TALK2BIOMODELS AND TALK2KNOWLEDGEGRAPHS: AI AGENT-BASED APPLICATION FOR PREDICTION OF PATIENT BIOMARKERS AND REASONING OVER BIOMEDICAL KNOWLEDGE GRAPHS

Gurdeep Singh,[*] Lilija Wehling,[*] Ahmad Wisnu Mulyadi,[*]
Rakesh Hadne Sreenath & Douglas McCloskey[†]
Virtual Patient Engine
BioMed X GmbH
Heidelberg, 69115, Germany
{gsingh, wehling, mulyadi, hadnesreenath, mccloskey}@bio.mx

Thomas Klabunde & Tommaso Andreani
Translational Medicine Unit (TMU)
Disease Modeling, Sanofi R&D
Frankfurt am Main, Germany
{Thomas.Klabunde, Tommaso.Andreani}@sanofi.com

ABSTRACT

In this study, we present Talk2Biomodels (T2B) and Talk2KnowledgeGraphs (T2KG) as open-source[1], user-friendly, large language model-based agentic AI platforms designed to democratize access to computational models of disease processes using natural language. T2B and T2KG eschew the traditional graphical user interface (GUI) and minimally adaptable workflow in favour of a modern agentic framework to provide a dynamic and immersive experience to explore the biology of disease in silico and how different treatment options can be efficacious in different virtual patient populations. T2B supports models encoded in the open-source community format Systems Biology Markup Language (SBML) for quantitative prediction of patient biomarkers and integrates with biomedical knowledge graphs to provide qualitative insights not captured in the computational model. A use case in precision medicine is presented to demonstrate how experts and non-experts in computational biology and data science can benefit from T2B and T2KG.

1 INTRODUCTION

Recent developments in generative artificial intelligence (AI), particularly in large language models (LLMs), have demonstrated substantial progress. Noteworthy advances in both algorithmic innovation and data-driven biomedical discovery are showcased by the advent of foundational learning models for protein structure, single cell omics and genomic data across species, such as, AlphaFold, scGPT, and Evo 2, to name a few (Brixi et al., 2025; Cui et al., 2024; Jumper et al., 2021). While LLMs' role is currently confined to narrow, task-specific applications, efforts are underway to enable open-source LLM-driven agentic workflows in biomedical discovery (Lobentanzer et al., 2025). Collaborative AI agents, with their task orchestration, reflective learning and sophisticated reasoning capabilities, have the potential to revolutionize biomedical research and patient treatment (Gao et al., 2024; Zou & Topol, 2025).

[*]These authors contributed equally to this work.
[†]Corresponding author.
[1]Source code available at: https://github.com/VirtualPatientEngine/AIAgents4Pharma

Computational models and biomedical knowledge graphs have become an essential component in enhancing our understanding of complex disease processes. The most common computational modeling frameworks in systems biology modeling include kinetic, constraint-based, logic, and agent-based modeling. The present study specifically focuses on ordinary differential equations (ODE) models, a subset of kinetic models, due to their widespread application and significance in the field. ODE models represent the most popular modeling approach, with approximately 1,662 models[2] currently deposited in the BioModels database (Malik-Sheriff et al., 2019) and a multitude of models not systematically stored in the scientific literature. These models cover a wide range of domains, as demonstrated by their use in precision medicine (e.g., inflammatory bowel disease treatment) (Dwivedi et al., 2014), COVID-19 pandemic studies (Tang et al., 2020), and theoretical analysis of dynamic systems (Markevich et al., 2004). Furthermore, ODEs are a standard tool in pharmaceutical research, particularly in the development of physiologically based pharmacokinetic/pharmacodynamic (PBPK/PD) and quantitative systems pharmacology and toxicology (QSP/QST) models (Chan et al., 2024; Kuepfer et al., 2016). In addition, the regulatory authorities, such as the U.S. Food and Drug Administration (FDA), recognize these computational models as a well-established and increasingly significant body of evidence in drug approval decision-making processes (Morrison et al., 2018). While computational models provide quantitative predictions within a limited biological scope, biomedical knowledge graphs offer qualitative reasoning across a broader biological spectrum. Biomedical knowledge graphs also play a critical role in pharmaceutical research (Bonner et al., 2022) where they are used to identify new drug targets (Middleton et al., 2024), repurpose existing drugs (Huang et al., 2024), explore the effects of drug-drug interactions (Zitnik et al., 2018), and refine patient endotypes that would most benefit from an existing or new therapy (Geleta et al., 2021).

To interact with mathematical models of biological systems, several popular graphical user interface (GUI)-based ODE modelling frameworks exist for biological and translational applications. These include proprietary options such as the MATLAB toolbox SimBiology (The MathWorks Inc, 2024) and Berkley Madonna (Marcoline et al., 2022), as well as open-source alternatives like COPASI (Hoops et al., 2006) and the Open Systems Pharmacology Suite (Kuepfer et al., 2016). However, both proprietary and open-source solutions typically require software installation, specialized training, and domain knowledge for effective simulation and analysis of models.

To overcome the limitations of existing modelling frameworks, and to expand the biological scope of computational models, we propose Talk2BioModels (T2B) and Talk2KnowledgeGraphs (T2KG) as easily accessible agentic AI applications. T2B and T2KG users do not need to possess expert knowledge in programming nor systems biology to simulate and perform basic analyses of ODE models, and qualitatively reason over and contextualize the results in the broader understanding of biology and disease captured in biomedical knowledge graphs. The platform supports the import and export of open and closed-source models encoded in SBML (Keating et al., 2020), and connectors to open and closed-source biomedical knowledge graphs such as the publicly available PrimeKG dataset (Chandak et al., 2023).

