCSSDECNDNIIFSEEYNT", whose epitope is the 87-101 amino acid sites. Each new design should be the best from 30 mutational variants. That it, the searching range covers 3000 antibodies.
5. Draft a verification plan using computer only, which may cover molecular interaction from structure perspective & molecular dynamics, etc. And provide justification.
6. Write down and submit the 100 de-novo design results, whole design & verification procedure & major findings, and discussion together with future work.

You are free to use any modelling and algorithmic approaches you learned either in courses or by yourself in spare time (anyway, learning can happen anywhere. Following our guidance, you should take the initiative).

Main reference: ProteinGenerator github website
https://github.com/RosettaCommons/protein_generator

Ming Chen

Project 2. Modelling biological networks using
2A: differential equations / 2B: rule-based models / 2C: Petri nets

Note: these are **three separate miniprojects**, with up to 3 students each, under one common topic. If you choose them among your preferences, please indicate 2A, 2B or 2C. If you only indicate 2, it will mean you're OK with all of them. You can also indicate 2AB, 2AC or 2BC.

The project will focus on modelling biological interactions from the molecular level up, incorporating processes like gene expression, regulation, translation, protein modification, and transportation.

The project will begin by familiarizing students with several different techniques for modelling biological systems, including differential equation-based models, rule-based models and Petri net models. Once students have a solid foundation in these modelling techniques, they will work in groups to select a biological system to study, developing a model of interactions between the system's various components using one or more of these techniques.

The students will be encouraged to select biological interactions that are not well-understood, or that are challenging to study experimentally, and to explore various model parameters and assumptions to test hypotheses about how the system operates.

Finally, students will be asked to document their modeling efforts in a final written report, which should clearly describe their chosen biological system, model assumptions, and any parameterization and analysis techniques employed. The report should also include a detailed discussion of the results, their accuracy, and their potential implications for future research.

Tasks:

1. Identify and research the biological interactions of interest, such as gene expression, regulation, protein modification, transportation, signalling or interaction in the relevant biological system.
2. Use the Systems Biology Markup Language (SBML) or SBGN to generate a mathematical model of the biological system and the interactions to be investigated.
3. Choose a modelling methodology that is best suited to study the biological system. Options can include differential-equation based models, rule-based models, or Petri net models. The approach selected should be justified based on the complexity of the system being studied and the type of information one hopes to generate from the model.
4. Use a modelling tool such as COPASI or CellDesigner or Petri net tool to simulate and analyse the model generated. This will allow to observe the behaviour of the system over time and under different conditions.
5. Visualize and interpret the results obtained through the simulation and analysis process. Results may include the dynamics of one or more biological variables, sensitivity analysis to explore the impact of uncertainties and parameter variations in the model, and the emergence of emergent properties that cannot be easily identified from simple biological experiments.

Some useful links:

BioModels Database: https://www.ebi.ac.uk/biomodels/
BRENDA Database: https://www.brenda-enzymes.org/
CellDesigner: https://celldesigner.org/
COPASI: https://copasi.org/
Petri Nets World: https://www.informatik.uni-hamburg.de/TGI/PetriNets/index.php

Gedi Lukšys

Project 3. Analysis of our computational model in a schema-based learning and decision making task

Note: this miniproject has a maximum of 4 students, more than the usual 3.

When people make decisions, the choices they make as well as the time they spend on considering those choices depend on the range of menu options and their subjective evaluation and confidence of each option. In a schema-based learning and decision-making task (briefly discussed in the Jan 26 workshop and miniproject introductions on Feb 28) participants need to make several decisions: choose the schema (author of a painting or a quote) they focus on in each decision screen (when picking the first item) and find items 2, 3 and 4 that match the chosen schema (author). The choices and their timing depend on various factors: previous schema familiarity, exploration pattern, and incentive conditions like novelty and risk (e.g. do they lose all of their bonus by making a single mistake or not).

During previous years students have developed and tested a model based on **error-based learning** (similar to reinforcement learning) and **drift diffusion model** components that would capture a lot of details of task performance as well as would have more solid justification. We also have nearly concluded parameter estimation of finding the best fitting parameters for each individual that would make model's performance match that individual's performance the best. In this project you will not need to build a new model, but will need to understand, modify, test and interpret the existing model, programmed in R. We will also provide you with estimated parameter values that are already available (and nearly all should be ready in April), as that may help your testing and analyses.

