Reinforcement learning-guided media design for optimised antibody production in Chinese hamster ovary cells

Anirudh Rao, Anirudh TP, Pooja Shankar, Pranathi R, Sidharthan S C

Abstract

Manufacturing biopharmaceuticals, particularly monoclonal antibodies (mAbs), is a complicated process impacted by various parameters, including the composition of the culture medium. The choice of media becomes extremely critical in the case of Chinese hamster ovary (CHO) cells, which are one of the most used mammalian cell lines for recombinant protein synthesis. The *in silico* design and optimisation of media to maximise antibody production offers an alternative to labour-intensive and trial-and-error-based laboratory experiments. In this work, we present a novel computational framework for designing media using a combination of reinforcement learning (RL) and constraint-based genome-scale metabolic modelling. We then demonstrate how our method can be used to design culture media to optimise the production of antibodies in CHO cells. We believe that this framework can be extended to other cell types and can also be used to optimise other performance parameters like growth rate and antibody glycosylation.

Keywords

Chinese hamster ovary cells, reinforcement learning, genome-scale metabolic modelling, media design, antibody production

---

Introduction

Biopharmaceuticals are recombinant proteins (including recombinant antibodies), nucleic acid-based products, or genetically engineered cell-based products. They are synthesised using live cells, unlike small molecules, which are chemically made. Chinese hamster ovary (CHO) cells are an important mammalian cell line in the biopharmaceutical industry, showing high stability in antibody expression and crucial post-translational modifications, mimicking the behaviour of human cells. These factors contribute to the stable production of large antibody yields, making it cost-effective for commercial use. CHO cells are highly compatible with genetic engineering techniques, allowing us to modify their genome and express recombinant proteins. The development of fermentation processes for production in CHO cells has reduced the risks of contamination and has simplified downstream processing. CHO cells are a gold-standard choice for producing a variety of biopharmaceuticals, including monoclonal antibodies, Fc fusion proteins, various hormones and even vaccines [1]. The titer levels of antibodies produced are affected by a variety of parameters, including cell line, culture conditions, medium composition, and processing standards. Among these, media plays a vital role in cell survival, proliferation, biopharmaceutical generation, and quality. Adequate nutrient supply and proper growth conditions are required for effective antibody production [2]. Traditional approaches to designing media for improving antibody production involve laborious and time-consuming trial-and-error-based experiments.

Using a computational approach is a useful alternative in terms of labour and time. However, the design space – different combinations of media components and concentrations, is too large to explore exhaustively. Machine learning methods, reinforcement learning (RL) in particular, offer a unique avenue to rationally explore the design space to find suitable optima. Applying constraint-based techniques like flux balance analysis (FBA) to a genome-scale metabolic model can be used for *in silico* simulations of the behaviour of cells in response to varying concentrations of media components. When coupled together, RL and FBA form a novel computational framework for the iterative design and testing of media in a rational manner, allowing for the optimisation of antibody production in CHO cells.

Methods

Flux balance analysis

A genome-scale metabolic model is a network representation of the metabolism occurring in an organism [3]. Flux balance analysis (FBA) is a mathematical approach used to analyse the flow of metabolites through a metabolic network. FBA simulates the steady-state distribution of fluxes in a metabolic network, which are the rates at which metabolites are produced or consumed by reactions [4].

This requires a stoichiometric matrix S, a matrix representing the stoichiometry of the metabolic network. The entry S_{ij} indicates the stoichiometric coefficient of the metabolite i in the reaction j. We then define the flux vector v to represent the fluxes / rates of each reaction in the network. Finally, we define the objective function Z as a linear combination of a set of reactions in the metabolic network.

As the system is assumed to be a steady-state, we impose the constraint that $S \cdot v = 0$. We can also impose linear constraints, called bounds, on the fluxes through the reactions in the network. This is represented as $v_L \leq v \leq v_U$, where v_L and v_U are vectors describing the lower and upper bounds on the flux vector v. This is visualised in Figure 1 in the Annexe. FBA solves the linear programming problem given by

$$\max Z$$

$$\text{such that } S \cdot v = 0 \text{ and } v_L \leq v \leq v_U$$

For our work, we have chosen the iCHO2291 genome-scale metabolic model of CHO cells [5]. This incorporates data about enzyme kinetics and capacity into the constraints for FBA. This is a state-of-the-art model that accurately captures the intracellular fluxes in a CHO cell. To perform the FBA, we use the COBRApy library available in Python [6]. This has various user-friendly functions that help perform constraint-based modelling.

Reinforcement learning

Reinforcement learning (RL) is a branch of machine learning focused on optimising decision-making to maximise cumulative "rewards" in a given scenario [7]. In this, a system (called an "agent") to learn the best "actions" to take in an "environment" to achieve a goal. It learns by trying actions and seeing the new "state" of the environment. Each action results in a reward (positive or negative) based on its usefulness toward the goal defined by the user. The agent's main objective is to learn a strategy, called a "policy", to maximise its long-term rewards. The agent has to balance exploration (trying new actions) with exploitation (using known actions that give high rewards) to find the best policy. The agent learns over many attempts, adjusting its actions based on feedback from the environment, improving its policy for achieving the goal. This is visualised in Figure 2.

For our work, the environment is defined as the CHO cell genome-scale metabolic model. The state of the environment is the uptake rates of a defined set of media components.

The actions performed by the agent represent adjustments to the bounds of these uptake reactions. The reward is the flux through a user-defined objective reaction. In our case, the objective reaction is related to antibody production.

