ProteinQure

# Protein design using quantum/classical hybrid algorithms

Mark Fingerhuth, CEO
mark@proteinqure.com

# Quantum machine learning incubator

## Quantum Machine Learning Program

### Mission

By 2022 the QML Program will have produced more well- capitalized, revenue generating quantum machine learning software companies than the rest of the world combined. The majority of these will be based in Canada.

CREATIVE DESTRUCTION LAB

40 participants

↓

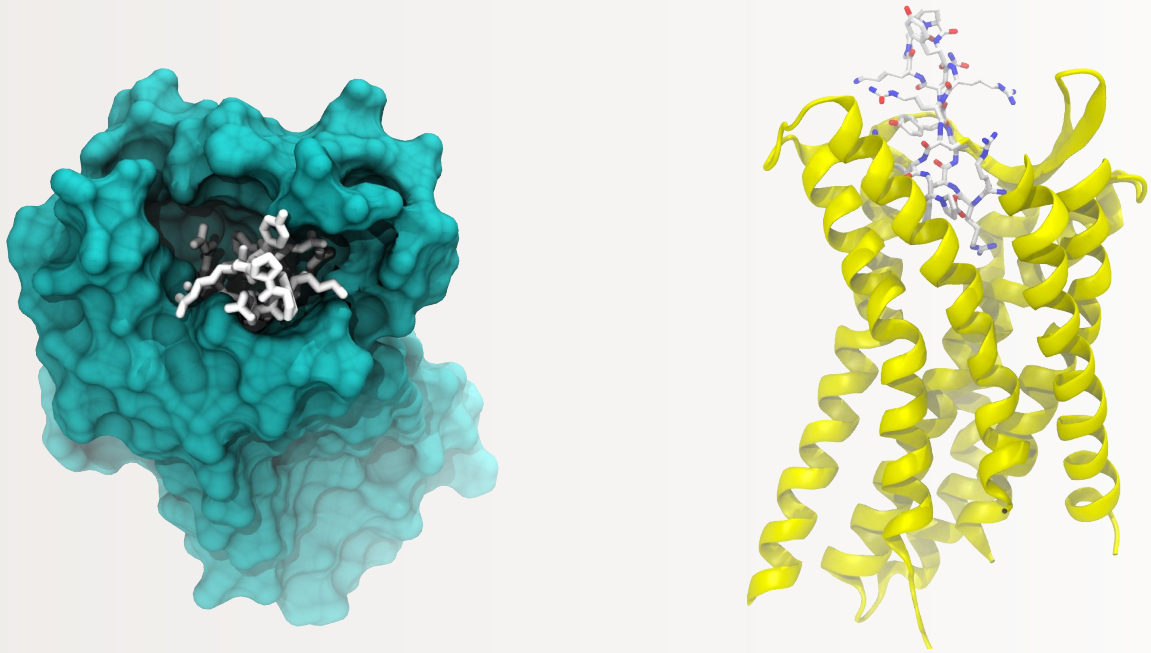
25 companies started

↓

ProteinQure + 2 companies ‘graduated’

ProteinQure

# Drug discovery:
# “finding the key to the lock”



CXCR4 chemokine receptor in complex
with cyclic peptide, Science 330 (2010)