By using an agentic AI framework, T2B and T2KG democratize access to expert knowledge captured in computational models of disease and biomedical knowledge graphs. T2B and T2KG promote accurate and explainable science by grounding LLM responses in computational modeling simulation results or in extracted biomedical knowledge graph subsets. Graph Retrieval-Augmented Generation (GraphRAG) (Lewis et al., 2020; He et al., 2024; Edge et al., 2025) provides the necessary explainability for deployment in regulated industries, including healthcare and pharmaceutical industries.

2 METHODS

We developed a multi-agent hierarchical system using the LangGraph library (v0.3.2)[3], which provides a framework for constructing stateful computational graphs. In this architecture, each node represents a discrete computational step, while the edges define the relationships between these steps. Each agent within the system maintains a state, a snapshot of its application, which is propagated and updated throughout the execution process (Fig. 1).

[2] Viewed on 24.02.2025

[3] https://github.com/langchain-ai/langgraph

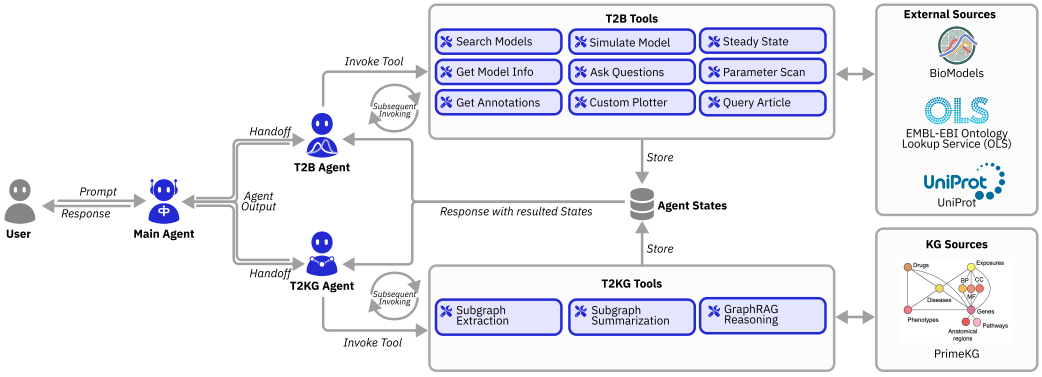

Figure 1: Architecture diagram of Talk2Biomodels and Talk2KnowledgeGraphs orchestrated by the main agent. Users can interact with T2B and T2KG agents via the main agent. The main agent acts as a supervisor agent and determines which agent (and its respective tools) to invoke based on the user input. The T2B agent specializes in quantitative analysis, including model search, simulation, plotting, parameter scanning, steady-state stability analysis, annotation extraction, and RAG for uploaded scientific articles. Meanwhile, the T2KG agent focuses on providing quantitative insights, such as subgraph extraction, summarization, and GraphRAG reasoning on the knowledge graph. These agentic AI tools interactively engage with resources, including BioModels, OLS, and UniProt databases as external databases and PrimeKG as a knowledge graph resource. Each agent maintains its state independently throughout the chat history.

For both T2B and T2KG agents, the computational nodes consist of a prebuilt reasoning and action (ReAct)-based agent, implemented following the ReAct paradigm (Yao et al., 2023), and a suite of tools available (Fig. 1). A central main agent developed using the LangGraph-Supervisor library (v0.0.4)[4], manages task delegation and communication, dynamically invoking T2B and/or T2KG based on task context. Directed edges within the graph facilitate structured interactions between agents and their tools. Agents autonomously select and execute tool invocations in response to user input, updating their internal state accordingly. The system tracks key-value pairs storing messages, analysis outputs, and intermediate results, ensuring consistency across computational steps. Upon completion, the final state is parsed to extract relevant outputs for rendering.

T2B executes operations on models using the BASICO library (v0.78)(Bergmann, 2023). T2KG extracts subgraphs from an initial biomedical knowledge graph based on user queries, employing the Prize Collecting Steiner Tree (PCST) algorithm (Hegde et al., 2015) to guarantee a connected subgraph. The retrieved subgraph in the form of a textualized subgraph (i.e., textual information of the selected nodes and triples) is further incorporated for RAG (He et al., 2024) to enhance response quality. We used PrimeKG (Chandak et al., 2023) as the knowledge graph dataset, enriching its nodes and edges with a set of textual information from Stark (Wu et al., 2024). Textual embeddings were generated using the nomic-embed-text model (Nussbaum et al., 2025) retrieved from Ollama[5]. The processed graph was stored using NetworkX[6] (Hagberg et al., 2008) and efficiently manipulated with the PyTorch Geometric[7] library (Fey & Lenssen, 2019).

For configuring agent-based text generation tasks, users can select between two LLMs: OpenAI's GPT-4o-mini (OpenAI et al., 2024) and NVIDIA NIM-optimized Meta/Llama-3.3-70B-Instruct [8]. For text-to-embedding transformations, the application supports two embedding models: OpenAI's text-embedding-ada-002 and NVIDIA NIM-optimized Llama-3.2-NV-EmbedQA-1B-V2[9]. The backend of the system is implemented in Python 3.12, while the front end was developed using

[4]https://github.com/langchain-ai/langgraph-supervisor-py
[5]https://ollama.com/
[6]https://networkx.org/
[7]https://pytorch-geometric.readthedocs.io/
[8]https://build.nvidia.com/meta/llama-3_3-70b-instruct/modelcard
[9]https://build.nvidia.com/nvidia/llama-3_2-nv-embedqa-1b-v2