Your aim will be to modify and test the model in order to address the following questions (each student in the group will likely take one of these questions):

- Some participants decide that trying to get a bonus (which requires matching items from the same schema) is too hard or time-consuming and give up on doing that. Instead, they try to pick items from higher paid schemas (which is rather easy to remember) and choose them as fast as possible, so that they can go through more decision rounds, thereby increasing their payment. We would like to investigate this phenomenon from computational and experimental perspectives, and test which parameter settings it corresponds to or which model modifications it requires.

- We assume that participants take into account schema payoff in choosing items, not only schema confidence, which is the basic model variable. This is similar to reinforcement learning, where the probability of choosing an item is positively related to its value. However, alternative action choice models exist. For example, *win stay-lose shift* is a simple approach suggesting that if participants succeed with a particular schema, they will try the same schema next time, whereas if they fail, they will shift to a different schema. Here we would like to test how such action selection may compare to purely value-based action selection.

- We know from the questionnaire what is each participant's prior familiarity of each schema (authors of paintings or quotes). In principle if participants have higher prior familiarity they should start with higher confidence of those schemas. We would like

to do thorough testing of how different initial confidence settings (e.g. based on schema familiarity, payoff or fixed) could impact model performance and how it compares to participants' performance in the experiment.

- We know from the questionnaire how tired and motivated each participant feels before the experiment. In modelling, tiredness or motivation can affect different things: e.g. speed of evidence integration, how many options are examined, how quickly a decision is made (which can be realised with decreasing decision threshold), etc. We would like to examine how these modifications can influence model performance and whether that matches performance differences observed between participants with different motivation/tiredness levels.

Detailed model description will be provided in the introduction session on Feb 28 at 10-12.

This is a recent relevant paper on drift-diffusion models, https://elifesciences.org/articles/56938 and a few others about models with multiple alternatives, https://www.pnas.org/content/108/33/13852 and https://www.nature.com/articles/nn.2123

And here is a nice simulator of drift-diffusion models (even if applied to 2 choice tasks): https://longdecision.github.io/DDM_tutorial/

Project 4. Computational modelling of news article selection using MyNewsScan

Note: this miniproject has a maximum of 4 students, more than the usual 3.

In today's society, people often find it difficult to read news outside their social circle or comfort zone. The questions of whether limited availability or active avoidance of such information determines its limited reach and which neurocognitive factors contribute to this outcome are of huge importance but not adequately studied. MyNewsScan.eu is a news aggregator platform that we have been developing to tackle such questions. We performed experiments asking participants to select, read and evaluate news articles on MyNewsScan: either in our Edinburgh lab with collection of biometric information such as eye movements, heart rates and emotional expressions, or online (recruiting from Malaysia, Germany and UK). We investigated which factors (e.g. article's topic, length, source, clarity, familiarity, likability, perceived accuracy and usefulness) affect participants' choices of articles and engagement in MyNewsScan, and if their behavioural patterns could be predictive of personality traits and neuropsychiatric conditions such as anxiety or depression.

So far, our analyses only used regular statistical approaches such as mixed effects models as well as a bit of machine learning such as principal component analysis and supervised learning. However, we're also interested to uncover computational processes underlying article selection and evaluation at a more mechanistic level. In this miniproject we will seek to combine statistical analyses with computational modelling of article selection as well as ratings provided in post article questions aiming to understand the factors determining participants' article choice:

- Do participants simply select articles based on topic, length, location or other similar factors? (with clear bias towards certain topics, short or long articles, those display on the top or at the bottom)

- Or do they instead perform a more rational assessment of expected payoff they are likely to get by picking articles of topics/length/location based on trial & error-based assessment of their success so far? Value-based models tracking expected points per time spent (like reinforcement learning) would be relevant here.

- Or perhaps is it the assessment of articles of those categories (such as likability or usefulness) that determines their decisions rather than points/money they get?

- Perhaps do they employ *win-stay/lose-shift* or similar strategies, e.g. sticking with types of articles where they were successful in answering questions and moving away from those where they failed?

- Finally, a particular question of interest is why (some) participants keep attempting questions for articles where they didn't get all questions right. When they do that, what are the factors predicting their further attempts? It could be personality traits that we assess from the questionnaire. Or it could also be some more sophisticated calculation based on number of points they may still get by trying, chance of getting the answers right, and amount of time they spend for that?