ProteinQure

# Computational design of protein therapeutics

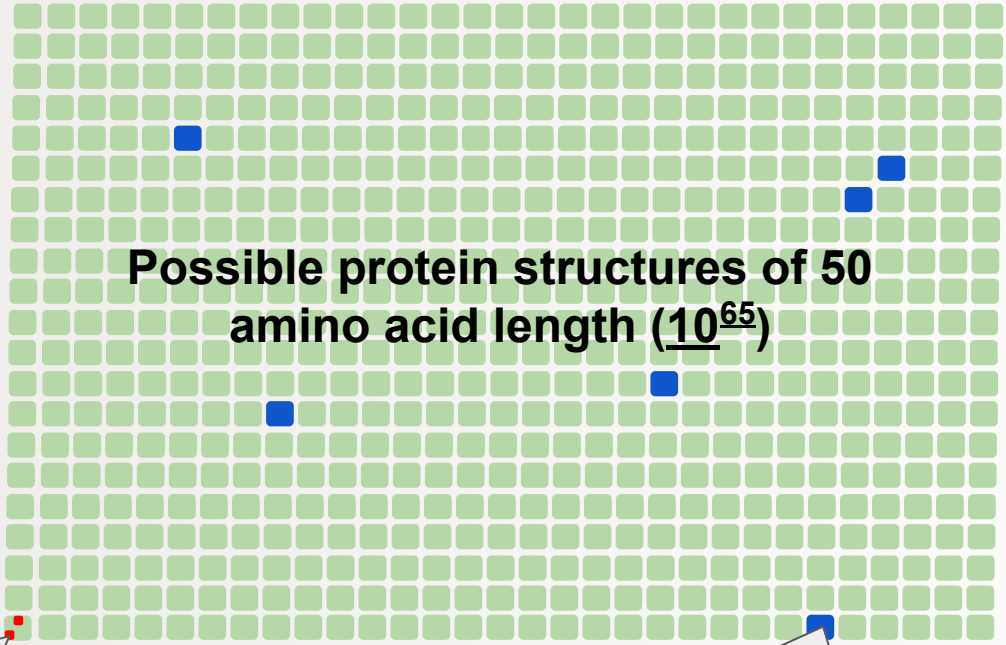

**Small molecule drugs**
(aspirin)

**Next-Gen Protein Therapeutics**
(novel recombinant proteins)

**Antibodies**
(Novoeight®)

**ProteinQure** is developing novel algorithms for designing next-generation protein therapeutics using quantum computing and machine learning.

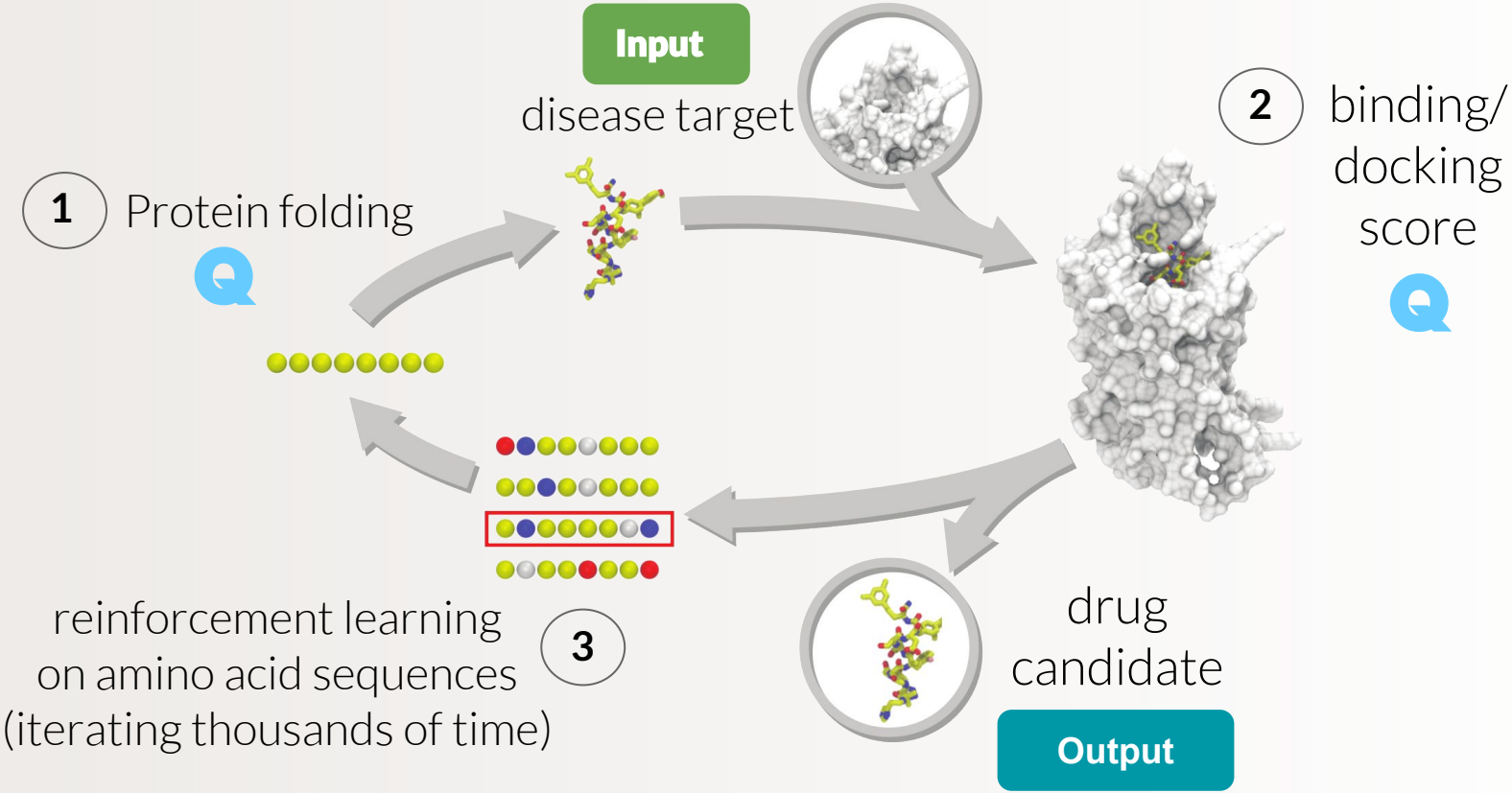
# Protein drug design is limited by data availability and potential search space