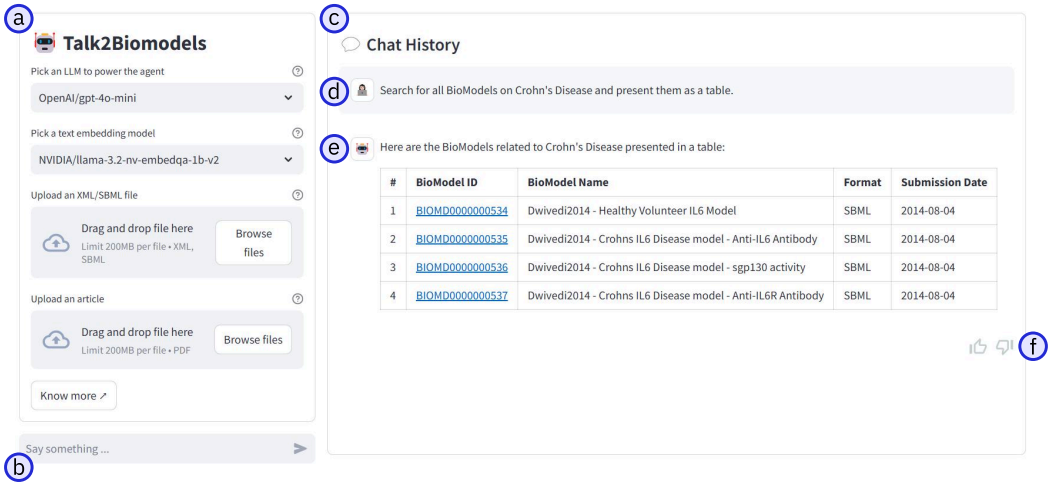

Figure 2: An example of the search models tool in the Talk2BioModels GUI. This GUI supports both open and closed-source LLMs, enabling users to upload SBML-format mathematical models and PDF-format scientific articles (a). T2B allows natural language prompts via the chat box interface (b). The GUI features a chat history window (c) displaying question-and-answer conversations (d, e). Additionally, T2B's GUI allows dynamic user feedback through upvote and downvote buttons. (f). We exemplify search models tool for keyword-based searches in the BioModels database (e).

Streamlit (v1.39.0)[10], Plotly (v5.24.1)[11], and Gravis (v0.1.0)[12] for an interactive and user-friendly interface.

The T2B, T2KG, and their unified version (Talk2AIAgents4Pharma) applications are open-source and available on GitHub[13], supporting community contributions and discussions.

3 Results and Discussion

T2B offers a comprehensive suite of functions for biological and translational ODE models. Users can upload ODE models in SBML format or retrieve curated models from the BioModels database using the *search models* tool (Fig. 2). The *get model information* tool extracts model details, including species, parameters, and model abstract. Concurrently, the *get annotations* tool retrieves and displays model species' identifiers by dynamically accessing Ontology Lookup Service (OLS) and UniProt data bases (Jupp et al., 2015; The UniProt Consortium, 2024) over APIs. This capability enables researchers to thoroughly explore and interpret the specific nomenclature of the model.

In T2B, SBML models can be directly accessed, simulated, and analyzed. The platform supports time-course simulations with adjustable time steps, initial concentrations and durations (*simulate model* tool), presenting results in dynamic tables and graphs that can be interactively modified to suit the users' needs (*custom plotter* tool). Additionally, users can interrogate model description and simulation results (*ask questions* tool) and alter model parameters using natural language.

Moreover, T2B offers several tools for model analysis: the *steady-state* tool determines whether a model reaches equilibrium, while the *parameter scan* tool performs multiple simulations by systematically varying model initial conditions or parameter values within user-specified range and intervals (Fig. A1). Finally, the *query article* tool leverages the RAG framework (Lewis et al., 2020) to extract supplementary information from user-uploaded research articles, enhancing model understanding.

[10]https://streamlit.io/
[11]https://plotly.com/
[12]https://robert-haas.github.io/gravis-docs/
[13]https://github.com/VirtualPatientEngine/AIAgents4Pharma

To promote greater explainability and contextualization of quantitative computational biology simulation predictions, users can invoke the T2KG to qualitatively reason over T2B results using a broader scope of disease biology captured in biomedical knowledge graphs. Specifically, users can extract out one or more subgraphs using the *subgraph extraction* tool based on the textual embedding similarity between the user's query and the embedded textual features of the biomedical knowledge graph. The textualized and summarized subgraph is then used for RAG in conjunction with previous computational biology model simulation results using the *graph RAG reasoning* tool. The results are presented to the user in natural language with citations to the entities and relations in the biomedical knowledge graph along with a dynamic widget of the extracted subgraph for interactive visual inspection.

Overall, this suite of agentic AI-driven tools offers a unique platform for working with biological models of disease, combining user-friendliness and interactive features.

3.1 USE-CASE IN PRECISION MEDICINE

The use cases in precision medicine[14] utilizes a multiscale model of interleukin 6 (IL-6) mediated immune regulation in Crohn's disease to examine drug-disease interaction mechanisms and drug treatment prediction (Dwivedi et al., 2014). The model was validated using clinical trial biomarker data from patients treated with tocilizumab, an anti IL6 receptor monoclonal antibody drug.

Users can interrogate model annotations, adjust the dosage of a drug, simulate drug administration at intervals of either 2 or 4 weeks, and view the value of the clinical readout parameter (C-reactive protein (CRP) in serum) at any given time point. Users can determine the time required for the depletion of serum anti-IL6Ralpha concentration by replicating the published figure (Dwivedi et al., 2014). The effect on CRP suppression can also be compared by simulating different antibodies with varying dissociation affinities, thereby performing in silico tests for potential antibodies. Additionally, users can explore alternative drugs that target IL-6 or even alternative targets of Crohn's diseases as a basis for follow-up studies. Users can qualitatively reason over the simulated results by developing a rational grounded in curated biomedical knowledge of genes and pathways associated with Crohn's disease for why certain treatments work and others do not. A snapshot of a use case on extracting subgraph that captures the IL6 node along with its interconnections to other relevant nodes (e.g., gene/protein, disease, pathway) within subsets of PrimeKG is presented in Fig. A2.