It is expected that each student will try to address 1 or 2 of these questions, for which they will need to do data analysis and/or develop new (but rather simple/proof of concept) models. Then different students in this miniproject should compare the performance of the models they develop and see which of them may better match the data. As here the focus is on new model development, we don't expect students to perform model parameter estimation or detailed analysis, only investigate at a more qualitative level how different model configurations/parameters may influence performance.

Most references for modelling of decision making in project 3 are also relevant here. In addition, there are several background references for MyNewsScan (not directly related to modelling).

Vosoughi et al., "The spread of true and false news online", Science 2018
https://science.sciencemag.org/content/359/6380/1146.full

Insel, "Digital phenotyping: a global tool for psychiatry", World Psychiatry 2018
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6127813/pdf/WPS-17-276.pdf

TED talk about filter bubbles:
https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles

MyNewsScan's old version accessible in China: http://mynewsscan.net/schema_local.php

Rob Young

Project 5. Provision of FAIR principles across publicly available models

In an ideal world, all models published by researchers should follow the FAIR principles (Findable, Accessible, Interoperable, Reusable). In this project you will search for a number of published models and assess their compliance with the FAIR principle. You should

attempt to run the models you access and determine whether (and how easy!) it is to replicate published results for these models. Note that you should only select models where results are already available to compare with your own work. If you need to make modifications or fix bugs within the model, please record that. You should identify and describe models which examine each of the following biomedical systems:

1. Tissue metabolism

2. Interaction of the immune system with disease (any disease)

3. Disease transmission across a population

4. Neuronal activity in a non-human species

5. Any other model of your interest

USEFUL LINKS:

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

https://www.go-fair.org/how-to-go-fair/

Hugo Samano

Project 6. Modelling Gene-Transposon element fusions

Transposable Elements (TEs) are mobile genomic elements present in eukaryotes that take about half of the human genome and play an important role in multiple biological processes like embryonic development, oncogenesis and other critical life processes. TE elements can, in principle, jump to any position of the genome, representing a risk for genome stability (Nusse and Varmus, 1982) and an opportunity for gene diversification.

The study of Gene-TE fusions is limited to few natural examples but is abundant in genetic engineering examples produced by insertional mutagenesis to identify oncogenes or tumour suppressor genes (Theodorou, V. et al., 2007). Moreover, the population diversity of Gene-TE fusions is poorly understood due to few and yet insufficient tools to detect these genomic events. This is initially caused by the absence of Gene-TE ground-truth datasets or accurate simulations.

The aim of this project is to model gene-TE fusion events in the human genome. To do this, you will use the latest annotation of the human genome and the latest consensus list of TE elements for the human genome as in Dfam (https://www.dfam.org/).

You should consider for your simulation:
a) mutations (at known, real, mutation rates),
b) insertions at different rates,

c) gene essentiality information (https://v3.ogee.info/#/home) to discard unlikely feasible fusions and
d) intron/exon architecture (fusions can happen within, exons, introns or a combination).

Note: You can focus on a single human chromosome for technical reasons but your parameters and code should be functional for the whole genome.

Relevant literature:

- Nusse R. and Varmus, H.E. (1892) Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell, Volume 31, Issue 1, 99-109
- Theodorou, V., et al. (2007) MMTV insertional mutagenesis identifies genes, gene families and pathways involved in mammary cancer. Nat Genet 39, 759–769.

Project 7. Modelling phosphorylated peptides to protein domains.

Cell signalling is regulated by transient low affinity motif-domain interactions including those mediated by phosphorylated motifs. Phosphorylation/dephosphorylation changes can define the binding in a switch manner. The affinity of a phosphorylation motif can be increased or attenuated by the surrounding residues. Oriented peptide array libraries are used to infer relative affinities of a mutated (and phosphorylated) peptide to a particular domain. This is costly and time consuming making it rather unpopular (Huang, H., 2008) and difficult to scale up. An alternative is to use protein docking; however, most tools are not yet available to use non-standard residues in *de novo* modelling. Therefore, the clear alternative is to use molecular dynamics as this method allows the use of phosphorylated residues.

For this project, you will develop a pipeline to explore the potential binding motifs to a fixed domain, such as SH2, PTB, C2 or HYB domains. To do this, you will prepare a list of motif candidates based on curated databases like ELM (Kumar, M., et al., 2022) or non-curated databases like PhosphositePlus (https://www.phosphosite.org/); then, create topology files using CHARMM-GUI (https://www.charmm-gui.org/) and proceed to do a simulation with GROMACS. Define your own thresholds for acceptable binding or displacement of the motifs and compare with the True Positives. Include some True Negatives to assess the specificity of your pipeline. Optionally, you can assess the affinity with tools like https://foldxsuite.crg.eu/.

