ETH zürich

Drug design with machine intelligence

Gisbert Schneider

Generative molecular design

©Jason Allen, “Théâtre D’opéra Spatial” (2022)

- The chemistry and pharma sectors **lag behind**
- **Rapidly growing field** (>1000 companies)
- ”Big pharma” is developing **in-house A.I. capacities**

The drug discovery cycle: The bots are coming!

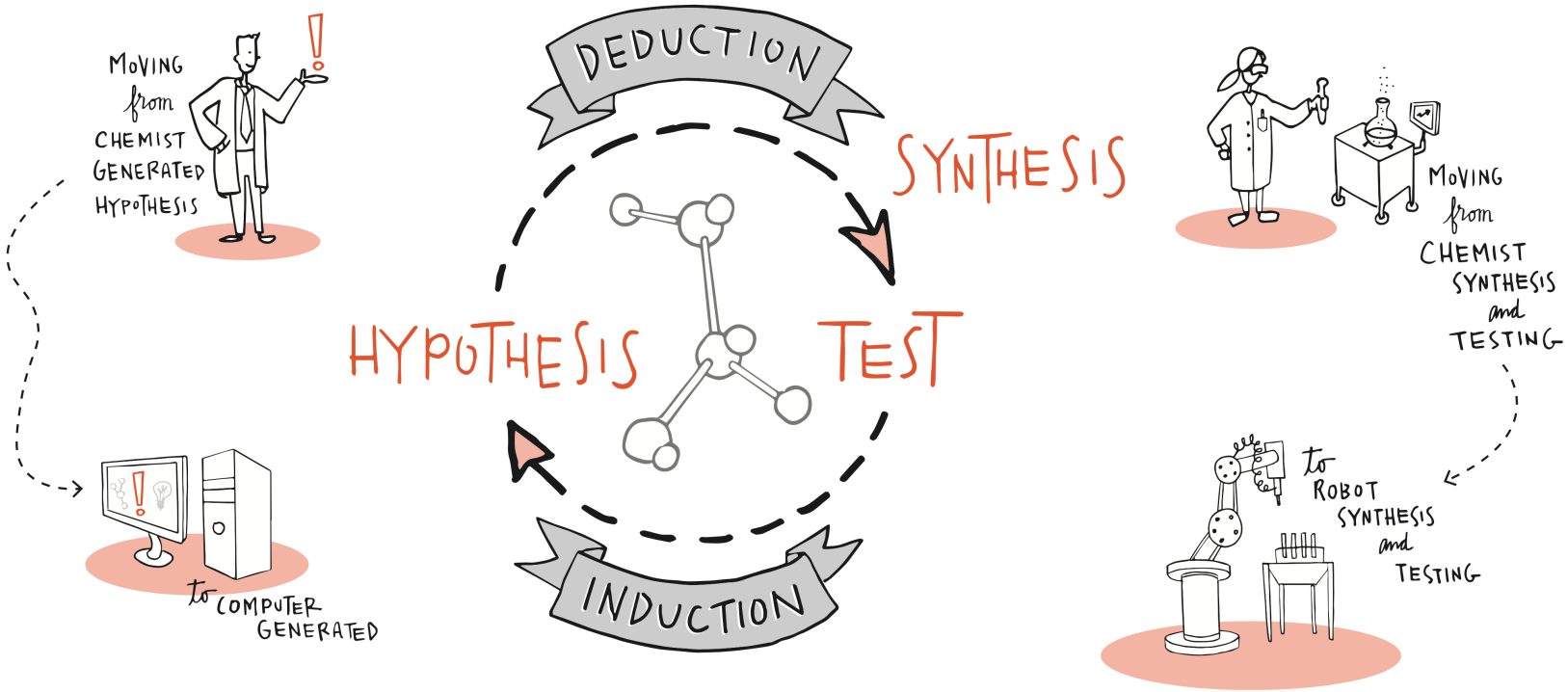


Image credit: Jack Burgess

Nat. Mach. Intell. **2019**, *1*, 128.
Nat. Rev. Drug Discov. **2020**, *19*, 353.