# Our solution

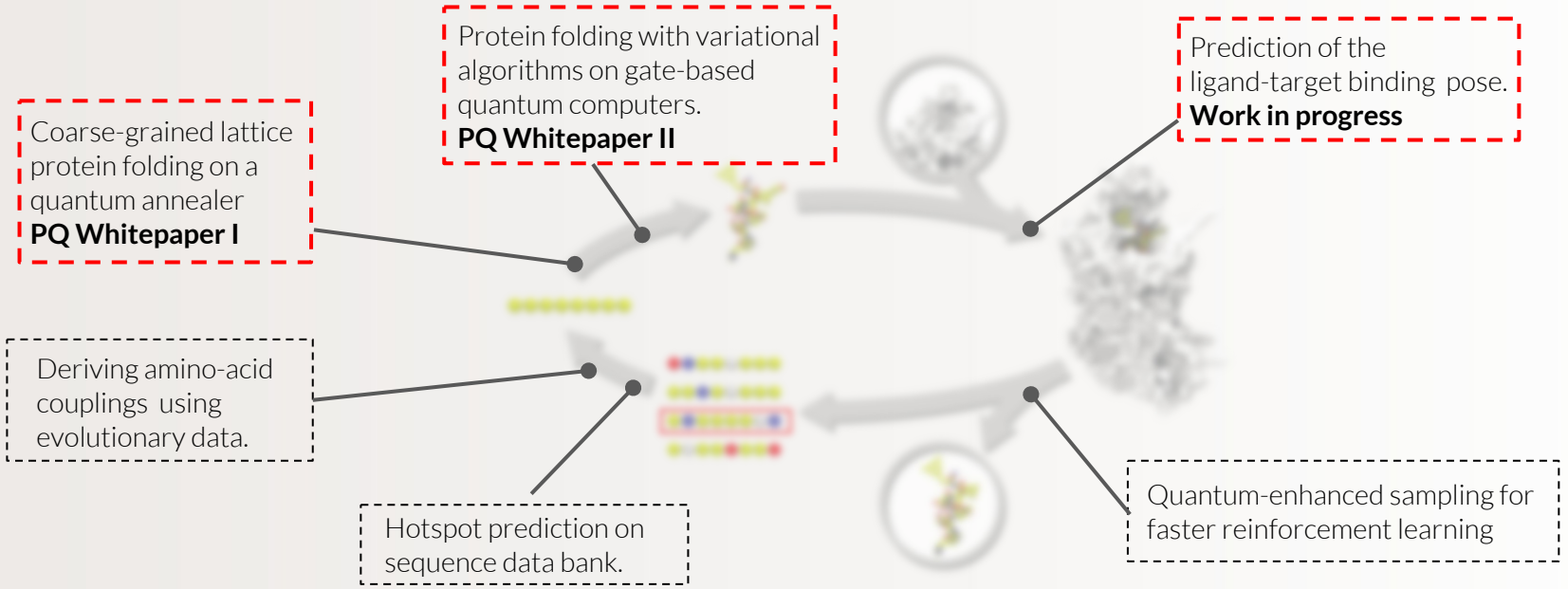
## Physical Simulations, Quantum Computing + Reinforcement Learning