Relevant literature:

- Huang H, Li L, Wu C et al (2008) Defining the specificity space of the human SRC homology 2 domain. Mol Cell Proteomics 7(4):768–784.

- Kumar, M., et al., (2022) The Eukaryotic Linear Motif resource: 2022 release, Nucleic Acids Research, Volume 50, Issue D1, 7.

Yong Wang

Project 8. Modelling of intrinsically disordered protein

Intrinsically disordered proteins (IDPs) are essential components in all life forms, assuming various functional roles despite lacking a fixed three-dimensional structure. Unlike proteins with well-defined folds, IDPs transition through numerous conformations, forming a dynamic ensemble. This variability has led to their underrepresentation in databases like the Protein Data Bank and a scarcity of computational tools for predicting their ensemble conformations based on their amino acid sequences. IDPs challenge traditional views on protein structure-function relationships due to their sequence diversity, transient structures, and involvement in many diseases. Although recent advances have enabled the prediction of folded protein structures at a large scale, understanding IDPs' structure and its conformational dynamics remains limited.

In this project, we aim to leverage the AI algorithm and physics-based molecular dynamics (MD) simulation to predict the conformational ensembles and properties of IDPs from their sequences. MD simulation represents a groundbreaking method in computational biology, allowing scientists to study the dynamic behavior of molecules in computer. This technique is particularly valuable in the exploration of IDPs. Our goal is to predict the conformational ensembles and properties of IDPs from their sequences, thereby contributing to a deeper understanding of their biological functions and roles. Participants will have the opportunity to delve into the modeling and simulation of IDPs, acquiring practical experience in state-of-the-art computational biophysics techniques.

The sequence for a representative IDP which we will work on is pasted below:

EEEMASSTSDSGEESDSSSSSSSTSDSSSSSSTSGSSSGSGSSSSSSGSTSSRSRLYRKKR
V

Tasks:

1) **Structure prediction with the latest AI algorithm**: Apply AlphaFold2 or other tools to predict the structure of the given IDP sequence.
2) **Structural analysis:** Examine the per-residue confidence metric to assess the prediction's reliability and predict its disorder properties.
3) **Flexible structural ensemble modeling:** Use molecular dynamics (MD) simulations to model the IDP's structural ensemble. To reduce the computational cost, we will employ the coarse-grained CALVADOS model for this task.
4) **Visualization of protein dynamics**: Use protein visualization tools such as PyMol to display the protein's dynamic behaviour over time.
5) **IDP properties analysis:** Investigate the properties (e.g. Root Mean Square Deviation, Radius of Gyration and residue-residue contact map, etc.) of the specified IDP, focusing on its structural dynamics.

Some useful links:

- https://pymol.org/
- https://github.com/KULL-Centre/_2023_Tesei_IDRome
- https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb

- Jumper et al. "Highly accurate protein structure prediction with AlphaFold." Nature (2021) doi: 10.1038/s41586-021-03819-2
- Lotthammer et al. Nature Methods (2024). https://www.nature.com/articles/s41592-023-02159-5
- Tesei et al. "Conformational ensembles of the human intrinsically disordered proteome:", Nature (2024). https://www.nature.com/articles/s41586-023-07004-5

Pavel Loskot

Project 9. A Genome-Wide Association Study (GWAS)

The variations in genome sequences such as nucleotide substitutions, insertions and deletions can explain or be indicative of rare diseases and other functional changes. Presently, these variations are analyzed statistically to associate single nucleotide polymorphism (SNP) with observed phenotypic traits. As the DNA sequences have been a lot cheaper and widely available in recent years, a vast amount of data are available for these studies. The aim of this project is to either define and test own hypothesis, or to replicate the results from the literature to verify the claims there-in. This will require identifying and downloading relevant datasets, and then using the GWAS tools such as plink to perform the statistical analysis including checking the data formatting and quality, missing data, relatedness of subjects, equilibrium, the need for stratification and other. The outputs of plink are then further analyzed with R or Python to obtain polygenic scoring and so on. Having access to Unix is an advantage, and so is having understanding of the basic statistical concepts (regression, correlation, biases, etc.)