To allow the reward to provide feedback to the agent and improve future actions, we have implemented the Actor-Critic with Experienced Replay (ACER) reinforcement learning framework [8]. This consists of two neural networks, the actor and critic. The input given to the actor is the current state of the environment. Thus, the input layer of the actor neural network has a number of neurons equal to the number of uptake reactions of media components. The actor produces a set of actions, from a user-defined superset, for each media component, by updating the bounds of their uptake rates. This is used to define the new state of the environment and perform FBA, yielding a reward. The critic neural network takes in the new state and reward as its input and estimates the advantage of this new state over the previous ones. This is used to reinforce the actor to take better actions that yield states with higher advantage, and thus, higher rewards in the long run. This is carried out iteratively with the actor and critic getting trained in the process. This is visualised in Figure 3.

Media optimisation

Using the genome-scale model and the ACER framework, we wish to optimise antibody production in CHO cells. Before this, we constrain the growth rate of the cells to be in exponential phase. We then define the set of media components whose uptake rates are to be optimised using the `Model.medium` attribute provided by COBRApy. Four different objective functions for maximising antibody production were chosen as the four IgG antibody reactions present in the iCHO2291 model – `igg_hc`, `igg_lc`, `DM_igg[g]`, `igg_formation`. The first two reactions are the synthesis reactions of the heavy chain and light chain of the antibody. The next reaction is a demand reaction in the Golgi apparatus. The last reaction is the final formation of the antibody, putting together the heavy and light chains along with various glycan molecules for glycosylation. The ACER framework was then used to find the optimal set of media uptake rates to maximise the chosen objective function. The change in the uptake rates and the objective value were tracked across all the iterations of the reinforcement learning process. The maximised objective value was compared to the globally maximal objective value, which was obtained by allowing the cells to have access to unlimited resources at a constrained growth rate. Thus, four different media were designed by the RL framework, one for each of the objective reactions.

Comparative analysis of designed media

We identified the media components that have similar or varied uptake rates across the four designed media. To better understand how similar or different the four media are from each other, we performed principal component analysis (PCA). This is a statistical approach that reduces dimensionality while maintaining most of the variation in the data. It allows us to identify the contribution of individual media components to achieving the objective function.

Results and Discussion

Growth rate prediction

Using the default bounds of the iCHO2291 model, FBA was used to predict the growth rate (the flux through the `biomass_cho` reaction). This was found to be 0.032 h^{-1}, which is close to the experimentally observed exponential growth rate of 0.033 h^{-1} [9]. For media optimisation, the CHO cell growth rate was constrained to be 0.032 h^{-1}.

Media optimisation

The uptake rates of 25 media components was optimised for the four different objective functions. The media components are arginine, asparagine, aspartate, choline, cysteine, iron, glucose, glycine, bicarbonate, histidine, hypoxanthine, isoleucine, leucine, lysine, methionine, oxygen, phenylalanine, proline, pyridoxine, serine, threonine, tryptophan, tyrosine, and valine. The results of this optimisation can be found in the Annexe, along with a visualisation of the iterations of the RL framework. The RL agent converges to an optimal value in fewer than 70 iterations for each of the objective reactions, after which the reward plateaus. We can also observe the actions taken be the agent while updating the uptake rates. The maximal fluxes for `igg_hc`, `igg_lc`, `DM_igg[g]`, `igg_formation` are 0.1, 0.1, 0.05, and 0.05 $\mathrm{mmol\ gDCW^{-1}\ h^{-1}}$ respectively.

The three media components with the highest uptake rates when IgG heavy chain production is optimised were found to be folate, bicarbonate, and pyridoxine. When IgG light chain production is optimised, the top three media components were phenylalanine, bicarbonate, and asparagine. When the IgG demand reaction in the Golgi was optimised, the top three components were aspartate, glucose, and asparagine. When the final IgG formation was optimised, the top three components were found to be proline, threonine, and choline.

Bicarbonate appears in the top three for both the IgG heavy chain and light chain optimisations. It plays an important role in the stabilisation of pH in the culture medium. It also plays a role in the replenishment of TCA cycle intermediates. For example, pyruvate carboxylase makes use of bicarbonate to form oxaloacetate from pyruvate. Asparagine is the only amino acid that appears in the top three more than once. It is involved in post-translational N-linked glycosylation of proteins.

Comparative analysis of designed media

Lysine, threonine, and iron have the least varied uptake rates across the four media. Asparagine, phenylalanine, and folate have the most varied uptake rates.

After mean centering the media data, principal component analysis was performed. The first 3 principal components captured 44.3%, 34.6% and 21.1% of the total variance in the data. When the first two PCs were plotted against each other, no direct clustering among the media was seen. However, we could observe that considering PC1 alone led to the clustering of the media for `igg_hc` and `igg_lc` optimisation. Considering PC2 alone led to the clustering of the media for `DM_igg[g]` and `igg_formation`. Plotting the PCA loadings as a biplot showed that all the media components contribute nearly equally to the first two PCs. There was no obvious clustering of media components. These results can be found in the Annexe.

Future Directions

The current model only optimises a set number of reactions, primarily concentrating on amino acid consumption. However, it is known that other key metabolites play a role in antibody production (e.g., lipids). Hence, the framework could be expanded to explore more reactions to provide better results. Furthermore, the current values obtained from the model are in terms of rates. Providing the results in terms of concentrations will simplify the experimental preparation of the media and aid future work along these lines. Intracellular flux prediction can be greatly improved by utilising enzyme capacity FBA or ecFBA. This method of prediction incorporates enzyme kinetics information into the GSMM and adds enzyme capacity constraints within the flux balance network to reduce the variation in flux in a biologically significant manner. This can be achieved using new Python libraries like Ginkgo Bioworks' GECKO [10].

