



Oneness of Target and Modality Selection in Drug Discovery

How system thinking may improve R&D productivity in Pharma

Jitao David Zhang

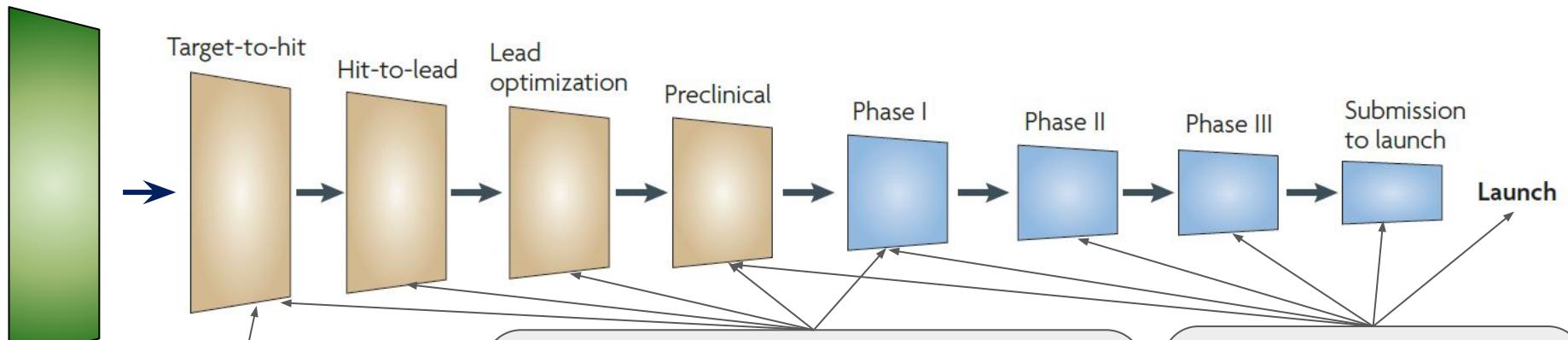
Pharma Research and Early Development (pRED)

Roche Innovation Center Basel

F. Hoffmann-La Roche Ltd

System thinking, interdisciplinary research, and quantitative science are critical to drug discovery

Target identification & assessment

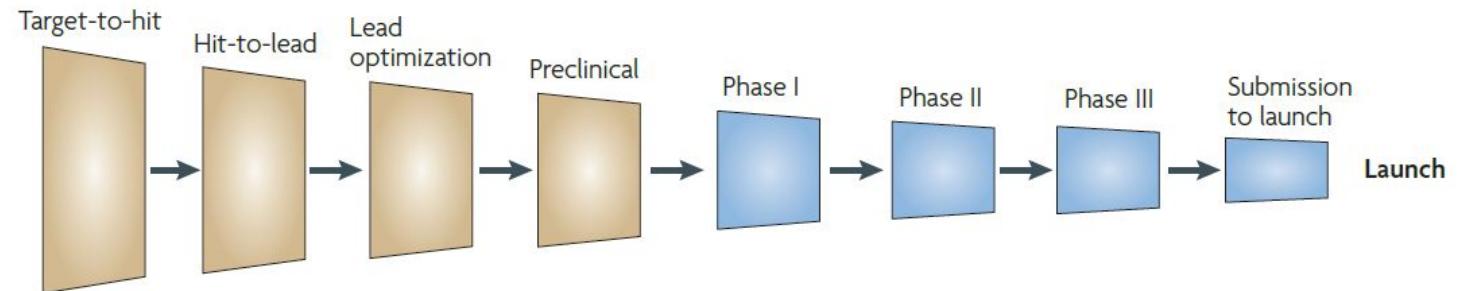


- Target identification, assessment, and validation
- Modality selection
- *Target Product Profile* design

- Compound and dose selection based on absorption, distribution, metabolism, excretion, and toxicity (ADMET), pharmacokinetics (PK), and pharmacodynamics (PD) profiles
- Mechanism of action (MoA) study
- Biomarker identification

- In vitro-in vivo and animal-human translation
- Mechanistic understanding of clinical findings
- Informing future projects with successes and failures

Cost of target selection and modality selection? Invisibly high

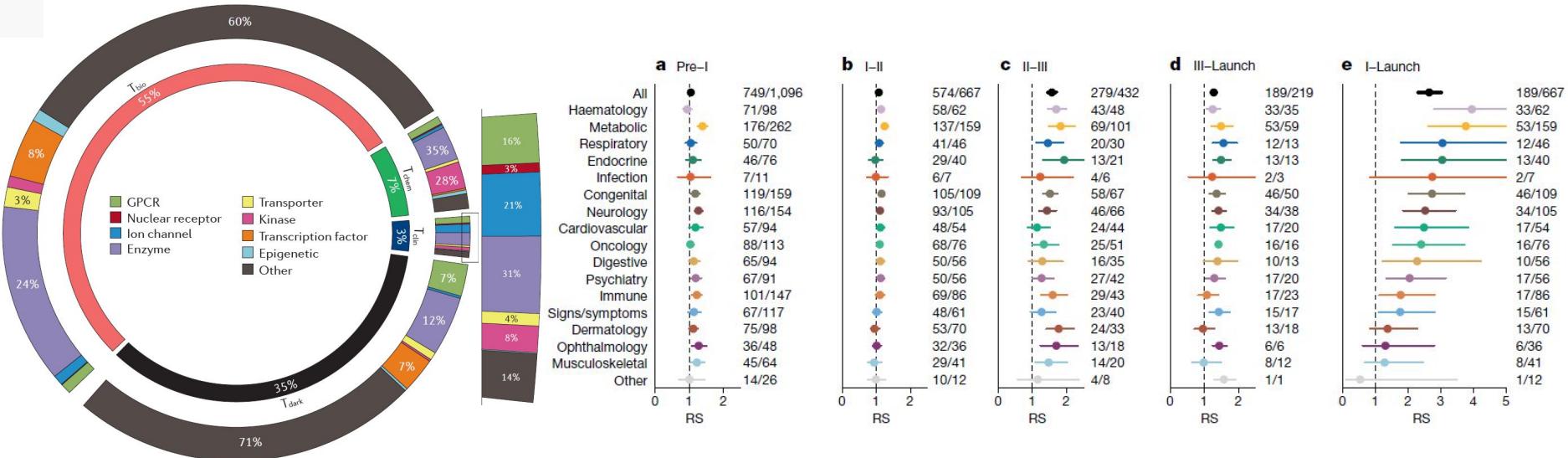


<i>p(TS)</i>	80%	75%	85%	69%	54%	34%	70%	91%	Total ≈ 5%
WIP needed for 1 launch	24.3	19.4	14.6	12.4	8.6	4.6	1.6	1.1	1
Cost per WIP per Phase	\$1	\$2.5	\$10	\$5	\$15	\$40	\$150	\$40	
Cycle time (years)	1.0	1.5	2.0	1.0	1.5	2.5	2.5	1.5	Total ≈ 12.5
Cost per launch (out of pocket)	\$24	\$49	\$146	\$62	\$128	\$185	\$235	\$44	\$873
% Total cost per NME	3%	6%	17%	7%	15%	21%	27%	5%	
Cost of capital	11%								
Cost per launch (capitalized)	\$94	\$166	\$414	\$150	\$273	\$319	\$314	\$48	\$1,778

\$ in the unit of Millions

■ Discovery ■ Development

Target identification: much unknown, lack of diversity, and support by genetic evidence is hardly discernible until Phase II/III