Modern drug discovery: From screening to designing new molecules

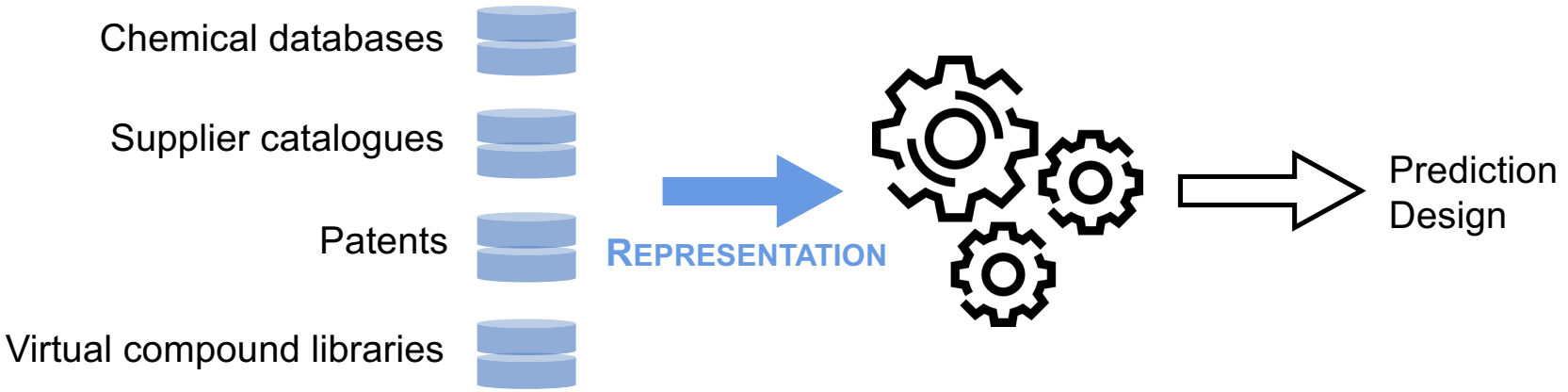
10^{100} → **1 – 10**

MACHINE LEARNING
ARTIFICIAL INTELLIGENCE

- Generative design supports chemists by suggesting **surprising novel ideas**
- Examples: Chemical language model, graph neural network, transformer, GAN
- **Complements or replaces high-throughput screening**

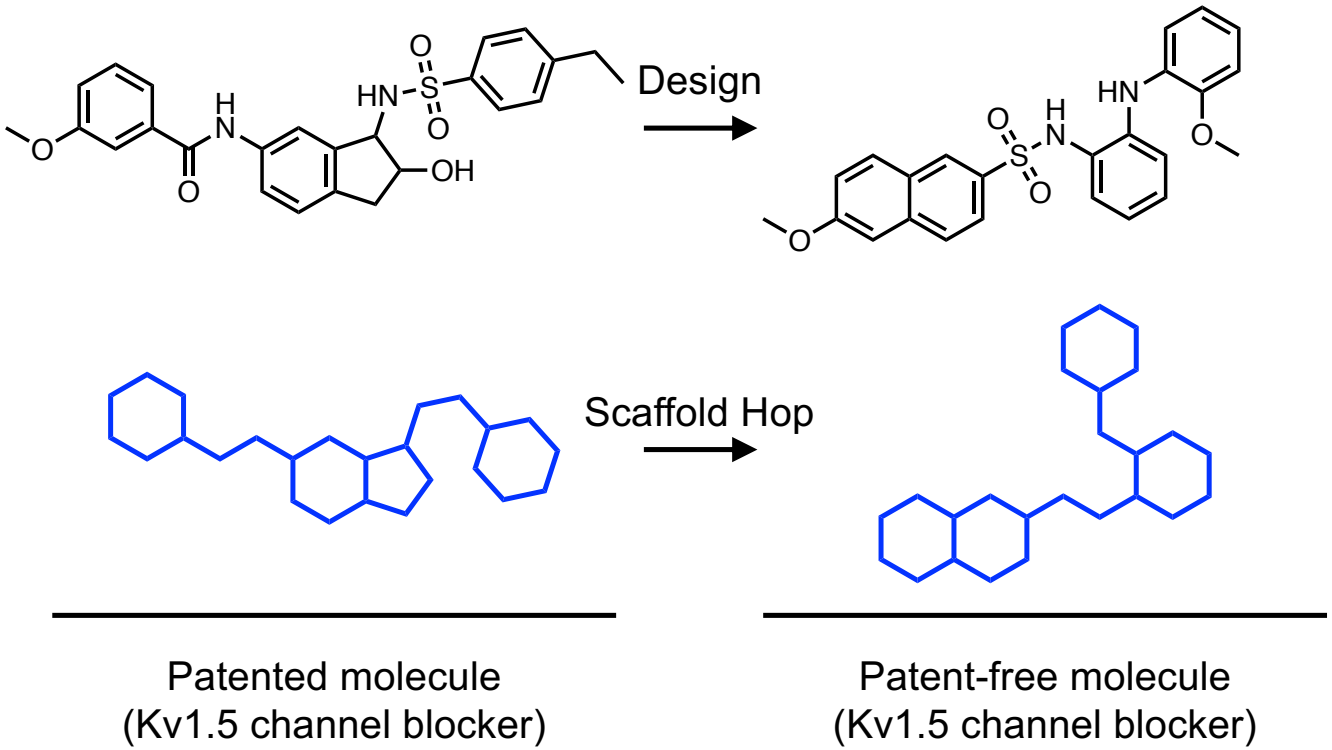
Nat. Mach. Intell. **2019**, *1*, 128.
Nat. Rev. Drug Discov. **2018**, *17*, 97.
J. Med. Chem. **2016**, *59*, 4077.