ProteinQure



# Quantum optimization
## can speed up protein design

ProteinQure



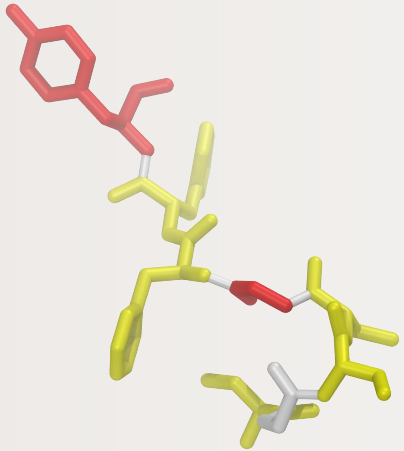

**Hybrid methods will lead to speedups and the ability to scale to new problem sizes in the future.**

# Lattice protein folding
## with quantum annealers

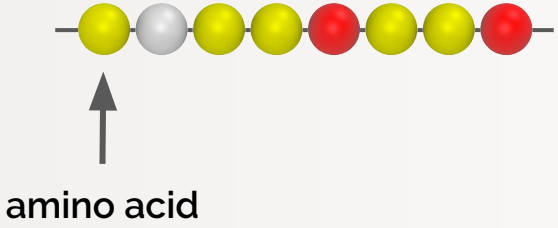
ProteinQure

# Coarse-grained lattice protein folding

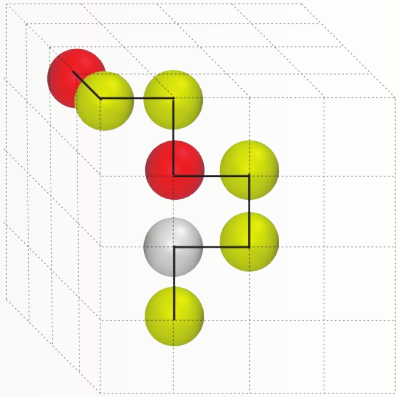
ProteinQure



Fully atomistic protein model




Coarse-grained protein model

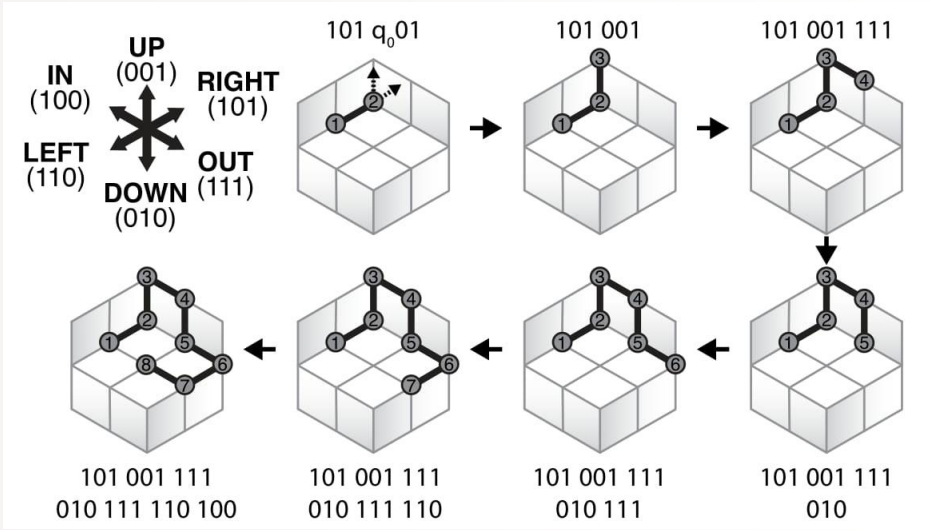


Lattice protein fold: self-avoiding walks on lattice

Babej, Fingerhuth & Ing (2018). PQ Whitepaper I (on arXiv soon)