References

[1] Tihanyi, B., & Nyitray, L. (2020). Recent advances in CHO cell line development for recombinant protein production. Drug Discovery Today Technologies, 38, 25–34. https://doi.org/10.1016/j.ddtec.2021.02.003

[2] Wurm, F. M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. Nature Biotechnology, 22(11), 1393–1398. https://doi.org/10.1038/nbt1026

[3] Passi, A., Tibocha-Bonilla, J. D., Kumar, M., Tec-Campos, D., Zengler, K., & Zuniga, C. (2021). Genome-Scale metabolic modeling enables In-Depth understanding of big data. Metabolites, 12(1), 14. https://doi.org/10.3390/metabo12010014

[4] Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? Nature Biotechnology, 28(3), 245–248. https://doi.org/10.1038/nbt.1614

[5] Yeo, H. C., Hong, J., Lakshmanan, M., & Lee, D. (2020). Enzyme capacity-based genome scale modelling of CHO cells. Metabolic Engineering, 60, 138–147. https://doi.org/10.1016/j.ymben.2020.04.005

[6] Ebrahim, A., Lerman, J. A., Palsson, B. O., & Hyduke, D. R. (2013). COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Systems Biology, 7(1). https://doi.org/10.1186/1752-0509-7-74

[7] Sutton, R., & Barto, A. (1998). Reinforcement Learning: An Introduction. IEEE Transactions on Neural Networks, 9(5), 1054. https://doi.org/10.1109/tnn.1998.712192

[8] Wang, Z., Bapst, V., Heess, N., Mnih, V., Munos, R., Kavukcuoglu, K., & De Freitas, N. (2016). Sample Efficient Actor-Critic with Experience Replay. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1611.01224

[9] Ahn, W. S., & Antoniewicz, M. R. (2011). Metabolic flux analysis of CHO cells at growth and non-growth phases using isotopic tracers and mass spectrometry. Metabolic Engineering, 13(5), 598–609. https://doi.org/10.1016/j.ymben.2011.07.002

[10] Muriel, J. C., Long, C., & Sonnenschein, N. (2023). Simultaneous application of enzyme and thermodynamic constraints to metabolic models using an updated Python implementation of GECKO. Microbiology Spectrum, 11(6). https://doi.org/10.1128/spectrum.01705-23

Acknowledgement

We would like to thank the course instructor Dr. Meiyappan Lakshmanan for his valuable inputs during this project and throughout the course. We would also like to thank the teaching assistants for their tutorial sessions and feedback on this project.