4 CONCLUSIONS

T2B and T2KG are open-source, user-friendly agentic AI platform designed to democratize access to computational models and biomedical knowledge graphs of disease. By integrating a natural language interface with agentic AI, these platforms lower the barrier for non-experts to leverage expert knowledge captured in computational biology models and biomedical knowledge graphs. A use case in precision medicine was presented to demonstrate the ease which the users can reproduce and explore questions of clinical translatable of drug assets.

As the user base of these platforms continues to grow and the complexity of agentic tools and workflows increases, ensuring transparency and validation of results becomes increasingly critical. Minimizing hallucinations through model simulation-specific tool calling is essential and must be rigorously tested using comprehensive testing suites. Further, we aim to extend T2KG's current functionality to support multi-modal embeddings for small molecules, proteins, and other biological entities, enhancing its utility in biomedical applications. Additionally, as new LLMs are developed and incorporated into the framework, it is important to continuously assess and validate their reasoning capabilities.

We would greatly appreciate contributions to our open-source project from the community.

ACKNOWLEDGMENTS

We acknowledge the contributions of Anil Kumar Kanasani and Maryam Najafian during the "AI Agents for Life Sciences" hackathon on October 14-15 2024. We acknowledge Rahuman S. Malik

[14] https://virtualpatientengine.github.io/AIAgents4Pharma/talk2biomodels/cases/Case_1/

Sheriff and the rest of the BioModels team for support in and discussions when developing T2B. We acknowledge Mohammed H. Mosa, Henrik Cordes, and Thomas Klabunde from Sanofi for testing T2B. We acknowledge the technical support from Dror Hilman, Jake Valsamis, Conor Mohan, and Ben Evan Tsur, gracious donations of computing resources during our hackathon from Shahar Frunkin, and Daniel Koster, and an going support and facilitation from Simon Adar, all from CodeOcean[15]. We acknowledge funding support for our Hackathons from the AI Health Innovation Cluster[16] and particular the help received from Daniela Beyer and the rest of the AIH Coordination Team. We thank Ornella Kossi, Ann-Kristin Mueller, and Stefanie Schimmel from BioLabs Heidelberg[17], for hosting our hackathons. We thank Isabel Wilkinson, Uwe Samer, Nick Venanzi, and David Ruau from NVIDIA for assisting with our integration of NVIDIA NIMs for chat and embedding inference and graciously providing free credits for use with T2B. We thank Vultr[18] for graciously providing free credits for deploying T2B.

Author Contributions

LW, GS, AWM, and DM conceptualized the study. GS developed the agentic AI framework and the agentic AI tools for T2B. AWM contributed to the development of the agentic AI framework, agentic AI tools for T2KG, and created the figures and icons. RHS implemented get annotations tool. GS and AWM developed the user interface. DM and LW conceptualized the precision medicine use case. LW wrote the manuscript. DM supervised the writing of the manuscript and the development of the agentic AI framework. All authors contributed to testing T2B and T2KG and reviewing the manuscript.

Funding

This study was funded by Sanofi.

Conflic interests

TK and TA are Sanofi employees and may hold shares and/or stock options in the company.

References

Frank T. Bergmann. Basico: A simplified python interface to COPASI. *Journal of Open Source Software*, 8(90):5553, 2023. doi: 10.21105/joss.05553.

Stephen Bonner, Ian P Barrett, Cheng Ye, Rowan Swiers, Ola Engkvist, Andreas Bender, Charles Tapley Hoyt, and William L Hamilton. A review of biomedical datasets relating to drug discovery: a knowledge graph perspective. *Briefings in Bioinformatics*, 23(6):bbac404, 2022.

Garyk Brixi, Matthew G. Durrant, Jerome Ku, Michael Poli, Greg Brockman, Daniel Chang, Gabriel A. Gonzalez, Samuel H. King, David B. Li, Aditi T. Merchant, Mohsen Naghipourfar, Eric Nguyen, Chiara Ricci-Tam, David W. Romero, Gwanggyu Sun, Ali Taghibakshi, Anton Vorontsov, Brandon Yang, Myra Deng, Liv Gorton, Nam Nguyen, Nicholas K. Wang, Etowah Adams, Stephen A. Baccus, Steven Dillmann, Stefano Ermon, Daniel Guo, Rajesh Ilango, Ken Janik, Amy X. Lu, Reshma Mehta, Mohammad R.K. Mofrad, Madelena Y. Ng, Jaspreet Pannu, Christopher Ré, Jonathan C. Schmok, John St. John, Jeremy Sullivan, Kevin Zhu, Greg Zynda, Daniel Balsam, Patrick Collison, Anthony B. Costa, Tina Hernandez-Boussard, Eric Ho, Ming-Yu Liu, Thomas McGrath, Kimberly Powell, Dave P. Burke, Hani Goodarzi, Patrick D. Hsu, and Brian L. Hie. Genome modeling and design across all domains of life with Evo 2. *bioRxiv*, 2025. doi: 10.1101/2025.02.18.638918.

Jason R. Chan, Richard Allen, Britton Boras, Antonio Cabal, Valeriu Damian, Francis D. Gibbons, Abhishek Gulati, Iraj Hosseini, Jeffrey D. Kearns, Ryuta Saito, Lourdes Cucurull-Sanchez, Jangir Selimkhanov, Andrew M. Stein, Kenichi Umehara, Guanyu Wang, Weirong Wang, and

[15] https://codeocean.com/

[16] https://www.aih-cluster.ai/

[17] https://www.biolabs.io/heidelberg

[18] https://www.vultr.com/

Susana Neves-Zaph. Current practices for QSP model assessment: an IQ consortium survey. *Journal of Pharmacokinetics and Pharmacodynamics*, 51(5):543–555, Oct 2024. doi: 10.1007/s10928-022-09811-1.