# Ising encoding of lattice protein folding

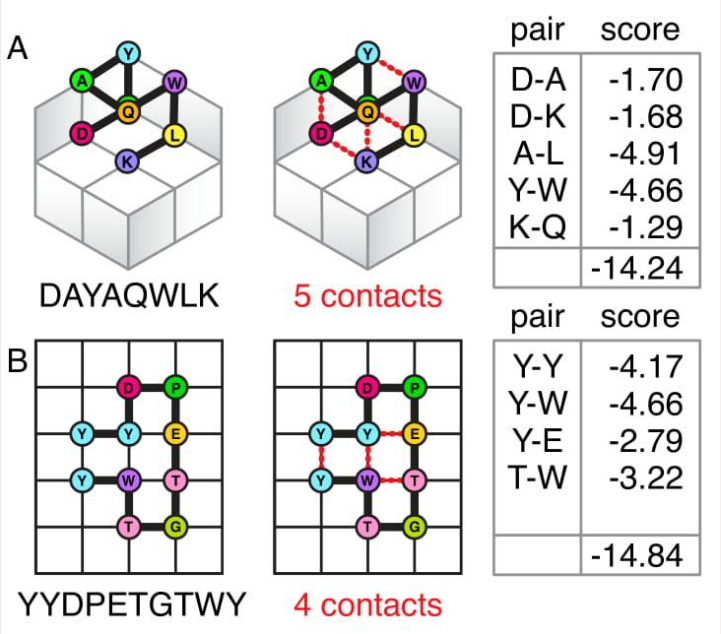
- Early work by Babbush, Perdomo-Ortiz & Aspuru-Guzik (Phys. Rev. A. (2008), Adv. Chem. Ph. (2012) & arXiv (2012)) on planar lattices
- They were first to map lattice protein folding to an **Ising-type Hamiltonian**
- We **generalized** the approach to cubic lattices
- **Improved** the implementation from quadratic to quasilinear [ O(n log n) ]




$$H(\mathbf{q}) = H_{\text{back}}(\mathbf{q}) + H_{\text{redun}}(\mathbf{q}) + H_{\text{olap}}(\mathbf{q}) + H_{\text{pair}}(\mathbf{q}).$$

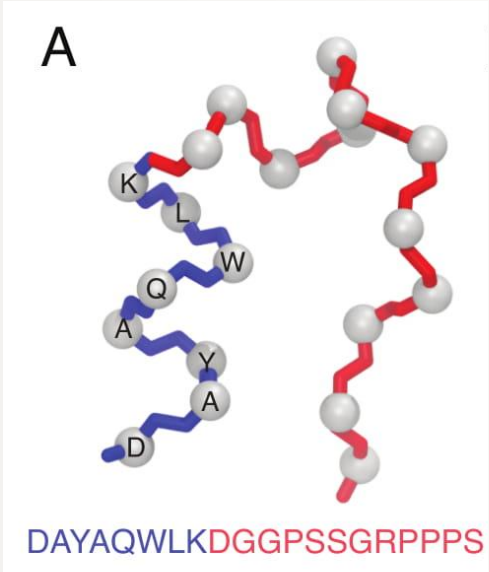
Babej, Fingerhuth & Ing (2018). PQ Whitepaper I (on arXiv soon)