Annexe

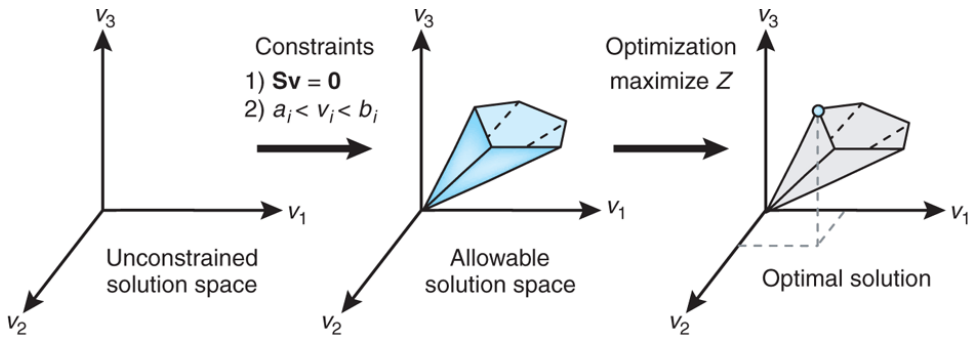

Figure 1: Flux balance analysis

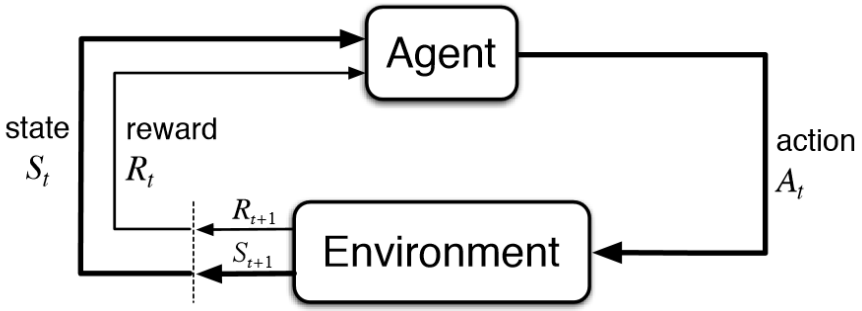

Figure 2: Reinforcement learning

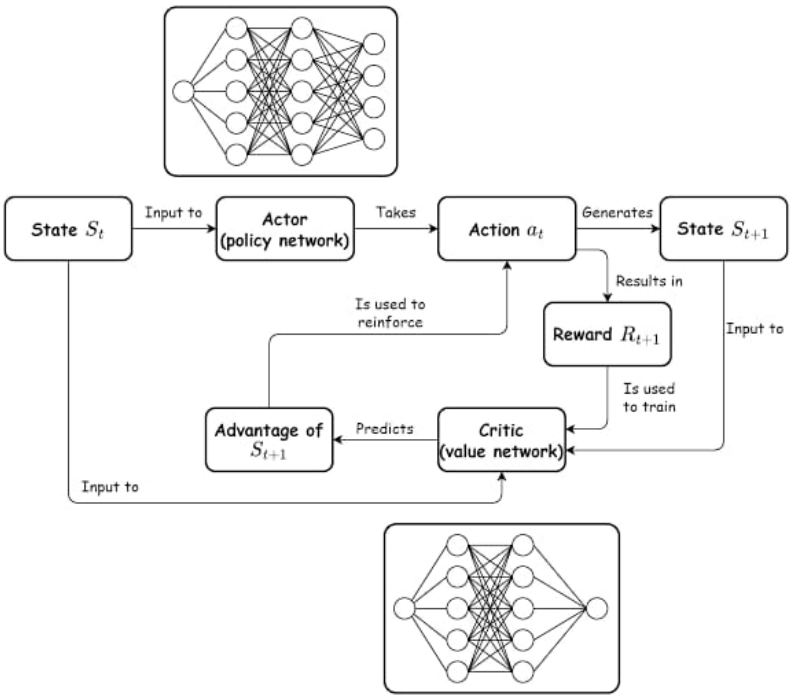

Figure 3: ACER framework

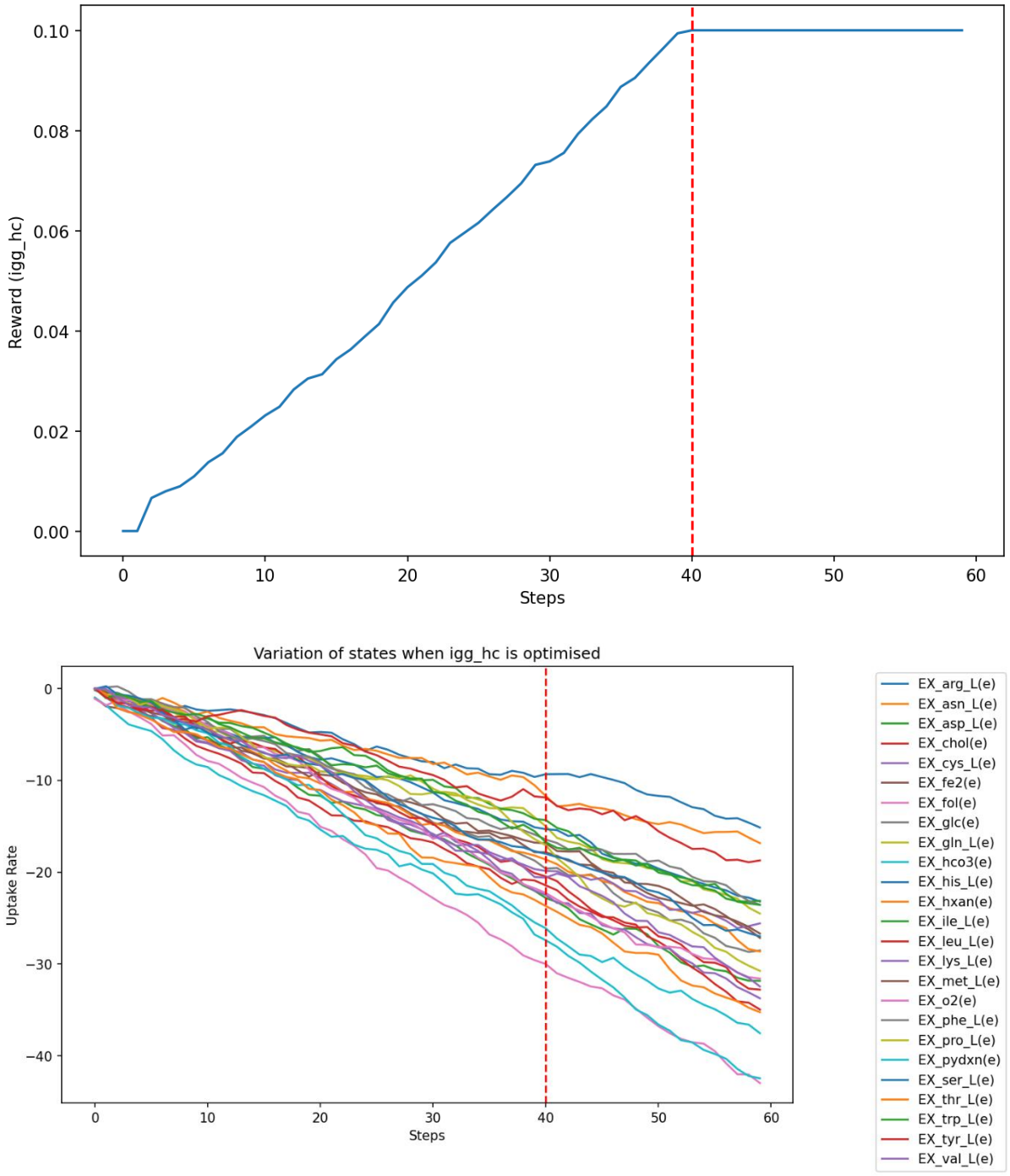

Figure 4: Optimisation of IgG heavy chain synthesis

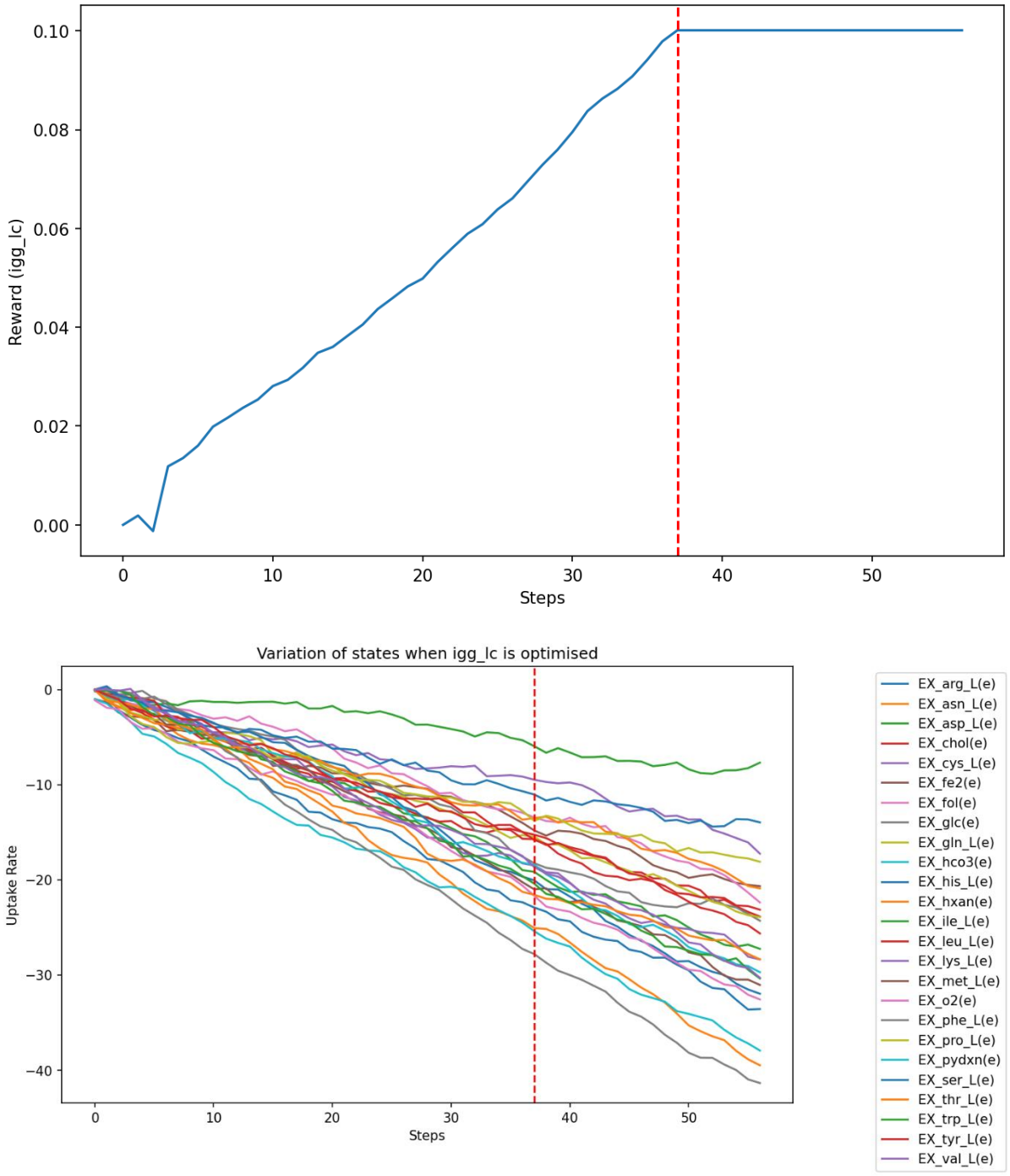

Figure 5: Optimisation of IgG light chain synthesis

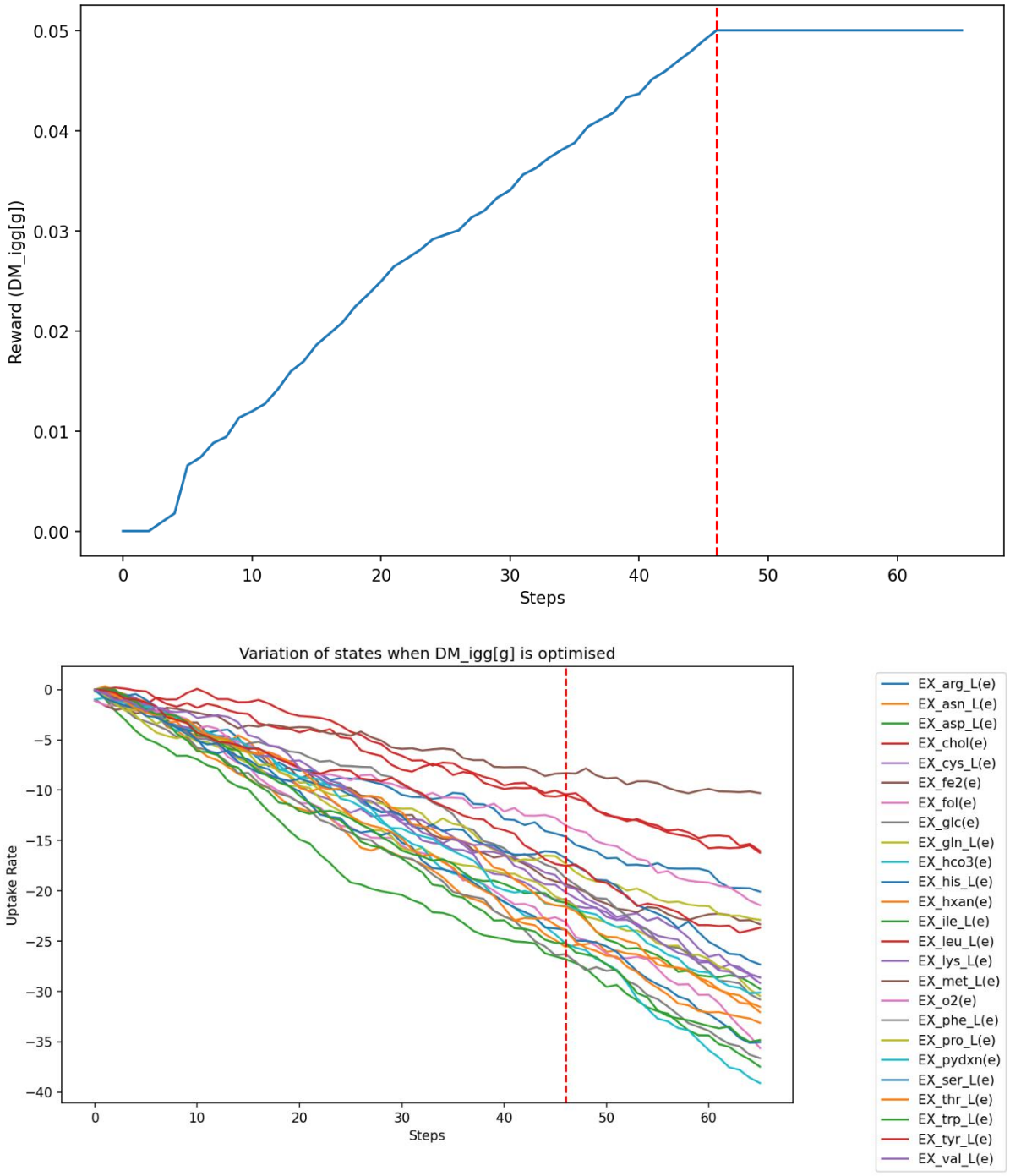

Figure 6: Optimisation of IgG Golgi demand

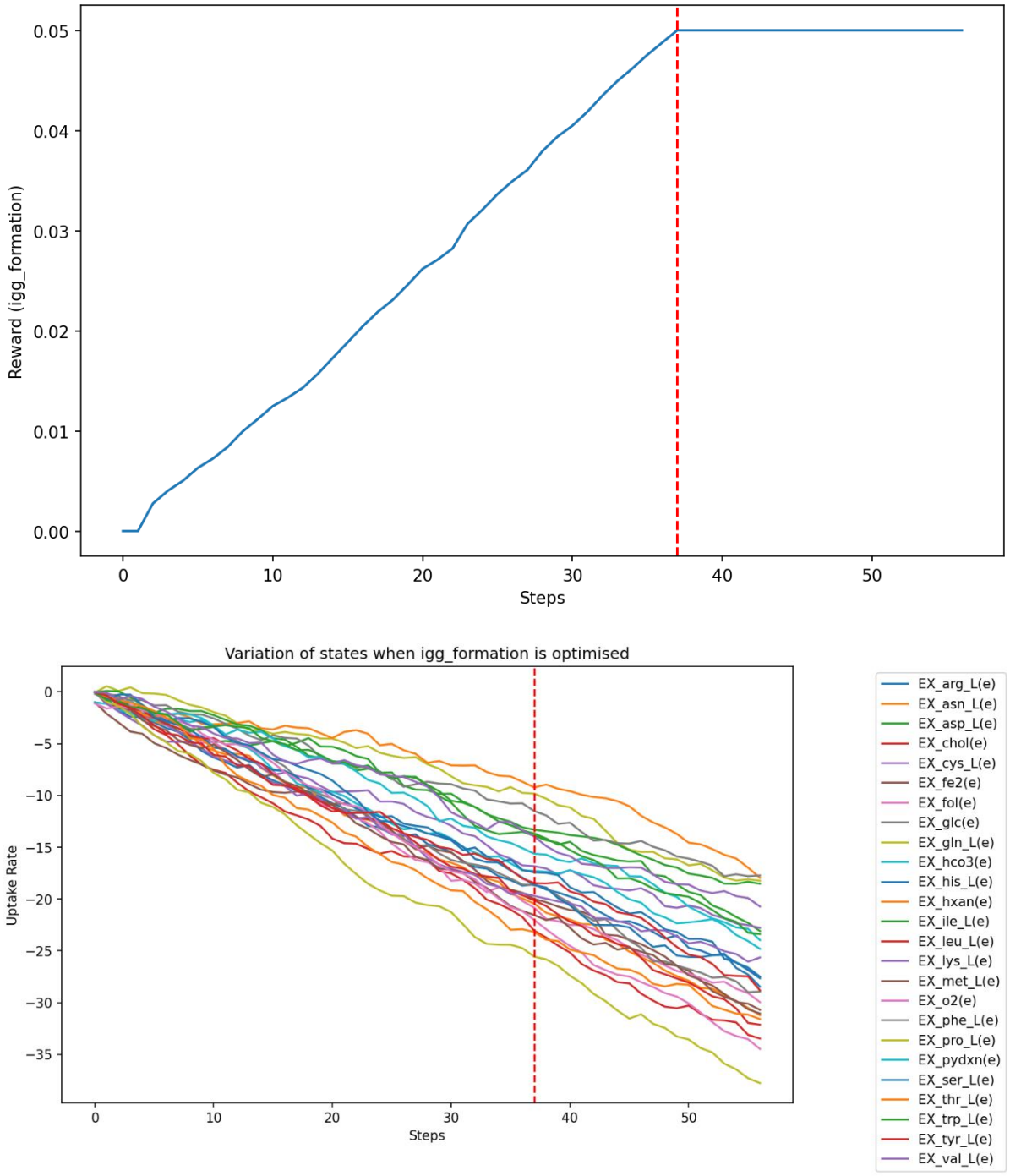

Figure 7: Optimisation of complete IgG formation

Table 1: RL-guided optimised uptake rates (in $mmol\ gDCW^{-1}\ h^{-1}$) of media components

	Objective Reaction			
	igg_formation	igg_hc	igg_lc	DM_igg[g]