Payal Chandak, Kexin Huang, and Marinka Zitnik. Building a knowledge graph to enable precision medicine. *Scientific Data*, 10(1):67, Feb 2023. doi: 10.1038/s41597-023-01960-3.

Haotian Cui, Chloe Wang, Hassaan Maan, Kuan Pang, Fengning Luo, Nan Duan, and Bo Wang. scGPT: Toward building a foundation model for single-cell multi-omics using generative AI. *Nature Methods*, 21(8):1470–1480, Aug 2024. doi: 10.1038/s41592-024-02201-0.

G Dwivedi, L Fitz, M Hegen, SW Martin, J Harrold, A Heatherington, and C Li. A multiscale model of interleukin-6–mediated immune regulation in crohn's disease and its application in drug discovery and development. *CPT: Pharmacometrics & Systems Pharmacology*, 3(1):89, 2014. doi: https://doi.org/10.1038/psp.2013.64.

Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, Dasha Metropolitansky, Robert Osazuwa Ness, and Jonathan Larson. From local to global: A graph RAG approach to query-focused summarization, 2025. URL https://arxiv.org/abs/2404.16130.

Matthias Fey and Jan Eric Lenssen. Fast graph representation learning with pytorch geometric, 2019. URL https://arxiv.org/abs/1903.02428.

Shanghua Gao, Ada Fang, Yepeng Huang, Valentina Giunchiglia, Ayush Noori, Jonathan Richard Schwarz, Yasha Ektefaie, Jovana Kondic, and Marinka Zitnik. Empowering biomedical discovery with AI agents. *Cell*, 187(22):6125–6151, Oct 2024. doi: 10.1016/j.cell.2024.09.022.

David Geleta, Andriy Nikolov, Gavin Edwards, Anna Gogleva, Richard Jackson, Erik Jansson, Andrej Lamov, Sebastian Nilsson, Marina Pettersson, Vladimir Poroshin, Benedek Rozemberczki, Timothy Scrivener, Michael Ughetto, and Eliseo Papa. Biological insights knowledge graph: an integrated knowledge graph to support drug development. *bioRxiv*, 2021. doi: 10.1101/2021.10.28.466262.

Aric Hagberg, Pieter J. Swart, and Daniel A. Schult. Exploring network structure, dynamics, and function using networkx. Los Alamos National Laboratory (LANL), Los Alamos, NM (United States), 01 2008. URL https://www.osti.gov/biblio/960616.

Xiaoxin He, Yijun Tian, Yifei Sun, Nitesh Chawla, Thomas Laurent, Yann LeCun, Xavier Bresson, and Bryan Hooi. G-retriever: Retrieval-augmented generation for textual graph understanding and question answering. In *Advances in Neural Information Processing Systems*, volume 37, pp. 132876–132907, 2024.

Chinmay Hegde, Piotr Indyk, and Ludwig Schmidt. A nearly-linear time framework for graph-structured sparsity. In *International Conference on Machine Learning*, pp. 928–937. PMLR, 2015.

Stefan Hoops, Sven Sahle, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. COPASI—a COmplex PAthway SImulator. *Bioinformatics*, 22(24):3067–3074, 10 2006. doi: 10.1093/bioinformatics/btl485.

Kexin Huang, Payal Chandak, Qianwen Wang, Shreyas Havaldar, Akhil Vaid, Jure Leskovec, Girish N. Nadkarni, Benjamin S. Glicksberg, Nils Gehlenborg, and Marinka Zitnik. A foundation model for clinician-centered drug repurposing. *Nature Medicine*, 30(12):3601–3613, Dec 2024. doi: 10.1038/s41591-024-03233-x.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, Aug 2021. doi: 10.1038/s41586-021-03819-2.

Simon Jupp, Tony Burdett, Catherine Leroy, and Helen E Parkinson. A new Ontology Lookup Service at EMBL-EBI. *SWAT4LS*, 2:118–119, 2015.

Sarah M Keating, Dagmar Waltemath, Matthias König, Fengkai Zhang, Andreas Dräger, Claudine Chaouiya, Frank T Bergmann, Andrew Finney, Colin S Gillespie, Tomáš Helikar, Stefan Hoops, Rahuman S Malik-Sheriff, Stuart L Moodie, Ion I Moraru, Chris J Myers, Aurélien Naldi, Brett G Olivier, Sven Sahle, James C Schaff, Lucian P Smith, Maciej J Swat, Denis Thieffry, Leandro Watanabe, Darren J Wilkinson, Michael L Blinov, Kimberly Begley, James R Faeder, Harold F Gómez, Thomas M Hamm, Yuichiro Inagaki, Wolfram Liebermeister, Allyson L Lister, Daniel Lucio, Eric Mjolsness, Carole J Proctor, Karthik Raman, Nicolas Rodriguez, Clifford A Shaffer, Bruce E Shapiro, Joerg Stelling, Neil Swainston, Naoki Tanimura, John Wagner, Martin Meier-Schellersheim, Herbert M Sauro, Bernhard Palsson, Hamid Bolouri, Hiroaki Kitano, Akira Funahashi, Henning Hermjakob, John C Doyle, Michael Hucka, null null, Richard R Adams, Nicholas A Allen, Bastian R Angermann, Marco Antoniotti, Gary D Bader, Jan Červený, Mélanie Courtot, Chris D Cox, Piero Dalle Pezze, Emek Demir, William S Denney, Harish Dharuri, Julien Dorier, Dirk Drasdo, Ali Ebrahim, Johannes Eichner, Johan Elf, Lukas Endler, Chris T Evelo, Christoph Flamm, Ronan MT Fleming, Martina Fröhlich, Mihai Glont, Emanuel Gonçalves, Martin Golebiewski, Hovakim Grabski, Alex Gutteridge, Damon Hachmeister, Leonard A Harris, Benjamin D Heavner, Ron Henkel, William S Hlavacek, Bin Hu, Daniel R Hyduke, Hidde de Jong, Nick Juty, Peter D Karp, Jonathan R Karr, Douglas B Kell, Roland Keller, Ilya Kiselev, Steffen Klamt, Edda Klipp, Christian Knüpfer, Fedor Kolpakov, Falko Krause, Martina Kutmon, Camille Laibe, Conor Lawless, Lu Li, Leslie M Loew, Rainer Machne, Yukiko Matsuoka, Pedro Mendes, Huaiyu Mi, Florian Mittag, Pedro T Monteiro, Kedar Nath Natarajan, Poul MF Nielsen, Tramy Nguyen, Alida Palmisano, Jean-Baptiste Pettit, Thomas Pfau, Robert D Phair, Tomas Radivoyevitch, Johann M Rohwer, Oliver A Ruebenacker, Julio Saez-Rodriguez, Martin Scharm, Henning Schmidt, Falk Schreiber, Michael Schubert, Roman Schulte, Stuart C Sealfon, Kieran Smallbone, Sylvain Soliman, Melanie I Stefan, Devin P Sullivan, Koichi Takahashi, Bas Teusink, David Tolnay, Ibrahim Vazirabad, Axel von Kamp, Ulrike Wittig, Clemens Wrzodek, Finja Wrzodek, Ioannis Xenarios, Anna Zhukova, and Jeremy Zucker. SBML Level 3: An extensible format for the exchange and reuse of biological models. *Molecular Systems Biology*, 16(8): e9110, 2020. doi: https://doi.org/10.15252/msb.20199110.