# Results obtained from the D-Wave 2000Q




| pair | score |
|---|---|
| D-A | -1.70 |
| D-K | -1.68 |
| A-L | -4.91 |
| Y-W | -4.66 |
| K-Q | -1.29 |
| | -14.24 |

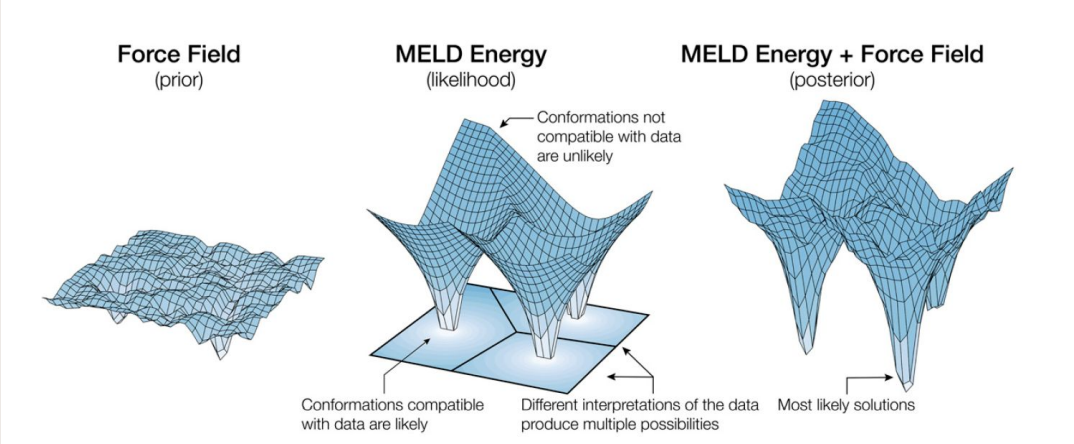
| pair | score |
|---|---|
| Y-Y | -4.17 |
| Y-W | -4.66 |
| Y-E | -2.79 |
| T-W | -3.22 |
| | -14.84 |



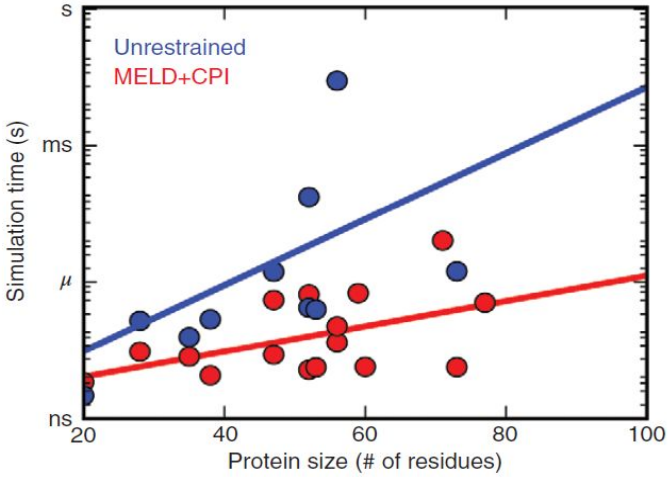

**We folded the largest proteins on today's quantum architectures (10 amino acids in 2D, 8 amino acids in 3D).**

Babej, Fingerhuth & Ing (2018). PQ Whitepaper I (on arXiv soon)