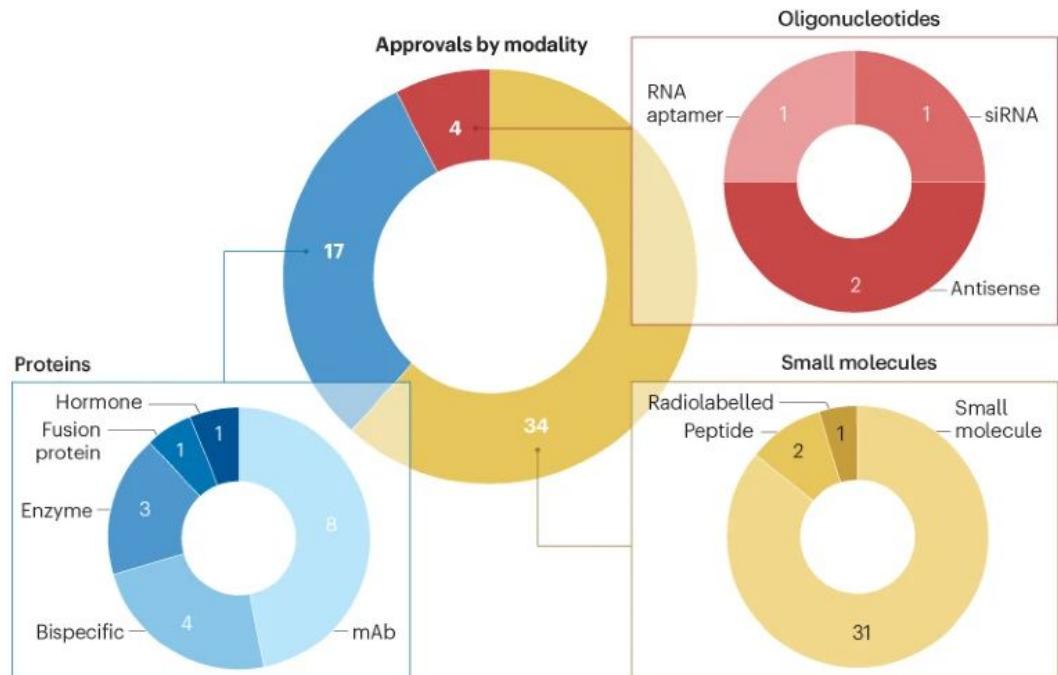


Left: T_{clin} proteins are linked to ≥ 1 approved drug. T_{chem} proteins bind to small molecules. T_{bio} have well-defined biological functions. GPCR: G-protein coupled receptors. Right: RS=Relative Success. Pre: preclinical development. Numbers: launched target-indication (T-I) pairs (numerator), and genetically supported T-I pairs (denominator).

References: Oprea, et al. 2018. [Unexplored Therapeutic Opportunities in the Human Genome](#). Nature Reviews Drug Discovery 17 (February):317–32.; Minikel, et al. 2024. [Refining the Impact of Genetic Evidence on Clinical Success](#). Nature 629 (8012): 624–29.

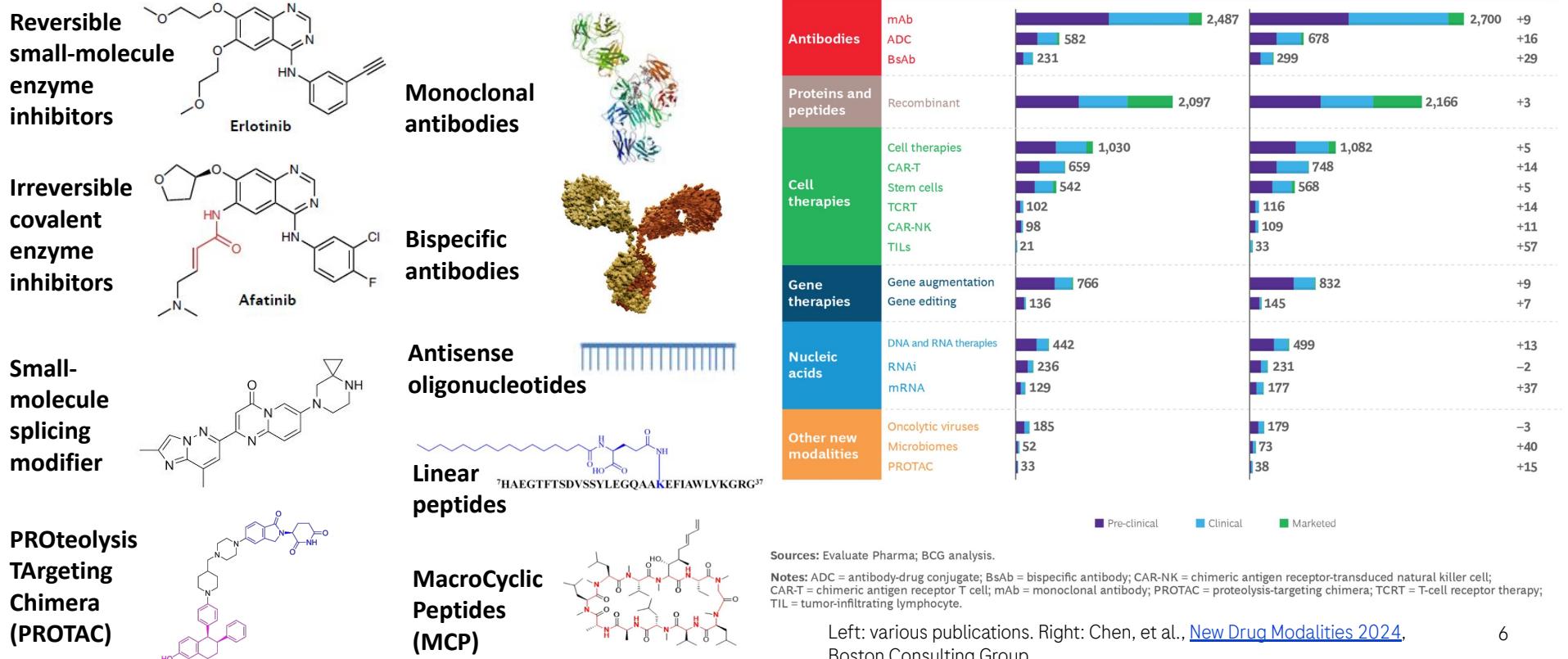
Novel drugs approved by the FDA's Center for Drug Evaluation and Research (CDER) in 2024

- Small molecules: molecular weight (MW) less than 1000 Daltons.
- Oligonucleotides: MW between 5 and 30 kDa (5000-30000 Da), negatively charged
 - siRNA: small interfering RNA
- Proteins: MW ~150 kDa
 - mAb: monoclonal antibody
 - Bispecific: antibodies that bind simultaneously to two antigens or two epitopes of the same antigen.

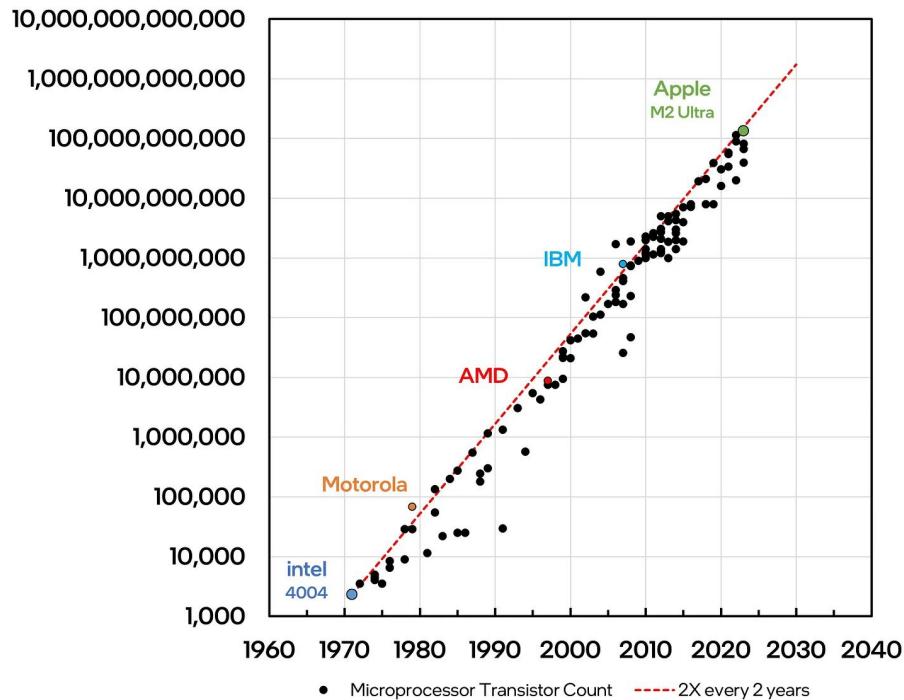


Source: [Asher Mullard, Nature Reviews Drug Discovery, 2024](#). The list can be found on [FDA's website](#)

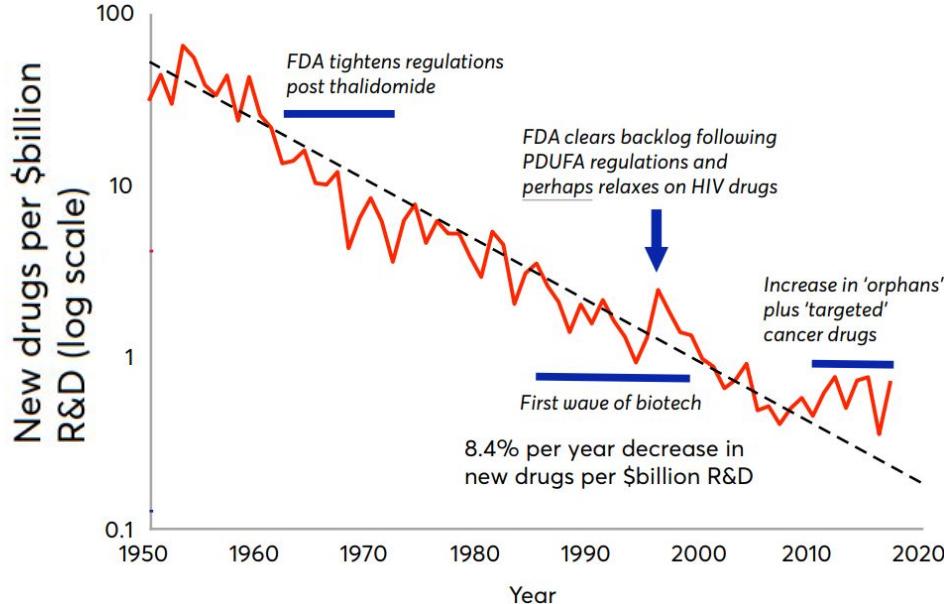
Modality selection: many choices, modality-specific challenges, and often implicit rationales