L Kuepfer, C Niederalt, T Wendl, J-F Schlender, S Willmann, J Lippert, M Block, T Eissing, and D Teutonico. Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model. *CPT: Pharmacometrics & Systems Pharmacology*, 5(10):516–531, 2016. doi: https://doi.org/10.1002/psp4.12134.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in neural information processing systems*, 33: 9459–9474, 2020.

Sebastian Lobentanzer, Shaohong Feng, Noah Bruderer, Andreas Maier, Adrián G. Díaz, Amy Strange, Anis Ismail, Anton Kulaga, Aurelien Dugourd, Barbara Zdrazil, Bastien Chassagnol, Cyril Pommier, Daniele Lucarelli, Ellen M. McDonagh, Emma Verkinderen, Fernando M. Delgado-Chaves, Georg Fuellen, Hannah Sonntag, Jonatan Menger, Lionel Christiaen, Ludwig Geistlinger, Luna Zacharias Zetsche, Marlis Engelke, Megan McNutt, Melissa Harrison, Melissa Hizli, Nikolai Usanov, Patrick Baracho, Sebastian Beier, Stefan Boeing, Taru A. Muranen, Trang T. Le, Valeriia Dragan, Xiao-Ran Zhou, Yasmin Nielsen-Tehranchian, Yuyao Song, Cankun Wang, Jan Baumbach, Jorge Abreu-Vicente, Nils Krehl, Qin Ma, Thomas Lemberger, Julio Saez-Rodriguez, and The BioChatter Consortium. A platform for the biomedical application of large language models. *Nature Biotechnology*, 43(2):166–169, Feb 2025. doi: 10.1038/s41587-024-02534-3.

Rahuman S Malik-Sheriff, Mihai Glont, Tung V N Nguyen, Krishna Tiwari, Matthew G Roberts, Ashley Xavier, Manh T Vu, Jinghao Men, Matthieu Maire, Sarubini Kananathan, Emma L Fairbanks, Johannes P Meyer, Chinmay Arankalle, Thawfeek M Varusai, Vincent Knight-Schrijver, Lu Li, Corina Dueñas-Roca, Gaurhari Dass, Sarah M Keating, Young M Park, Nicola Buso, Nicolas Rodriguez, Michael Hucka, and Henning Hermjakob. Biomodels—15 years of sharing computational models in life science. *Nucleic Acids Research*, 48(D1):D407–D415, 11 2019. doi: 10.1093/nar/gkz1055.

Frank V. Marcoline, John Furth, Smita Nayak, Michael Grabe, and Robert I. Macey. Berkeley Madonna Version 10–A simulation package for solving mathematical models. *CPT: Pharmacometrics & Systems Pharmacology*, 11(3):290–301, 2022. doi: https://doi.org/10.1002/psp4.12757.

Nick I. Markevich, Jan B. Hoek, and Boris N. Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *Journal of Cell Biology*, 164 (3):353–359, 01 2004. doi: 10.1083/jcb.200308060.

Lawrence Middleton, Ioannis Melas, Chirag Vasavda, Arwa Raies, Benedek Rozemberczki, Ryan S. Dhindsa, Justin S. Dhindsa, Blake Weido, Quanli Wang, Andrew R. Harper, Gavin Edwards, Slavé Petrovski, and Dimitrios Vitsios. Phenome-wide identification of therapeutic genetic targets, leveraging knowledge graphs, graph neural networks, and UK Biobank data. *Science Advances*, 10(19):eadj1424, 2024. doi: 10.1126/sciadv.adj1424.

Tina M. Morrison, Pras Pathmanathan, Mariam Adwan, and Edward Margerrison. Advancing Regulatory Science With Computational Modeling for Medical Devices at the FDA's Office of Science and Engineering Laboratories. *Frontiers in Medicine*, 5, 2018. doi: 10.3389/fmed.2018.00241.