# Restraint-based molecular dynamics simulations

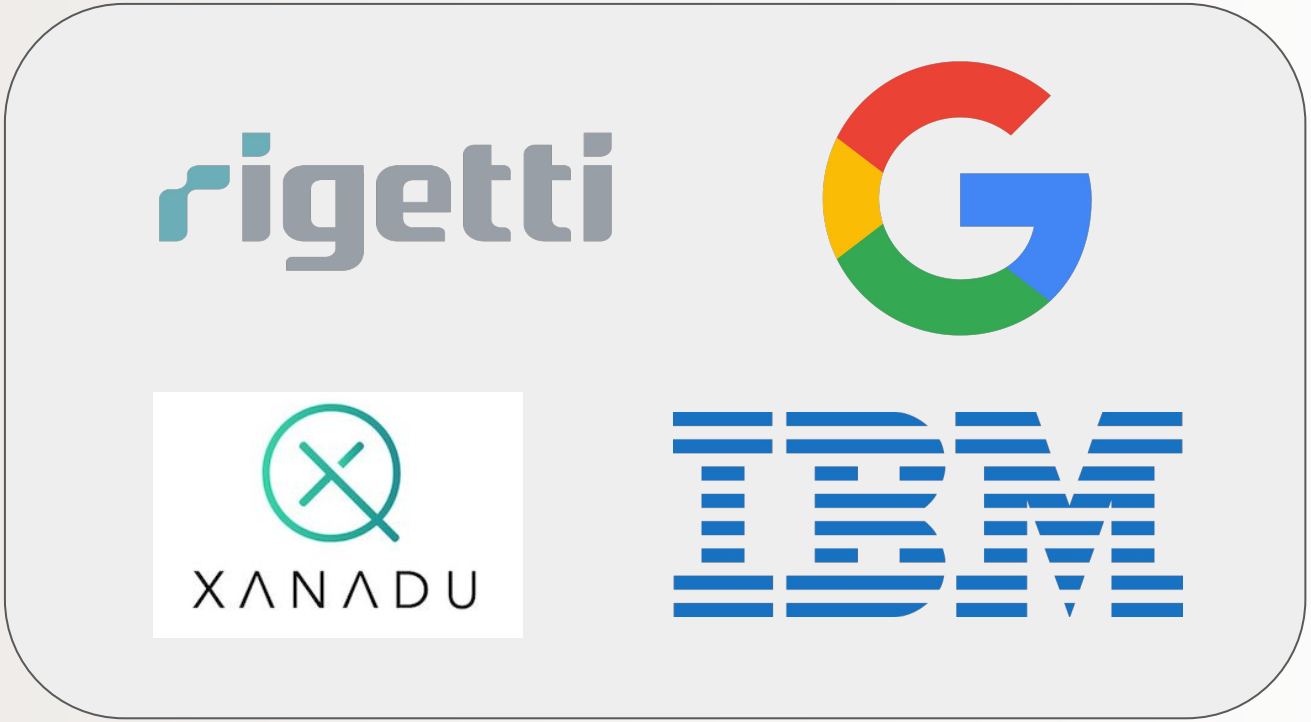



$$\overbrace{p(\mathbf{x}|\mathbf{D})}^{\text{posterior}} = \frac{p(\mathbf{D}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{D})} \sim \overbrace{p(\mathbf{D}|\mathbf{x})}^{\text{likelihood}}\overbrace{p(\mathbf{x})}^{\text{prior}}.$$




MacCallum et al. (2015). *Proc. Natl. Acad. Sci*, *112*(22), 6985-6990.