Moore's versus Eroom's Law

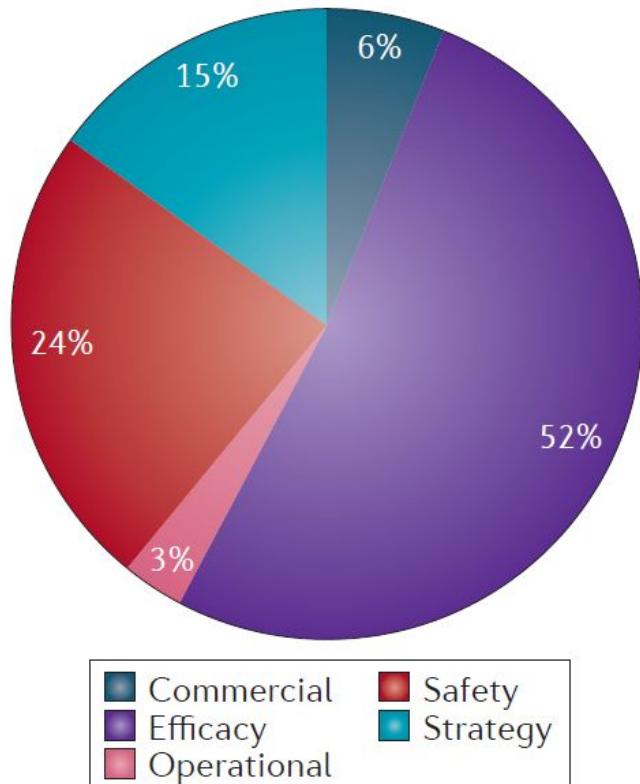


Adapted from graphs by [Hannah Ritchie and Max Roser](#) and by [Jim Keller](#)



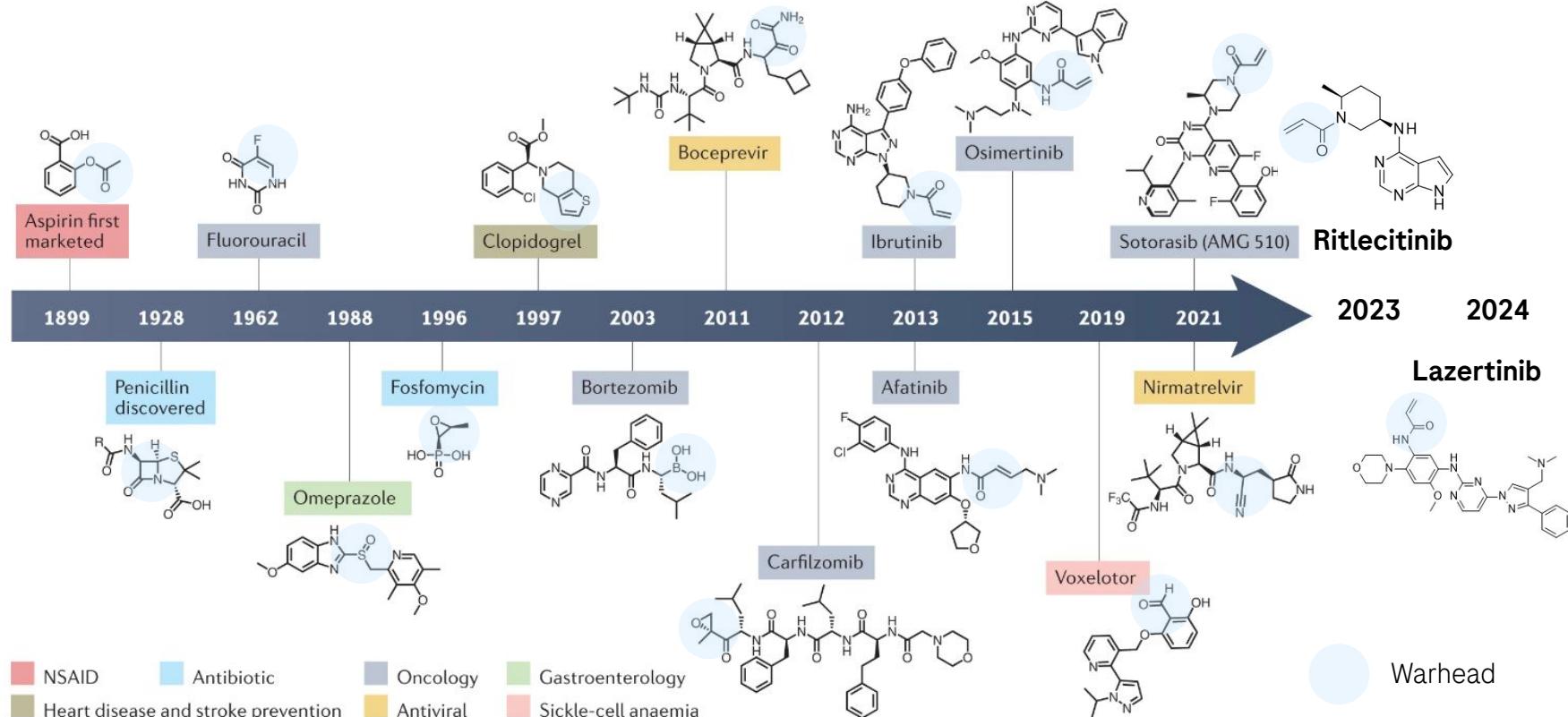
Data come from Scannell, et al. (2012) Diagnosing the decline in pharmaceutical R&D efficiency. Nature Reviews Drug Discovery, and personal communication. Figure by [Richard Jones and James Wilsdon](#)

Lack of efficacy and insufficient risk-benefit profiles are main reasons of failure in Phase II/III trials



We hypothesize that challenges in target and modality selection contribute to many efficacy/safety failures, and dwindling productivity overall.

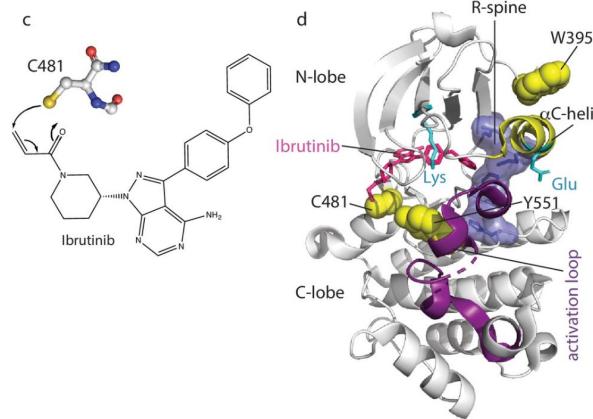
Covalent drugs have gained renewed interests



Ibrutinib, a first-in-class inhibitor of BTK (Bruton's Tyrosine Kinase)



Ibrutinib	
Approval	AbbVie (2013)
Binding type	Irreversible covalent binding
Binding site	C481, ATP-binding domain
Warhead	Acrylamide
Half-life	~4-6 hours
Indication	CLL, MCL, MZL, WM, GVHD, BN
Dosage	420 mg, qd (CLL/SLL, WM); 560 mg, qd (MCL, MZL)
Administration	Oral

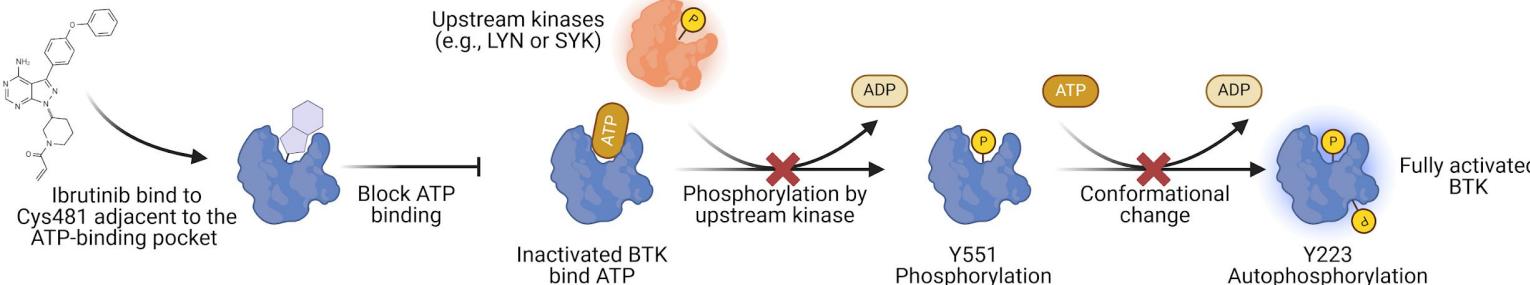


The ATP-binding pocket:

- A highly conserved region within the kinase domain.