Zach Nussbaum, John X. Morris, Brandon Duderstadt, and Andriy Mulyar. Nomic embed: Training a reproducible long context text embedder, 2025. URL `https://arxiv.org/abs/2402.01613`.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Jake Berdine, Gabriel Bernadett-Shapiro, Christopher Berner, Lenny Bogdonoff, Oleg Boiko, Madelaine Boyd, Anna-Luisa Brakman, Greg Brockman, Tim Brooks, Miles Brundage, Kevin Button, Trevor Cai, Rosie Campbell, Andrew Cann, Brittany Carey, Chelsea Carlson, Rory Carmichael, Brooke Chan, Che Chang, Fotis Chantzis, Derek Chen, Sully Chen, Ruby Chen, Jason Chen, Mark Chen, Ben Chess, Chester Cho, Casey Chu, Hyung Won Chung, Dave Cummings, Jeremiah Currier, Yunxing Dai, Cory Decareaux, Thomas Degry, Noah Deutsch, Damien Deville, Arka Dhar, David Dohan, Steve Dowling, Sheila Dunning, Adrien Ecoffet, Atty Eleti, Tyna Eloundou, David Farhi, Liam Fedus, Niko Felix, Simón Posada Fishman, Juston Forte, Isabella Fulford, Leo Gao, Elie Georges, Christian Gibson, Vik Goel, Tarun Gogineni, Gabriel Goh, Rapha Gontijo-Lopes, Jonathan Gordon, Morgan Grafstein, Scott Gray, Ryan Greene, Joshua Gross, Shixiang Shane Gu, Yufei Guo, Chris Hallacy, Jesse Han, Jeff Harris, Yuchen He, Mike Heaton, Johannes Heidecke, Chris Hesse, Alan Hickey, Wade Hickey, Peter Hoeschele, Brandon Houghton, Kenny Hsu, Shengli Hu, Xin Hu, Joost Huizinga, Shantanu Jain, Shawn Jain, Joanne Jang, Angela Jiang, Roger Jiang, Haozhun Jin, Denny Jin, Shino Jomoto, Billie Jonn, Heewoo Jun, Tomer Kaftan, Łukasz Kaiser, Ali Kamali, Ingmar Kanitscheider, Nitish Shirish Keskar, Tabarak Khan, Logan Kilpatrick, Jong Wook Kim, Christina Kim, Yongjik Kim, Jan Hendrik Kirchner, Jamie Kiros, Matt Knight, Daniel Kokotajlo, Łukasz Kondraciuk, Andrew Kondrich, Aris Konstantinidis, Kyle Kosic, Gretchen Krueger, Vishal Kuo, Michael Lampe, Ikai Lan, Teddy Lee, Jan Leike, Jade Leung, Daniel Levy, Chak Ming Li, Rachel Lim, Molly Lin, Stephanie Lin, Mateusz Litwin, Theresa Lopez, Ryan Lowe, Patricia Lue, Anna Makanju, Kim Malfacini, Sam Manning, Todor Markov, Yaniv Markovski, Bianca Martin, Katie Mayer, Andrew Mayne, Bob McGrew, Scott Mayer McKinney, Christine McLeavey, Paul McMillan, Jake McNeil, David Medina, Aalok Mehta, Jacob Menick, Luke Metz, Andrey Mishchenko, Pamela Mishkin, Vinnie Monaco, Evan Morikawa, Daniel Mossing, Tong Mu, Mira Murati, Oleg Murk, David Mély, Ashvin Nair, Reiichiro Nakano, Rajeev Nayak, Arvind Neelakantan, Richard Ngo, Hyeonwoo Noh, Long Ouyang, Cullen O'Keefe, Jakub Pachocki, Alex Paino, Joe Palermo, Ashley Pantuliano, Giambattista Parascandolo, Joel Parish, Emy Parparita, Alex Passos, Mikhail Pavlov, Andrew Peng, Adam Perelman, Filipe de Avila Belbute Peres, Michael Petrov, Henrique Ponde de Oliveira Pinto, Michael, Pokorny, Michelle Pokrass, Vitchyr H. Pong, Tolly Powell, Alethea Power, Boris Power, Elizabeth Proehl, Raul Puri, Alec Radford, Jack Rae, Aditya Ramesh, Cameron Raymond, Francis Real, Kendra Rimbach, Carl Ross, Bob Rotsted, Henri Roussez, Nick Ryder, Mario Saltarelli, Ted Sanders, Shibani Santurkar, Girish Sastry, Heather Schmidt, David Schnurr, John Schulman, Daniel Selsam, Kyla Sheppard, Toki Sherbakov, Jessica Shieh, Sarah Shoker, Pranav Shyam, Szymon Sidor,

Eric Sigler, Maddie Simens, Jordan Sitkin, Katarina Slama, Ian Sohl, Benjamin Sokolowsky, Yang Song, Natalie Staudacher, Felipe Petroski Such, Natalie Summers, Ilya Sutskever, Jie Tang, Nikolas Tezak, Madeleine B. Thompson, Phil Tillet, Amin Tootoonchian, Elizabeth Tseng, Preston Tuggle, Nick Turley, Jerry Tworek, Juan Felipe Cerón Uribe, Andrea Vallone, Arun Vijayvergiya, Chelsea Voss, Carroll Wainwright, Justin Jay Wang, Alvin Wang, Ben Wang, Jonathan Ward, Jason Wei, CJ Weinmann, Akila Welihinda, Peter Welinder, Jiayi Weng, Lilian Weng, Matt Wiethoff, Dave Willner, Clemens Winter, Samuel Wolrich, Hannah Wong, Lauren Workman, Sherwin Wu, Jeff Wu, Michael Wu, Kai Xiao, Tao Xu, Sarah Yoo, Kevin Yu, Qiming Yuan, Wojciech Zaremba, Rowan Zellers, Chong Zhang, Marvin Zhang, Shengjia Zhao, Tianhao Zheng, Juntang Zhuang, William Zhuk, and Barret Zoph. GPT-4 technical report, 2024. URL https://arxiv.org/abs/2303.08774.

