Machine learning enables engineering of potent, specific, and therapeutically developable proteases

Jung-Eun Shin
Seismic Therapeutic
Watertown, MA 02472
june.shin@seismictx.com

Nathan J. Rollins
Seismic Therapeutic
Watertown, MA 02472
nathan.rollins@seismictx.com

Purvi Mande
Seismic Therapeutic
Watertown, MA 02472
purvi.mande@seismictx.com

Jordan M. Anderson
Seismic Therapeutic
Watertown, MA 02472
jordan.anderson@seismictx.com

Allison Colthart
Seismic Therapeutic
Watertown, MA 02472
allison.colthart@seismictx.com

Soumya Bengeri
Seismic Therapeutic
Watertown, MA 02472
soumya.bengeri@seismictx.com

Emily Hoyt
Seismic Therapeutic
Watertown, MA 02472
emily.hoyt@seismictx.com

Alex Pellerin
Seismic Therapeutic
Watertown, MA 02472
alex.pellerin@seismictx.com

Ivan Mascanfroni
Seismic Therapeutic
Watertown, MA 02472
ivan.mascanfroni@seismictx.com

Jyothsna Visweswaraiah
Seismic Therapeutic
Watertown, MA 02472
jyothsna.visweswaraiah@seismictx.com

Yi Xing
Seismic Therapeutic
Watertown, MA 02472
yi.xing@seismictx.com

Kevin L. Otipoby
Seismic Therapeutic
Watertown, MA 02472
kevin.otipoby@seismictx.com

Nathan Higginson-Scott
Seismic Therapeutic
Watertown, MA 02472
nathan.higginson-scott@seismictx.com

Ryan Peckner
Seismic Therapeutic
Watertown, MA 02472
ryan.peckner@seismictx.com

Abstract

The development of biologic therapeutics is a labor-intensive, time-consuming, and expensive process. From the discovery of a protein with the desired function to optimizing the protein for manufacturability and removing immunogenicity, each step requires many iterations and large screening efforts. In order to accelerate this process, efforts have been increasingly focused on methods for protein engineering and design, both experimentally and computationally. Here we show that the

AI for New Drug Modalities at NeurIPS 2024.

careful application of machine learning methods in combination with structure-based and data-driven rational design allows for successful engineering of proteases with desired target specificity, low immunogenicity, and potent cleavage potential. We present two case studies describing the impact of machine learning in the development of two potential therapeutics. First, we apply machine learning models to optimize an IgG-degrading protease for drug-like properties while minimizing immunogenicity and other liabilities and maintaining therapeutic activity. Second, we leverage both machine learning and rational design methods to engineer an IgE-degrading protease with high specificity and activity. These results demonstrate the power of applying machine learning methods to accelerate the discovery and development of protein therapeutics.