Cysteine residue (Cys481):

- Located adjacent to the ATP-binding pocket
- Cys481 is a relatively unique cysteine, enabling selective covalent inhibition.



*Does it make sense to develop
covalent drugs for any target?
If not, which targets should be
covalently targeted?*

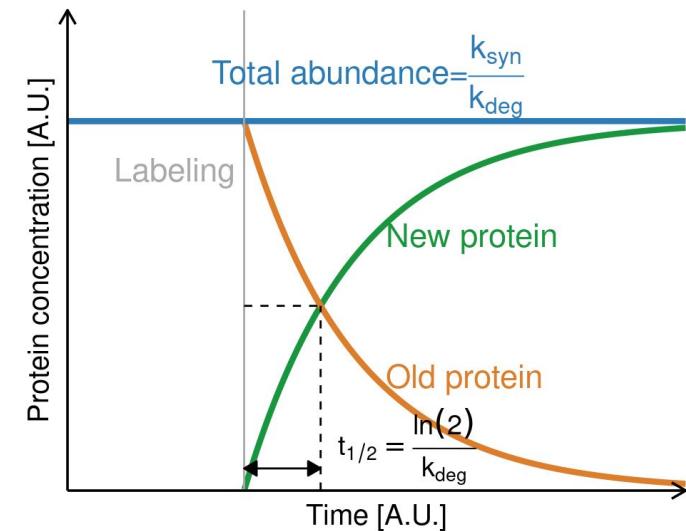
Importance of protein turnover

+

Turnover visualized, repurposing the *lilac tracer*
demonstrating Troxler's effect (Jeremy Hinton, [CC-BY 3.0](#))

Protein turnover is critical for drug discovery & development

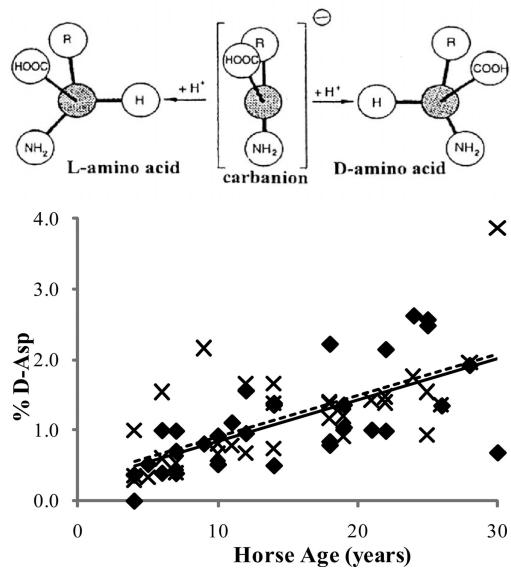
- Protein turnover affects efficacy, potency, ADME properties, and safety profiles of drug candidates.
- Protein turnover is essential for target prioritization and modality selection, for instance covalent binders and/or targeted protein degraders.
- Understanding protein turnover helps to translate pharmacokinetic and pharmacodynamic (PK/PD) relationships between systems.



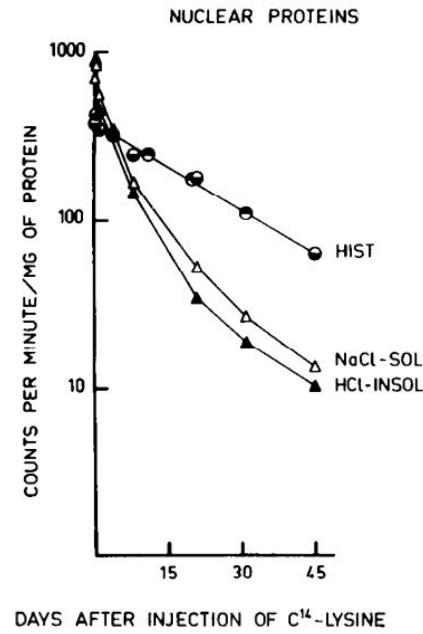
Assumptions: zero-order synthesis (rate k_{syn}), first-order degradation (rate k_{deg}), and steady state (i.e. no expression changes).

Quantifying protein turnover and long-living proteins

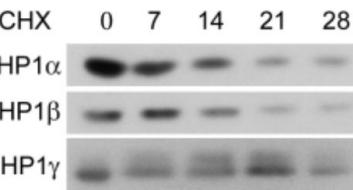
L- and D-Asp racemization



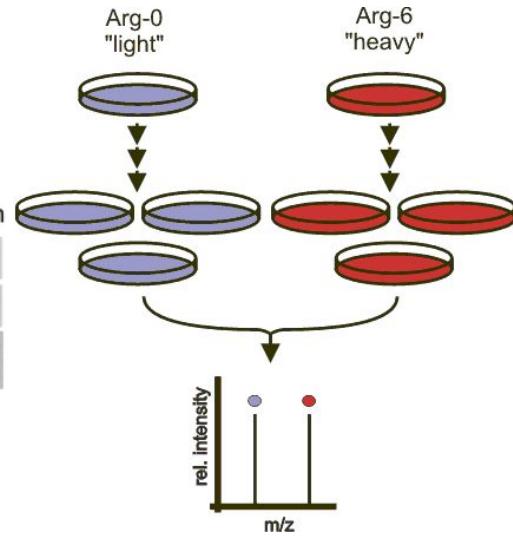
Radio isotope pulse-labelling



Western blotting following cycloheximide (CHX) treatment

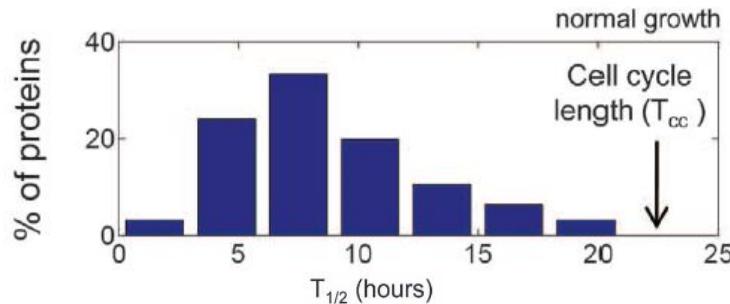


Stable isotope with amino acids in cell culture (SILAC) and mass spectrometry (MS)

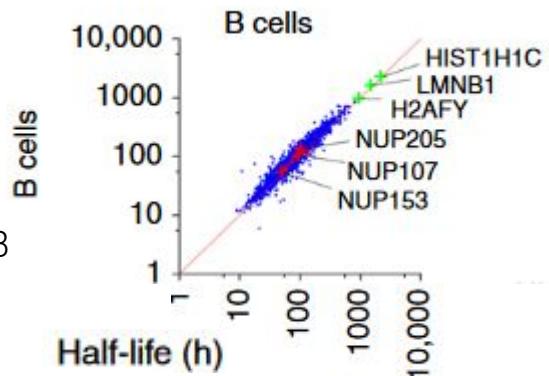


Source: various publications.