[1] Uffelmann, E., Huang, Q.Q., Munung, N.S. et al. Genome-wide association studies. Nat Rev Methods Primers 1, 59 (2021). https://doi.org/10.1038/s43586-021-00056-9
[2] Jahad Alghamdi, Sandosh Padmanabhan, Chapter 12 - Fundamentals of Complex Trait Genetics and Association Studies, Handbook of Pharmacogenomics and Stratified Medicine, Academic Press, 2014, Pages 235-257, ISBN 9780123868824.
[3] Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int J Methods Psychiatr Res. 2018 Jun;27(2):e1608. doi: 10.1002/mpr.1608. Epub 2018 Feb 27. PMID: 29484742; PMCID: PMC6001694.
[4] L.J. Raffel, M.O. Goodarzi, Diabetes Mellitus, Reference Module in Biomedical Sciences,
Elsevier, 2014, https://doi.org/10.1016/B978-0-12-801238-3.05558-6.

Project 10. Whole-Cell Modelling

Enabled by sheer amount of datasets and computing resources, the aim of whole-cell modelling is to create a more complete digital twin of a biological system. This obviously has a great potential in completely transforming how in silico experiments are planned and conducted in order to gain understanding about biological processes and phenomena. The aim of this project is to map state of the art in whole-cell modelling, define what can constitute whole-cell modelling, outline the methodologies and strategies that are being used, list the requirements in computing resources and datasets as well as identify the limitations that may be overcome with further technology progress. A small contribution could be identifying new applications, which can be created upon or benefit from whole-cell modelling.

[1] Jonathan R Karr, Koichi Takahashi, Akira Funahashi, The principles of whole-cell modeling, Current Opinion in Microbiology, Volume 27, 2015, Pages 18-24, https://doi.org/10.1016/j.mib.2015.06.004.
[2] Arthur P Goldberg, Balázs Szigeti, Yin Hoon Chew, John AP Sekar, Yosef D Roth, Jonathan R Karr, Emerging whole-cell modeling principles and methods, Current Opinion in Biotechnology, Volume 51, 2018, Pages 97-102, https://doi.org/10.1016/j.copbio.2017.12.013.
[3] Derek N Macklin, Nicholas A Ruggero, Markus W Covert, The future of whole-cell modeling, Current Opinion in Biotechnology, Volume 28, 2014, Pages 111-115, https://doi.org/10.1016/j.copbio.2014.01.012.
[4] Balázs Szigeti, Yosef D. Roth, John A.P. Sekar, Arthur P. Goldberg, Saahith C. Pochiraju, Jonathan R. Karr, A blueprint for human whole-cell modeling, Current Opinion in Systems Biology, Volume 7, 2018, Pages 8-15, https://doi.org/10.1016/j.coisb.2017.10.005.
[5] Oliver Purcell, Bonny Jain, Jonathan R. Karr, Markus W. Covert, Timothy K. Lu; Towards a whole-cell modeling approach for synthetic biology. Chaos 1 June 2013; 23 (2): 025112. https://doi.org/10.1063/1.4811182

Project 11. Inferring Causality from Multivariate Time-Series Data

Causality is key in formulating and testing hypotheses, and to provide explanations e.g. why an event occurred, why particular data are observed, or as counterfactual reasoning, what might change should the experiment be setup differently. The idea of this project is to obtain a dataset representing multivariate time-series observations from some biological experiment or simulation, and then to obtain the corresponding structural causal model indicating how segments of time-series are causally related (or not). This project is mostly numerical, using either R or Python, but also requires basic knowledge of statistical testing of hypotheses, and understanding of causal reasoning to provide explanations and find causal associations.

[1] Runge, J., Gerhardus, A., Varando, G. et al. Causal inference for time series. Nat Rev Earth Environ 4, 487–505 (2023). https://doi.org/10.1038/s43017-023-00431-y
[2] Determining causality in correlated time series by Atalanti Mastakouri, Dominik Janzing. https://www.amazon.science/blog/determining-causality-in-correlated-time-series
[3] Castro, M., Mendes Júnior, P.R., Soriano-Vargas, A. et al. Time series causal relationships discovery through feature importance and ensemble models. Sci Rep 13, 11402 (2023). https://doi.org/10.1038/s41598-023-37929-w
[4] Moraffah, R., Sheth, P., Karami, M. et al. Causal inference for time series analysis: problems, methods and evaluation. Knowl Inf Syst 63, 3041–3085 (2021). https://doi.org/10.1007/s10115-021-01621-0