EX_met_L(e)	-21.55483055	-16.85058212	-21.00032043	-8.304737091
EX_leu_L(e)	-20.13520432	-20.46927643	-15.81911755	-10.42608261
EX_chol(e)	-23.06950378	-21.44428825	-15.25889874	-10.48564053
EX_o2(e)	-22.03465652	-22.27285004	-21.66770744	-13.50377083
EX_arg_L(e)	-18.58504868	-9.338561058	-22.91124153	-14.60818672
EX_ser_L(e)	-18.66612053	-18.00676155	-11.03372288	-16.74649429
EX_pro_L(e)	-25.5796299	-17.06755066	-13.80424786	-17.30696297
EX_tyr_L(e)	-18.47743034	-11.86990356	-15.72382355	-17.52214622
EX_phe_L(e)	-11.5103035	-16.45296478	-27.79987526	-18.71604156
EX_fe2(e)	-19.99963379	-17.84293938	-14.81157398	-19.30975342
EX_lys_L(e)	-19.71253967	-19.8715992	-18.49480438	-19.54016495
EX_val_L(e)	-13.95855427	-20.60046768	-18.65909958	-20.13986397
EX_cys_L(e)	-16.82940102	-22.48837852	-9.481326103	-20.82244492
EX_gln_L(e)	-9.822813034	-15.00414085	-15.48595333	-20.95724678
EX_ile_L(e)	-13.31459236	-14.33504772	-5.957713604	-21.16860199
EX_pydxn(e)	-15.63251209	-26.11641502	-18.43297768	-21.49677849
EX_thr_L(e)	-23.18590736	-23.69534874	-21.572752	-21.56272316
EX_fol(e)	-20.84468651	-30.00884628	-13.43671227	-23.085989
EX_his_L(e)	-17.4277668	-15.31976604	-20.05235672	-23.84898949
EX_hxan(e)	-9.226536751	-18.62839127	-13.51349449	-23.91611671
EX_trp_L(e)	-13.88022709	-16.7432251	-20.40450287	-25.20640564
EX_hco3(e)	-17.2918663	-27.42977524	-25.39838791	-25.22895241
EX_asn_L(e)	-20.45246696	-11.70901489	-25.06777573	-25.56642723
EX_glc(e)	-18.65927315	-19.62451935	-18.25354767	-26.29706764
EX_asp_L(e)	-13.67327881	-22.8206234	-19.15742874	-26.78807449