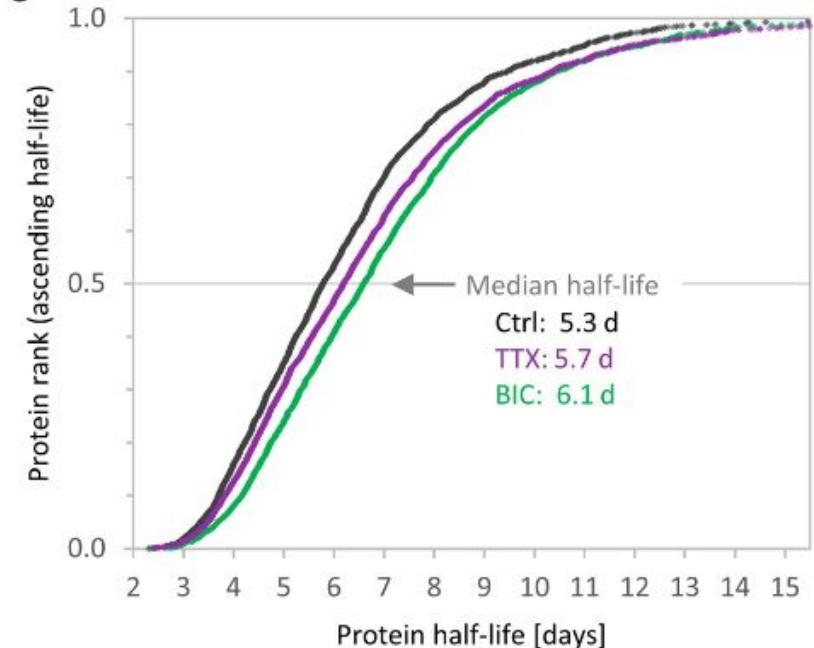
Protein half-life *in vitro* ranges between hours and days



Data from a human non-small cell lung cancer cell line
(Eden et al., 2011)

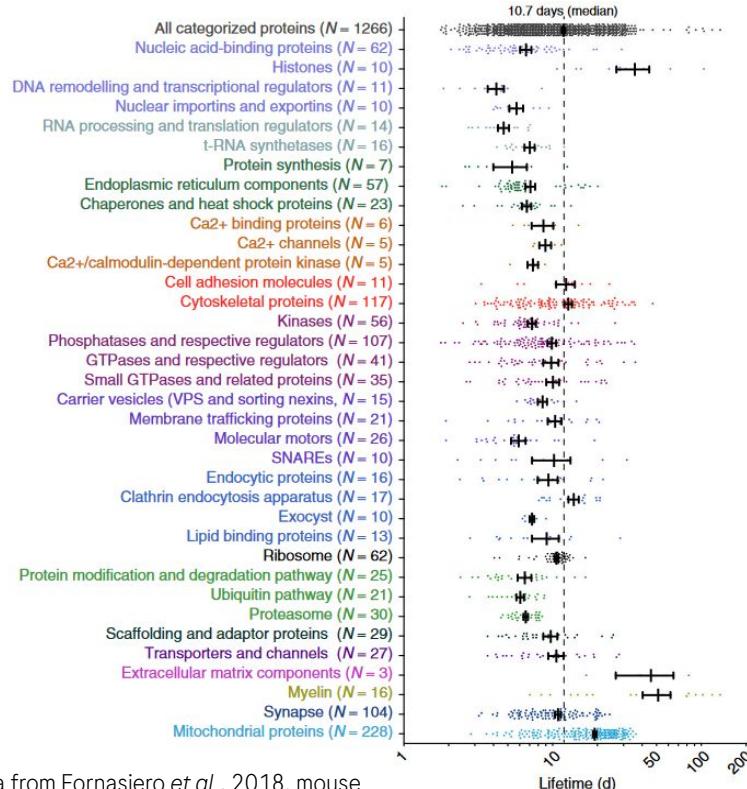


Data from
primary human B
cells (Mathieson
et al., 2018)



Data from primary hippocampal neuronal cells from rat
(Dörrbaum et al., 2020)

Protein half-life *in vivo* ranges between days and years



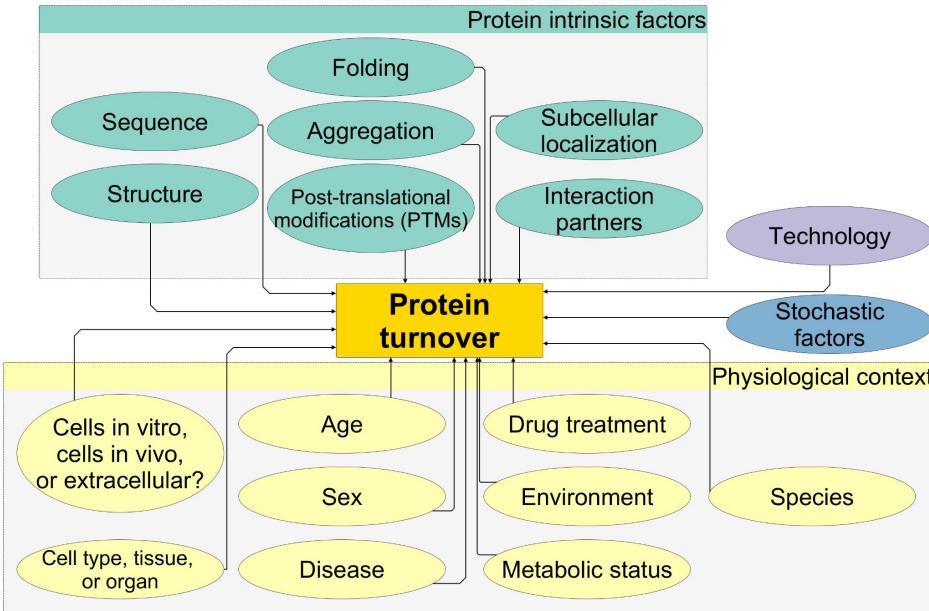
Data from Fornasiero et al., 2018, mouse brain

Table 1 | Known long-lived proteins and molecules

Protein or molecule*	Age [‡]	Measure	Organism
Eyelens crystallin	>70 years	Lifetime	Human
Collagen	117 years	Half-life	Human
Elastin	>78 years	Lifetime	Human
Enamel and dentine	>70 years	Lifetime	Human
Histones	223 days	Half-life	Mouse
	117 days	Half-life	Mouse
	218 days	Half-life	Rat
Nuclear pore proteins	>1 month	Lifetime	Worm
	>1 year	Lifetime	Rat
Myelin	95 days	Half-life	Rat
	>100 days	Half-life	Mouse
Myelin proteolipid protein	>100 days	Half-life	Mouse
REC8	Weeks	Lifetime	Mouse
mRNA	Possibly indefinite	Lifetime	Plant seed
	>2 years	Half-life	Frog oocyte
Cholesterol	>18 months	Lifetime	Rabbit
Phospholipids	>192 days	Lifetime	Rabbit

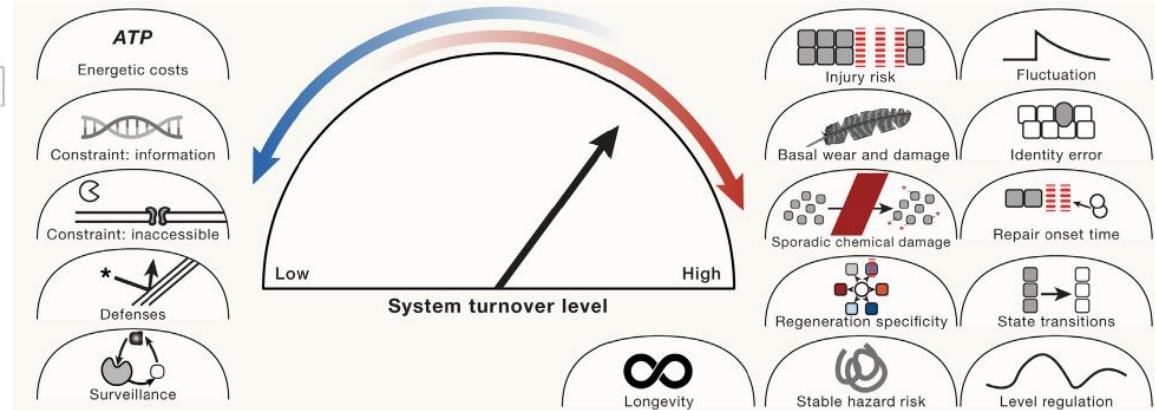
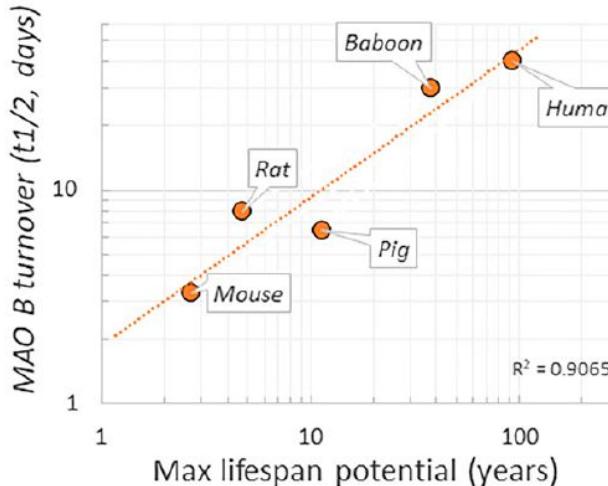
Toyama, Brandon H., and Martin W. Hetzer. “[Protein Homeostasis: Live Long, Won’t Prosper](#).” Nature Reviews Molecular Cell Biology, 2013

Half-life varies between proteins and contexts: influencing factors and an example