Sources of information for drug design with machine intelligence



- Computational drug design suffers from **scarce data**
- **Context-specific data representation** is essential

“Patent busting” by automated “Scaffold Hopping”

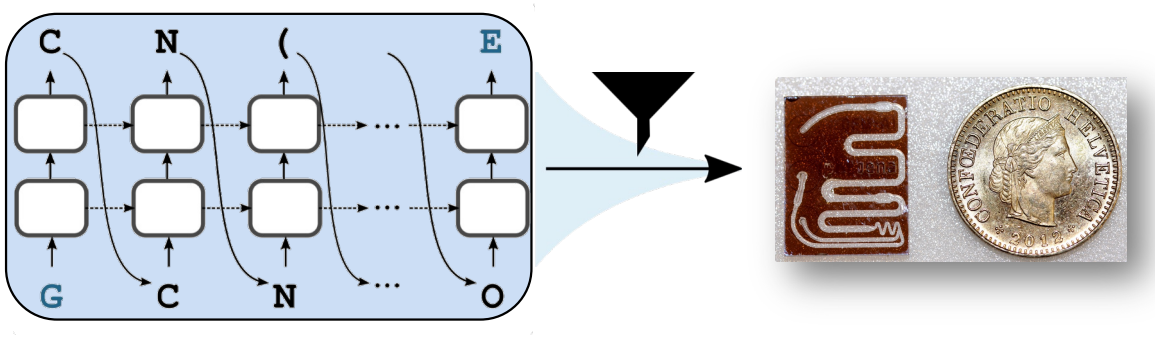


- **De novo design** generates new molecules
- **Fast and reliably** (>50% success)
- **Novel I.P.** and chemical scaffold classes

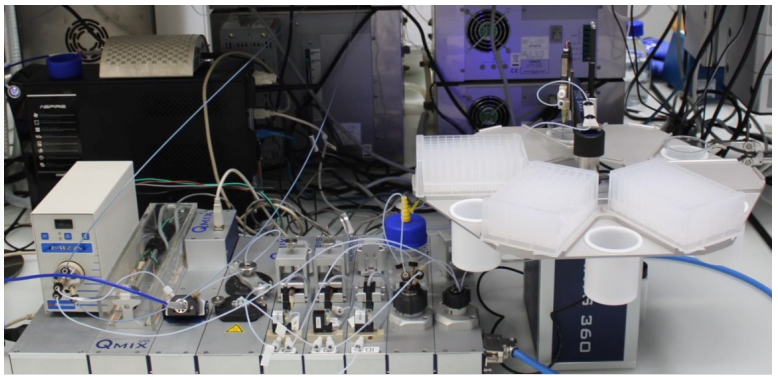
Angew. Chem. Int. Ed. **2000**, *39*, 4130.

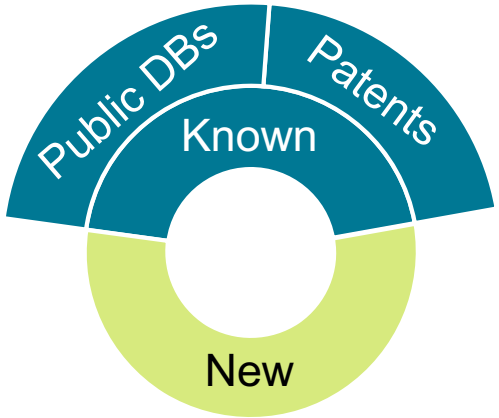
Integrated drug discovery laboratory

Chemical language model



Automated synthesis

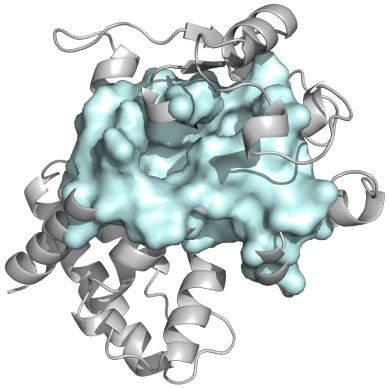




- 55% novel chemical structures
- 21% recreated patented molecules
- 64% successfully synthesized
- 68% bioactive