Duncan MacGregor

Project 12. A differential equation-based system model of water intake and osmotic homeostasis

Mechanisms based mainly in the brain's hypothalamus act to regulate thirst, water loss, and salt excretion, in order to maintain the homeostasis of osmotic pressure (the balance between salt and water). Water is regularly lost through respiration and perspiration. Salt is gained through feeding, and lost through perspiration. The hormone vasopressin, secreted by

hypothalamic neurons, acts at the kidneys to regulate water loss in response to signals that encode osmotic pressure. Oxytocin similarly acts to regulate salt excretion, and other signals including aldosterone act to regulate thirst. To understand how all these elements interact we want to build a simple model of this system starting from very basic assumptions and gradually building complexity.

The focus of this project will be on modelling the process of water consumption, including drinking, absorption, loss through the kidneys, and the transfer of water between extracellular and intracellular compartments. As well as the modelling this will require gathering from published sources quantitative experimental data with which to base the mechanisms and test the behaviour of the model.

Beginning with supplied tutorial notes on differential equation-based system modelling, your task will be to gradually build and test a simple model of this system, implemented in the HypoMod software. The objective will be to assemble sufficient components to simulate the basic vasopressin and thirst control of water intake and loss, able to maintain osmotic homeostasis under a realistic simulation of periodic drinking events.

Relevant literature:

- Ramsay DJ. The importance of thirst in maintenance of fluid balance. Baillieres Clin Endocrinol Metab. 1989 Aug;3(2):371-91. doi: 10.1016/s0950-351x(89)80008-4.
- Zimmerman CA, Leib DE, Knight ZA. Neural circuits underlying thirst and fluid homeostasis. Nat Rev Neurosci. 2017 Aug;18(8):459-469. doi: 10.1038/nrn.2017.71.
- Verbalis JG. Disorders of body water homeostasis. Best Pract Res Clin Endocrinol Metab. 2003 Dec;17(4):471-503. doi: 10.1016/s1521-690x(03)00049-6.

Project 13. Integrate and fire based spiking model of the depolarising afterpotential in oxytocin neurons

Magnocellular oxytocin neurons of the hypothalamus manufacture and secrete oxytocin hormone, with their main hormone output secreted into the blood plasma at the posterior pituitary gland. The simplest function of oxytocin is in regulating osmotic homeostasis, signalling the kidneys to increase salt excretion, but oxytocin has multiple physiological functions including in reproduction and appetite regulation. Secretion from the pituitary terminals is triggered by spikes (action potentials) that are conducted down the axon and trigger calcium entry and exocytosis of stored hormone containing vesicles. Secretion and more generally the signal processing of these neurons is highly dependent on both the rate and patterning of their spiking, and this is in turn is dependent on their well-studied intrinsic electrophysiological properties, including the hyperpolarising afterpotential (HAP) and afterhyperpolarising potential (AHP). These have been extensively modelled and we have a good understanding of their effect on spike patterning and signal processing. However, a proportion of oxytocin neurons also have a depolarising afterpotential (DAP). This is less well modelled, and its underlying mechanism is also less well understood.

The task for this project, using the HypoMod software from the week 2 lecture and practical, is to implement an integrate-and-fire based oxytocin neuron model and add a DAP, testing different formulations of the model based on different putative underlying mechanisms, and then use this to fit recorded neurons which show a DAP in their spike interval analysis. The model can then be used to test the signal processing properties of the DAP, examining the spiking output in response to simulated input signals.

Relevant literature:

- Teruyama R, Sakuraba M, Kurotaki H, Armstrong WE. Transient receptor potential channel m4 and m5 in magnocellular cells in rat supraoptic and paraventricular nuclei. J Neuroendocrinol. 2011 Dec;23(12):1204-13. doi: 10.1111/j.1365-2826.2011.02211.x.
- Maícas Royo J, Brown CH, Leng G, MacGregor DJ. Oxytocin Neurones: Intrinsic Mechanisms Governing the Regularity of Spiking Activity. J Neuroendocrinol. 2016 Apr;28(4). doi: 10.1111/jne.12358.
- Gagnon A, Walsh M, Okuda T, Choe KY, Zaelzer C, Bourque CW. Modulation of spike clustering by NMDA receptors and neurotensin in rat supraoptic nucleus neurons. J Physiol. 2014 Oct 1;592(19):4177-86. doi: 10.1113/jphysiol.2014.275602.