Condition	Half-life of protein X	Source
Human neurons <i>in vitro</i>	38.6h	Roche in-house data
Mouse neurons <i>in vitro</i>	34.1h (standard error: 3.9h)	Fornasiero et al., Nature Communications, 2018
Mouse cortex <i>in vivo</i>	619.2h, or 25.8d	Kluever et al., Science Advances, 2022

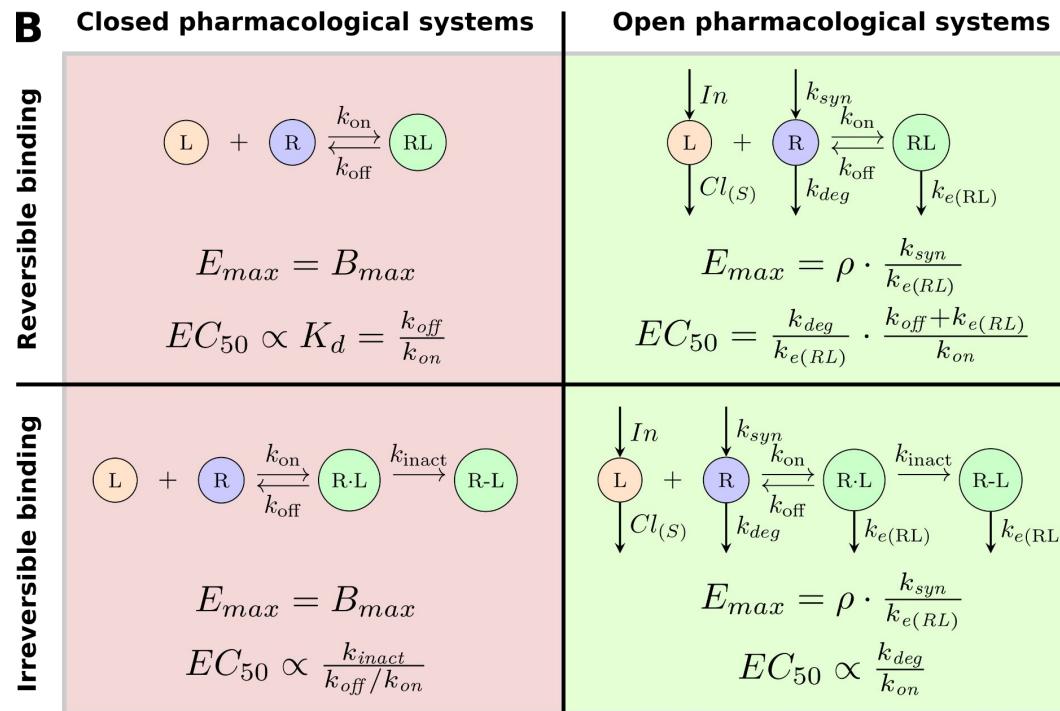
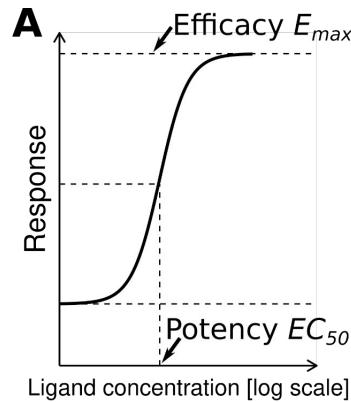
Nothing in Biology Makes Sense Except in the Light of Evolution: the purpose and ubiquity of turnover



Left: Gabrielsson, J., and S. Hjorth. 2023. "Turn On, Tune In, Turnover! Target Biology Impacts In Vivo Potency, Efficacy, and Clearance." *Pharmacological Reviews* 75 (3): 416–62. <https://doi.org/10.1124/pharmrev.121.000524>. Right: Reddien, Peter W. 2024. "The Purpose and Ubiquity of Turnover." *Cell* 187 (11): 2657–81. <https://doi.org/10.1016/j.cell.2024.04.034>. Quote: Theodosius Dobzhansky

Open models integrate protein turnover into pharmacological modeling

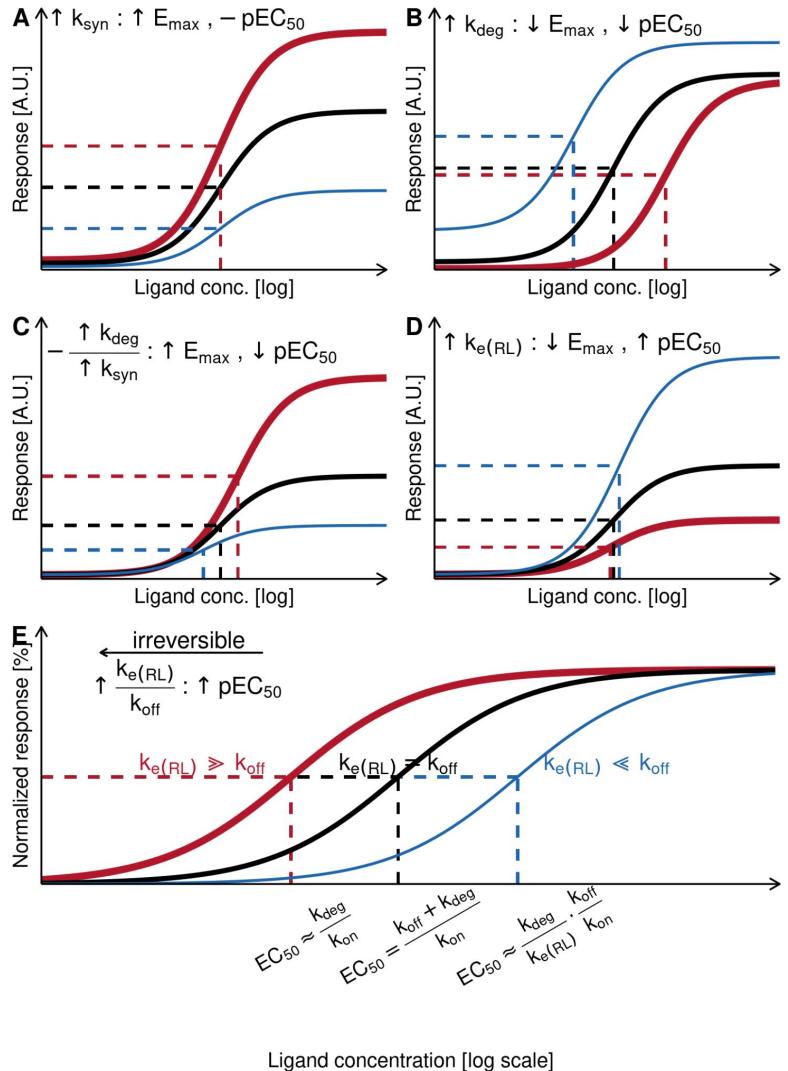
According to open models (see the comprehensive review by [Gabrielsson and Hjorth](#)), target turnover impacts *in vivo* potency, efficacy, and clearance.



Predictions by open models

Highlighted in blue: particularly relevant for covalent binders

- A. Higher target synthesis rate increases efficacy while potency remains unchanged.
- B. Higher degradation rate decreases both efficacy and potency.
- C. Keeping the steady-state abundance fixed, increasing both synthesis & degradation rates increase both efficacy and potency.
- D. Higher ligand-target complex elimination rate reduces efficacy while increases potency.
- E. Potency of covalent inhibitors is dictated by k_{deg}/k_{on} : slow turnover and fast on-rate are preferred.



Ligand concentration [log scale]

Roche's Protein Turnover Database integrates external and internal data

The table shows the protein half-life datasets that David curated for the turnover database. The curation contains following steps:

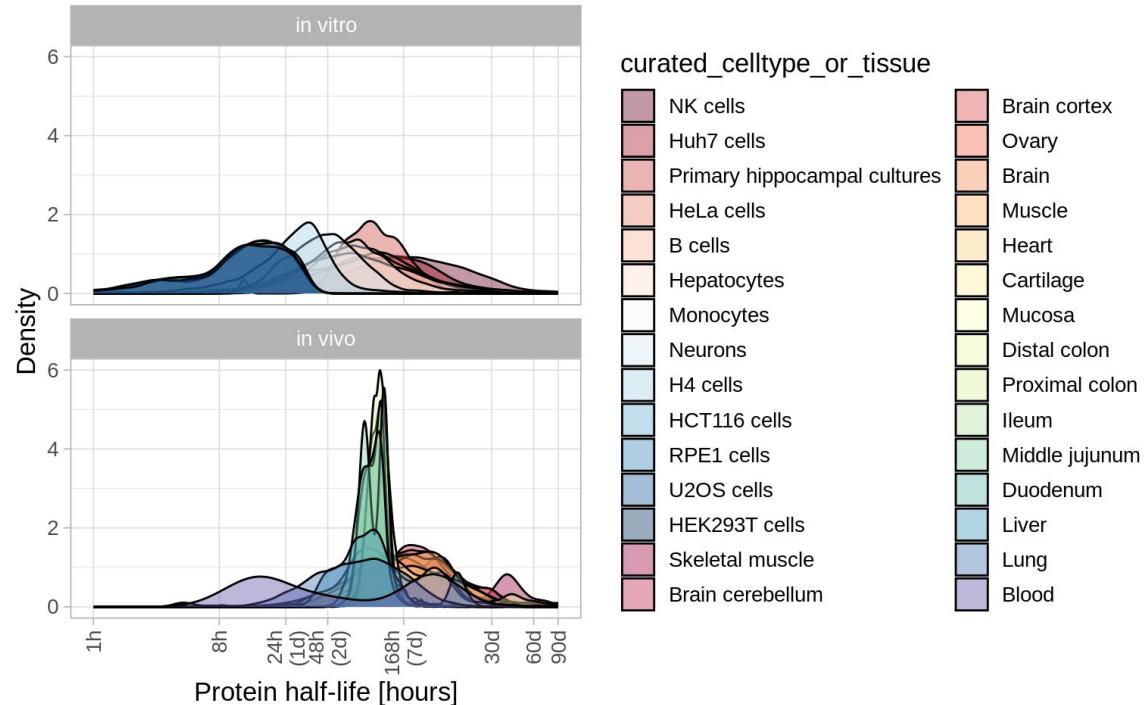
1. The data were curated from individual studies.
2. Features (uniprot IDs, protein groups, etc.) were harmonized and mapped to genes of the respective genome as well as to human orthologues.
3. Units of measurements were harmonized to hours.
4. Sample annotations are harmonized.

Roche Protein Turnover Database			
Dataset overview (v202407)			
	organism	assay_type	celltype_or_tissue
Doerrbaum-2018	rat	in vitro	Primary hippocampal cultures
Fornasiero-2018	mouse	in vivo	Brain cortex, Brain cerebellum, Heart, Muscle
Mathieson-2018-human	human	in vitro	NK cells, Hepatocytes, Monocytes, B cells
Mathieson-2018-mouse	mouse	in vitro	Neurons
Arike-2020	mouse	in vivo	Duodenum, Middle jejunum, Ileum, Proximal colon, Distal colon
Li-2021	human	in vitro	U2OS cells, HEK293T cells, HCT116 cells, RPE1 cells
Morgenstern-2021	human	in vitro	HeLa cells, Huh7 cells
Rolfs-2021	mouse	in vivo	Cartilage, Skeletal muscle, Mucosa, Liver, Blood
Kluever-2022	mouse	in vivo	Brain cortex, Brain cerebellum
Chen-2023	mouse	in vivo	Lung, Heart, Brain
Harasimov-2024	mouse	in vivo	Ovary
Lothar-H4	human	in vitro	H4 cells

We observe in general longer half-life *in vivo* than *in vitro*, with variations between cell/tissue types

Right: density plot of protein half-life, stratified by assay type (*in vitro* versus *in vivo*) and by cell type or tissue.

Most *in vivo* studies tend to report longer half-life than at least some *in vitro* studies, though considerable variability is observed in both categories.

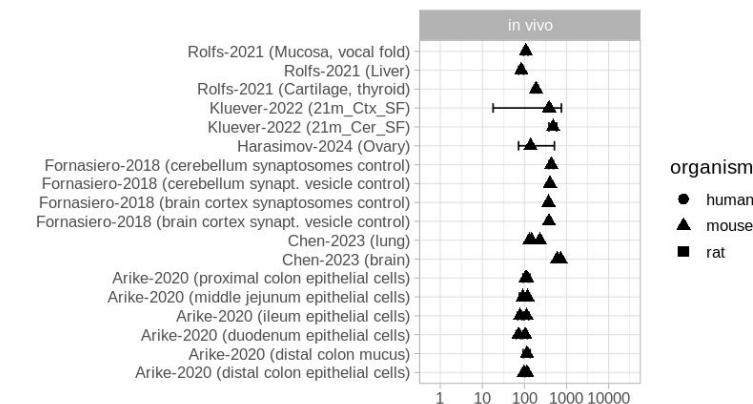
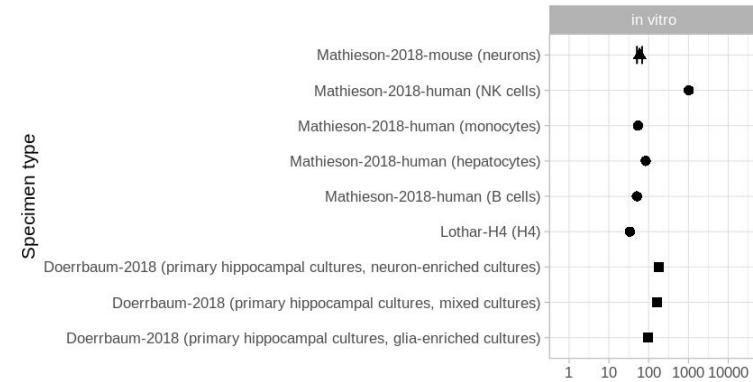


A survey of half-life of covalent binder targets

We curated 31 covalent binders which are either in clinical development or approved, targeting a total of 26 human and 7 viral or bacterial proteins.

The table summarizes half-life data for 24 human proteins. Turnover data of KRAS is visualized with boxplots.

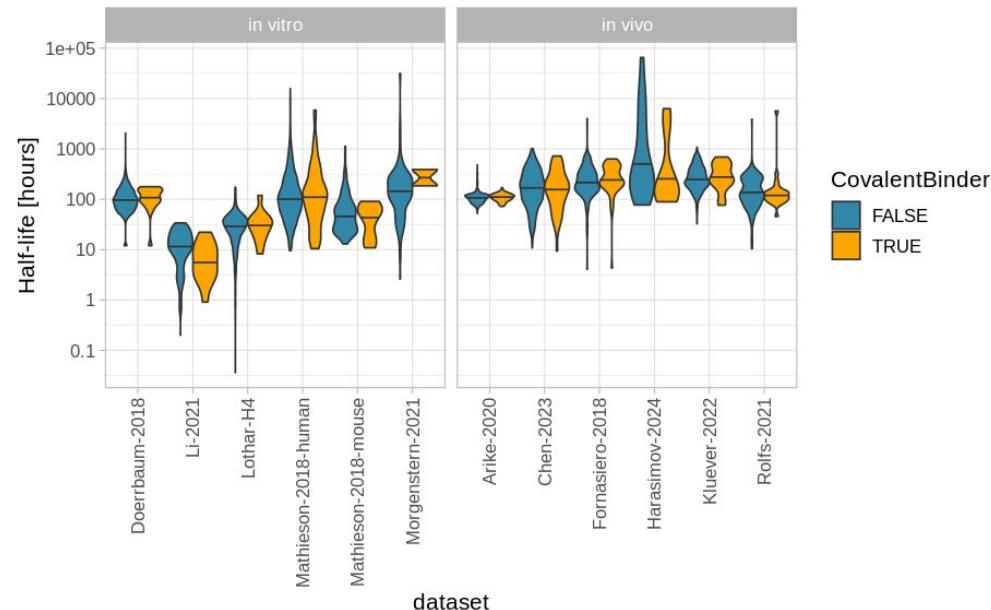
	Median half-life of targets of covalent binders [hours]		
	unique_drugs	in vitro	in vivo
EGFR	8	35.0	64.9
ERBB2	4	17.8	NA
BTK	3	79.2	NA
KRAS	3	84.3	124.6
PSMB5	3	109.7	212.6
ERBB4	2	19.3	70.2
MAOB	2	111.8	272.6
P2RY12	2	194.8	167.3
PSMB1	2	129.9	163.9
ABAT	1	185.0	433.6
ATP4A	1	NA	136.8
FGFR4	1	7.7	NA
HBA1	1	119.2	NA
HMGCR	1	10.6	5648.8
JAK1	1	12.5	89.1
JAK2	1	57.0	NA
JAK3	1	10.4	NA
PSMB10	1	160.7	46.9
PSMB2	1	133.3	197.5
PSMB8	1	119.8	128.5
PSMB9	1	203.2	266.1
PTGS1	1	667.6	145.1
PTGS2	1	8.2	NA
TYK2	1	20.5	NA



Targets of covalent binders have comparable half-life with targets of non-covalent binders, yet short-living proteins are less targeted by the covalent approach

The violin plot compares the half-life of targets of covalent binders ($N=24$) with the half-life of targets of non-covalent molecules for which a high potency or functional inhibition ($pACT \geq 8$, $N=788$).

Targets of covalent binders and those of non-covalent drugs have in general comparable half-lives. However, covalent drug targets are devoid of shortest-living proteins.