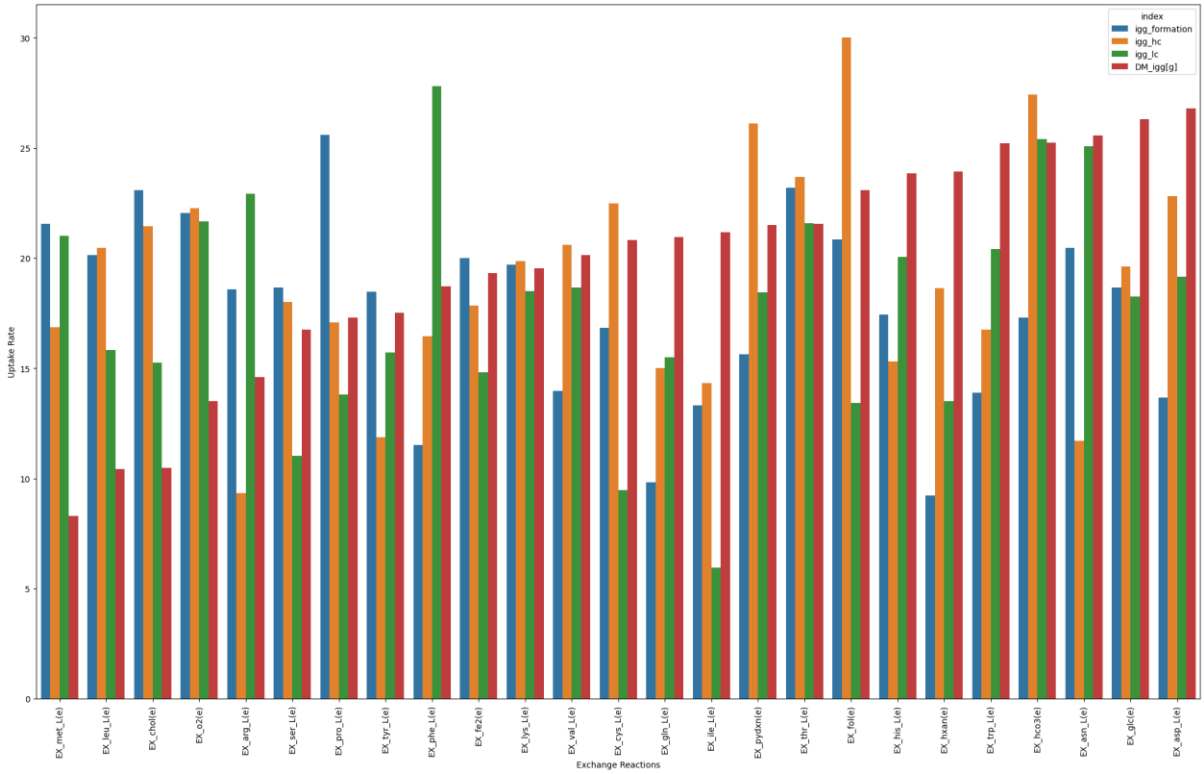

Figure 8: Comparison of uptake rates (in $mmol\ gDCW^{-1}\ h^{-1}$) across the four designed media