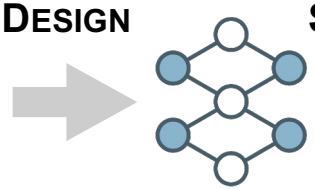
Sci. Adv. **2021**, *7*, eabg3338.

Protein structure-based drug design

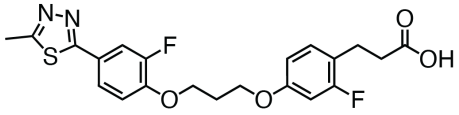


PPARγ
Ligand binding pocket

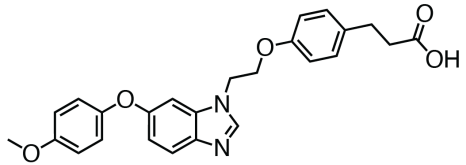
DESIGN **SYNTHESIS**



3D graph network
Message passing network
Transformer network
Chemical language model



De novo molecule **1**



De novo molecule **2**

TESTING

EC_{50} = 1.5±0.2 µM
14.6±0.6-fold activation

EC_{50} = 2.3±0.7 µM
4.4±0.4-fold activation

- ✓ Novel chemical structures
- ✓ LXRα, RARα, RXRα: inactive
- ✓ CYP panel: inactive
- ✓ Anti-target panel: inactive

Nat. Commun. **2024**, *15*, 3408.

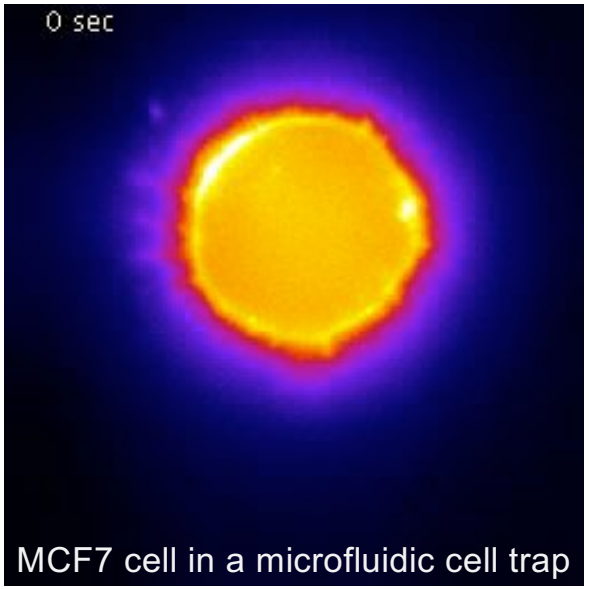
Automated A.I.-driven design of selective anticancer peptides

Known peptide sequences

P(**X**,**Y**)

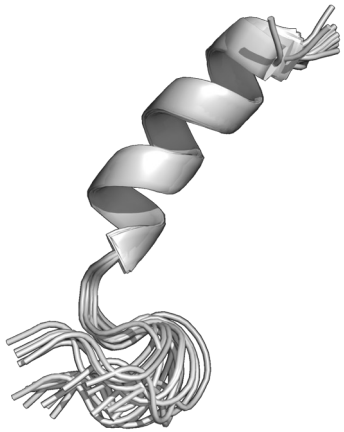
New sequence(s)

KLWKKIEKLIKKLLTSIR
YIWARAERVWLWWGKFLSL
ELAKKLTKLKRQLHRIW
DLFKQLQRLFLGILYCLYKIW
KLIDQWKKVLYHVE
...



MCF7 cell in a microfluidic cell trap

Video: Petra Dittrich, Lukas Armbrechter



KVWKIKKNIRRLLHGIKRGWKG

- **Chemical language model** for peptide design
- Novel amino acid **sequences not found in nature**

Angew. Chem. Int. Ed. **2019**, *58*, 1674.
J. Chem. Inf. Model. **2018**, *58*, 472.

ETH zürich | CURE

Gisbert SCHNEIDER
gisbert@ethz.ch
ETH Zurich

www.modlab.ethz.ch