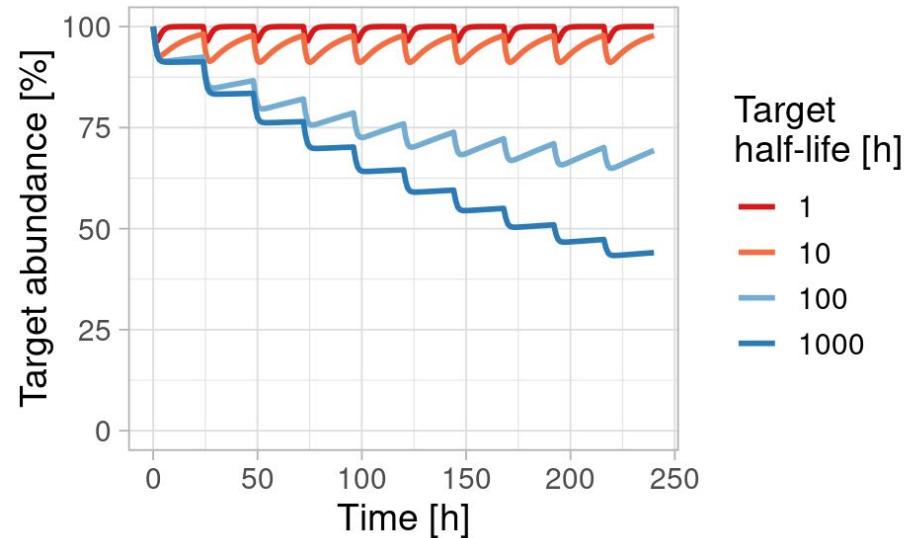


Protein half-life can be integrated into PK/PD models

Example: [target degradation PK/PD model of covalent binding](#) by Andrés Olivares

Modelling and simulation suggests that the PD effect of target degradation by a covalent binder is sensitive to target's turnover. Long-living proteins are more likely to become successful targets for covalent inhibitors.

[Bayer colleagues](#) also reported that half-life is a key parameter affecting the predictions of mechanistic PD models for targeted protein degraders.



Further points for consideration

1. Open Models and the importance of protein turnover does not only affect covalent binders: they are applicable to reversible and irreversible drug-target interactions, as well as to all protein targeting modalities including small molecules, large molecules (for instance antibodies), and PROTACs.
 - a. By taking consideration of the dynamics of RNAs, the Open Models can be extended to RNA-targeting modalities as well as gene therapies.
2. Protein turnover does not only affect drug's potency and duration of response *in vivo*: turnover of enzymes and transporters also affects metabolism and transport.
3. Looking forward, we believe open models, together with experimental data and/or predictions based on modeling and simulation and machine learning/generative models, can help us rationally select modalities. Many experiments are on-going or being planned. We look forward to collaborations.

Protein turnover is one slice of our work in drug discovery

1. Roche's research projects and collaborations
2. Disease model characterization and validation
3. Phenotypic drug discovery
4. Preclinical assay development
5. New experimental and computational tools for clinical drug candidate selection
6. Predicting and translating pharmacology and safety profiles between systems
7. Method and software development
8. Analysis of attritions in drug discovery and development

Selected references:

1. Van der Vries *et al.*, 2015; Zaidan *et al.*, 2020; Gatti *et al.*, 2022;
2. Grabole *et al.*, 2016; Reich *et al.*, 2021; Bosch *et al.* 2024
3. Moisan *et al.*, 2015; Roudnicky *et al.*, 2020; Wang *et al.*, 2023; Rodriguez-Iglesias, 2024
4. Zhang *et al.*, 2015; Moisan *et al.*, 2017; Jaklin *et al.*, 2020; Jaklin *et al.*, 2022; Ruegger *et al.* 2025
5. Zhang *et al.*, 2014; Boess *et al.* 2017; Mueller *et al.*, 2018; Chen *et al.*, 2022; Klughammer *et al.*, 2023
6. Zhang *et al.*, 2025 (in revision)
7. Zhang *et al.*, 2017; Choobdar *et al.*, 2019; Sturm *et al.*, 2019; Mädler *et al.*, 2021; Crouzet, *et al.* 2024; Rot *et al.*, 2024
8. Zhang *et al.* (in preparation)

Acknowledgement

Neil John Parrott, Martin Ebeling, Kenichi Umehara, Holger Fischer, Roland Schmucki, Andrés Olivares, Uwe Grether, Marcus Bantscheff, Miro Eigenmann, Björn Bartels, Joachim Rudolph, Lothar Lindemann, Sarah Morillo Leonardo, Bernd Kuhn, Torsten Schindler, Lizzie Gill, Eugenio Fornasiero, Christophe Fromont, Frederik Rode, Stewart Fisher, Alessio D'Addio, Arne Rufer, Bioinformatics Club, PBPK-ADME-DPL (PADCo) community, Discovery Safety (DISCo) community, Chemical Biology Sphere, Small Molecule Data Analytics (SMDA) Network, Roche PMDA Summer School 2024 cohort.





Predicting Protein Half-life for Drug Discovery

Jitao David Zhang on behalf of the Predictive Modeling And Data Analytics (PMDA) chapter, Pharmaceutical Sciences, Pharma Research and Early Development (pRED), Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd.

Doing now what patients need next

Overview of the dataset and the tasks

We share with you seven curated datasets, which are summarized in the table below.

Together the dataset contains about 100,000 triplets of mouse proteins, samples, and protein half-life data. We share 99% of the data with you. About 1% of the data was held back.

Your tasks: (1) performing analysis to explore the dataset, and (2) building models to predict the held-back test data.

The held-back test data will be shared with you on Thursday. Each team is expected to make a 15-minute presentation about their methods, achievements, and learnings on Friday afternoon.

PMDA Summer School 2024			
Dataset overview			
	organism	assay_type	celltype_or_tissue
Fornasiero-2018	mouse	in vivo	Brain cortex, Brain cerebellum, Heart, Muscle
Mathieson-2018-mouse	mouse	in vitro	Neurons
Arike-2020	mouse	in vivo	Duodenum, Middle jejunum, Ileum, Proximal colon, Distal colon
Rolfs-2021	mouse	in vivo	Cartilage, Skeletal muscle, Mucosa, Liver, Blood
Kluever-2022	mouse	in vivo	Brain cortex, Brain cerebellum
Chen-2023	mouse	in vivo	Lung, Heart, Brain
Harasimov-2024	mouse	in vivo	Ovary

Organisers, Guests, and Agenda



Flavia
Spielvogel



Nina
Lareida



Niklas
Trapp



Jannick
Lippuner



Chih-Hsuan
Hsin



Ercan
Sükür

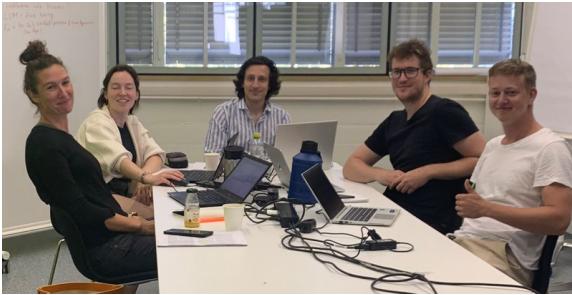


Jitao David
Zhang

	Monday (5th of August)	Tuesday (6th of August)	Wednesday (7th of August)	Thursday (8th of August)	Friday (9th of August)
8:30 - 10:00	Introduction	Group Work	Group Work	Group Work	Group Work
10:00 - 10:30	Snack				
10:30 - 13:00	Introduction (continued)	Group Work	Group Work	Group Work	Group Work
13:00 - 14:00	Lunch Break				
14:00 - 17:00	Group Work	Group Work	Group Work	Group Work	Team Presentation
16:00 - 16:30			Tune out	15:00 - 15:45 Talk by Ercan Sükür	
16:30 - 18:00	Group Work	Group Work	15:45 - 16:30 Roche Tour	16:00 - 16:30 Talk by Chih-hsuan Hsin	16:00 - 16:30 Tune out
17:00 - 17:30			Tune out	16:30-17:00 Tune out	16:30-17:30 Apero

Wednesday from 18:30 on: Social Dinner @ Don Camillo

Impressions



Scientific outcome: a summary

- Three teams came up with distinctive analysis and predictive models.
- A large range of machine-learning models were used, from linear regression, random forest, to protein language models.
- The teams explored also the biology of protein turnover, using available information as input for machine-learning models.
- Most models performed similar to the baseline ANOVA model, however more in-depth analysis is needed. Some features fit our previous knowledge (e.g. mitochondrial proteins tend to have longer half-lives), while some features are interesting (e.g. methionine content seems to be correlated with half-life).

“As I settle back into my lab, I already miss the inspirational atmosphere from our tune-out meetings. The dedicated group work and the exchange with other students were incredibly motivating.”

Except from the follow-up email of a participant

Doing now what patients need next