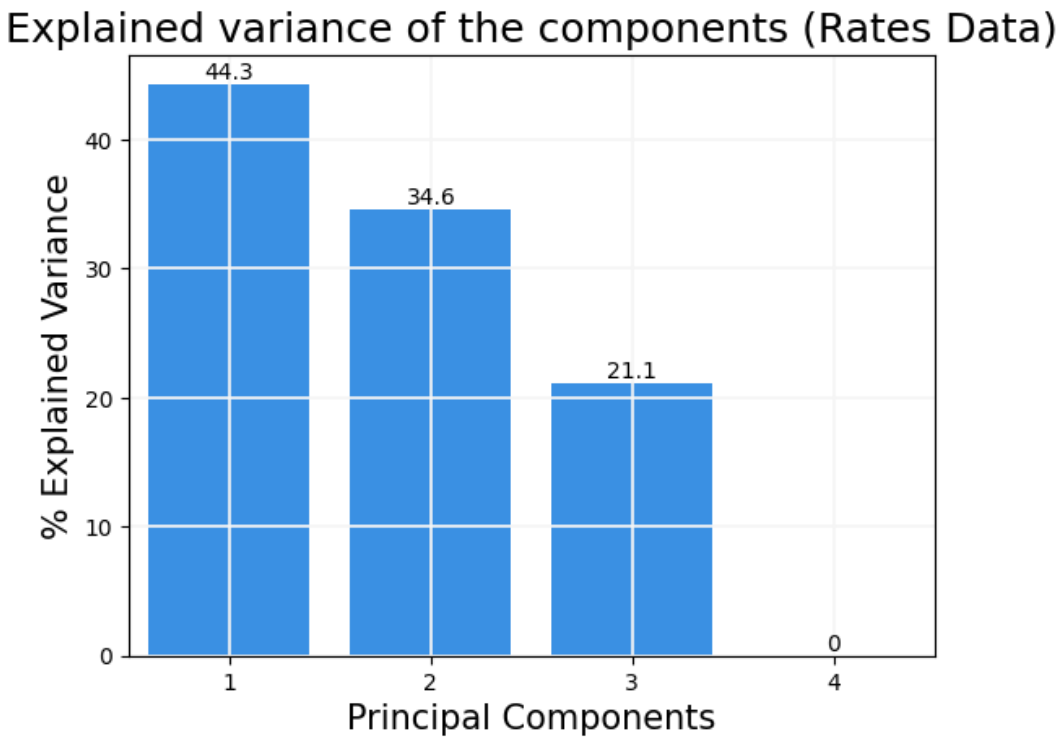

Figure 9: Variance explained by the principal components

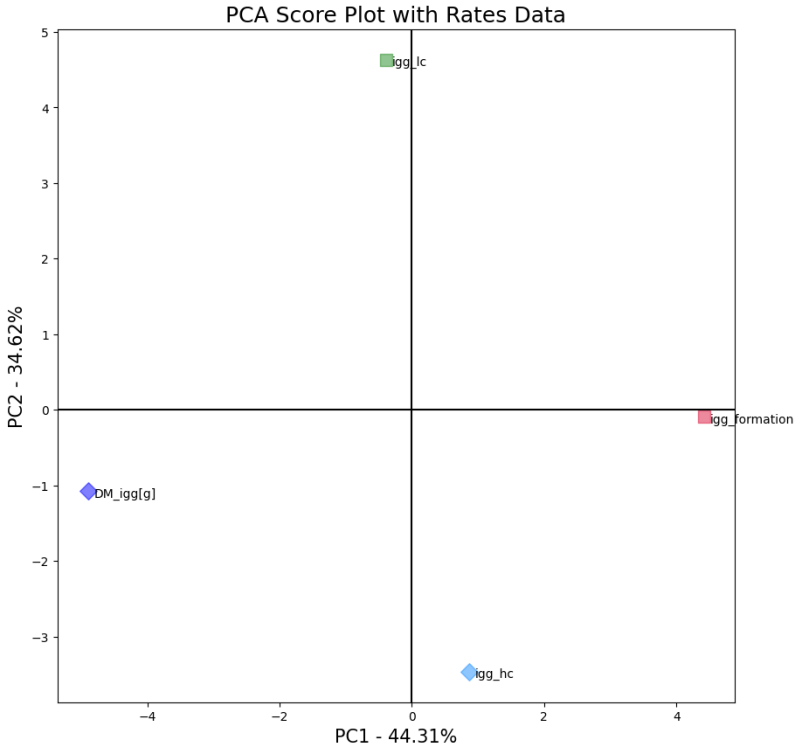

Figure 10: PCA score plot

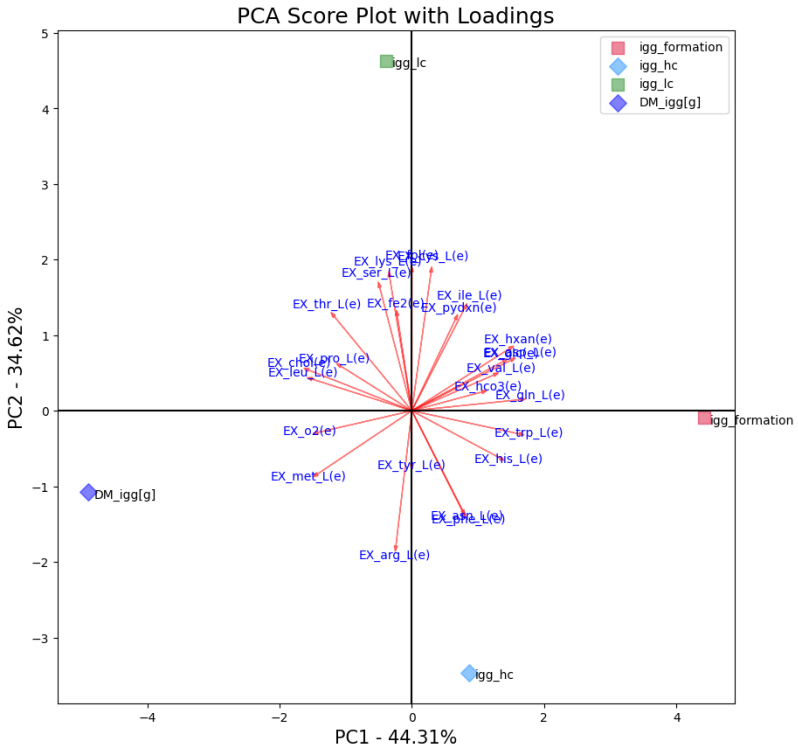

Figure 11: PCA biplot

All files and codes used for this project can be found here –

https://drive.google.com/drive/folders/16LGk73kSI_Er3-JKechOPIiSa_Rl42a6?usp=sharing

Contributions

Name	Roll No.	Contribution
Anirudh Rao	BE21B004	Conceptualization, reinforcement learning, constraint-based modelling, report writing
Anirudh TP	BS21B007	Conceptualization, reinforcement learning, constraint-based modelling, report writing
Pooja Shankar	BE21B027	Conceptualization, comparative media analysis, future directions, report writing
Pranathi R	BS21B023	Conceptualization, comparative media analysis, future directions, report writing
Sidharthan S C	BE21B039	Conceptualization, reinforcement learning, constraint-based modelling, report writing