Biao Tang, Xia Wang, Qian Li, Nicola Luigi Bragazzi, Sanyi Tang, Yanni Xiao, and Jianhong Wu. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. *Journal of Clinical Medicine*, 9(2), 2020. doi: 10.3390/jcm9020462.

The MathWorks Inc. SimBiology Toolbox version: 24.2 (R2024b) (Computer Software). 2024.

The UniProt Consortium. UniProt: the universal protein knowledgebase in 2025. *Nucleic Acids Research*, 53(D1):D609–D617, 11 2024. doi: 10.1093/nar/gkae1010.

Shirley Wu, Shiyu Zhao, Michihiro Yasunaga, Kexin Huang, Kaidi Cao, Qian Huang, Vassilis Ioannidis, Karthik Subbian, James Y Zou, and Jure Leskovec. Stark: Benchmarking LLM retrieval on textual and relational knowledge bases. *Advances in Neural Information Processing Systems*, 37: 127129–127153, 2024.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. ReAct: Synergizing Reasoning and Acting in Language Models, 2023. URL https://arxiv.org/abs/2210.03629.

Marinka Zitnik, Monica Agrawal, and Jure Leskovec. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*, 34(13):i457–i466, 2018.

James Zou and Eric J. Topol. The rise of agentic AI teammates in medicine. *The Lancet*, 405 (10477):457, Feb 2025. doi: 10.1016/S0140-6736(25)00202-8.

A APPENDIX

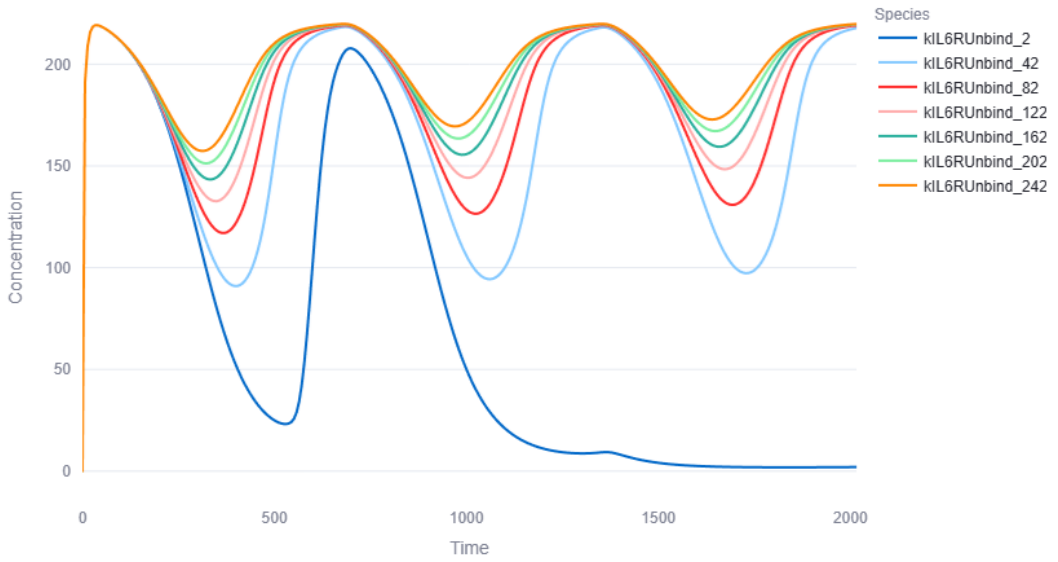

Figure A1: Parameter scan of a precision medicine model (Dwivedi et al., 2014) within the GUI. A parameter 'kIL6Unbind' is varied in steps of 40 units from 2 to 242 to simulate an unbinding event of an antibody from its target interleukin 6 receptor (IL6R). Each curve represents a separate simulation of this model.

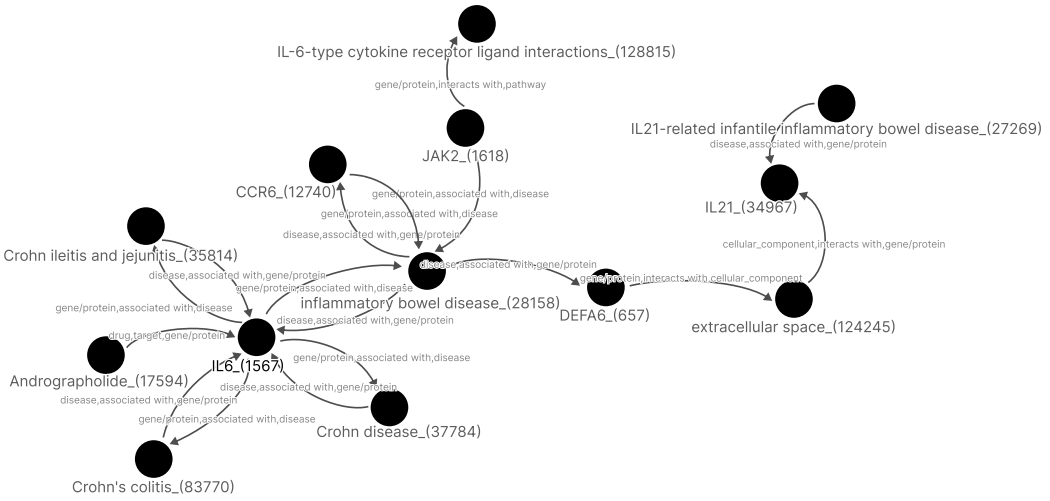

Figure A2: Extracted subgraph of IL-6 association with Crohn's Disease within the GUI. Nodes represent biological entities such as genes or pathways, and links represent relations between biological entities such as gene interaction with pathways or disease association with a gene. Note the disease association between IL-6 and Crohn's Disease shown on the bottom left and induced cytokine/chemokine network on the middle and right.