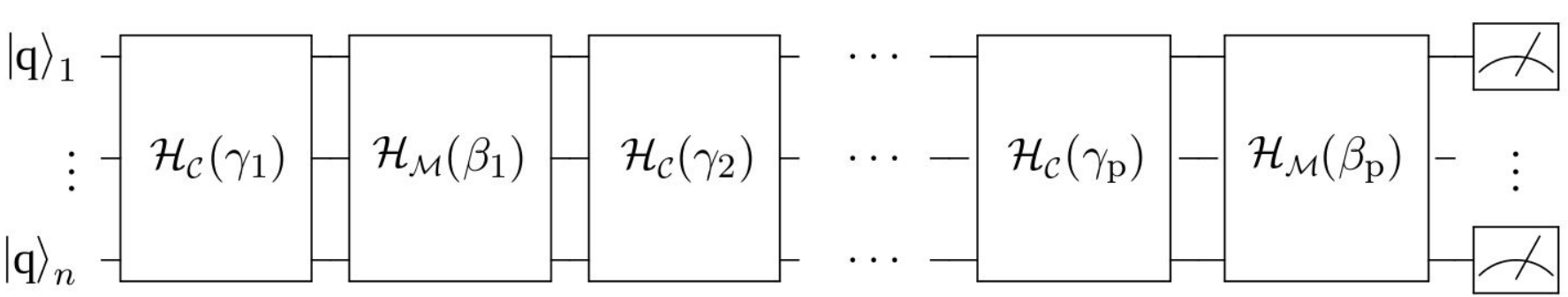
# Lattice protein folding
# with universal gate-based QCs




ProteinQure

# Quantum approximate optimization algorithm (QAOA)

ProteinQure




**Maximize** $\langle \gamma, \beta | C | \gamma, \beta \rangle$

Farhi et al. (2014). arXiv:1411.4028

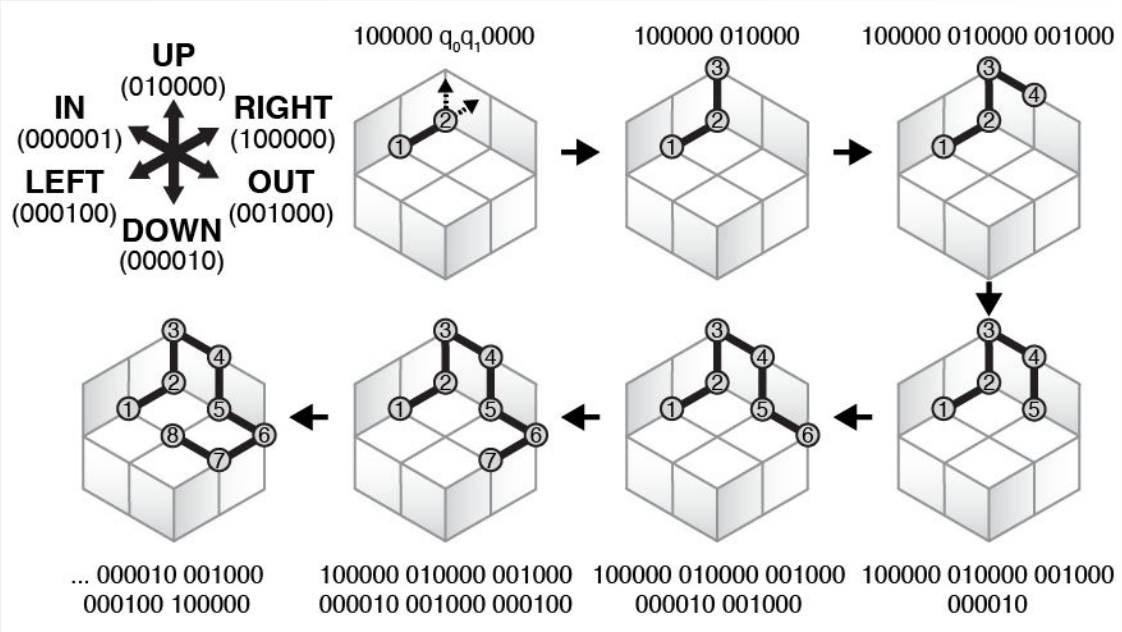
# One-hot encoding turns on a cubic lattice

ProteinQure

We one-hot encode the 6 spatial directions to ensure

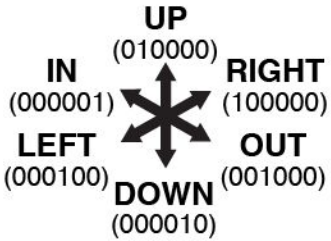
**constant Hamming weight**

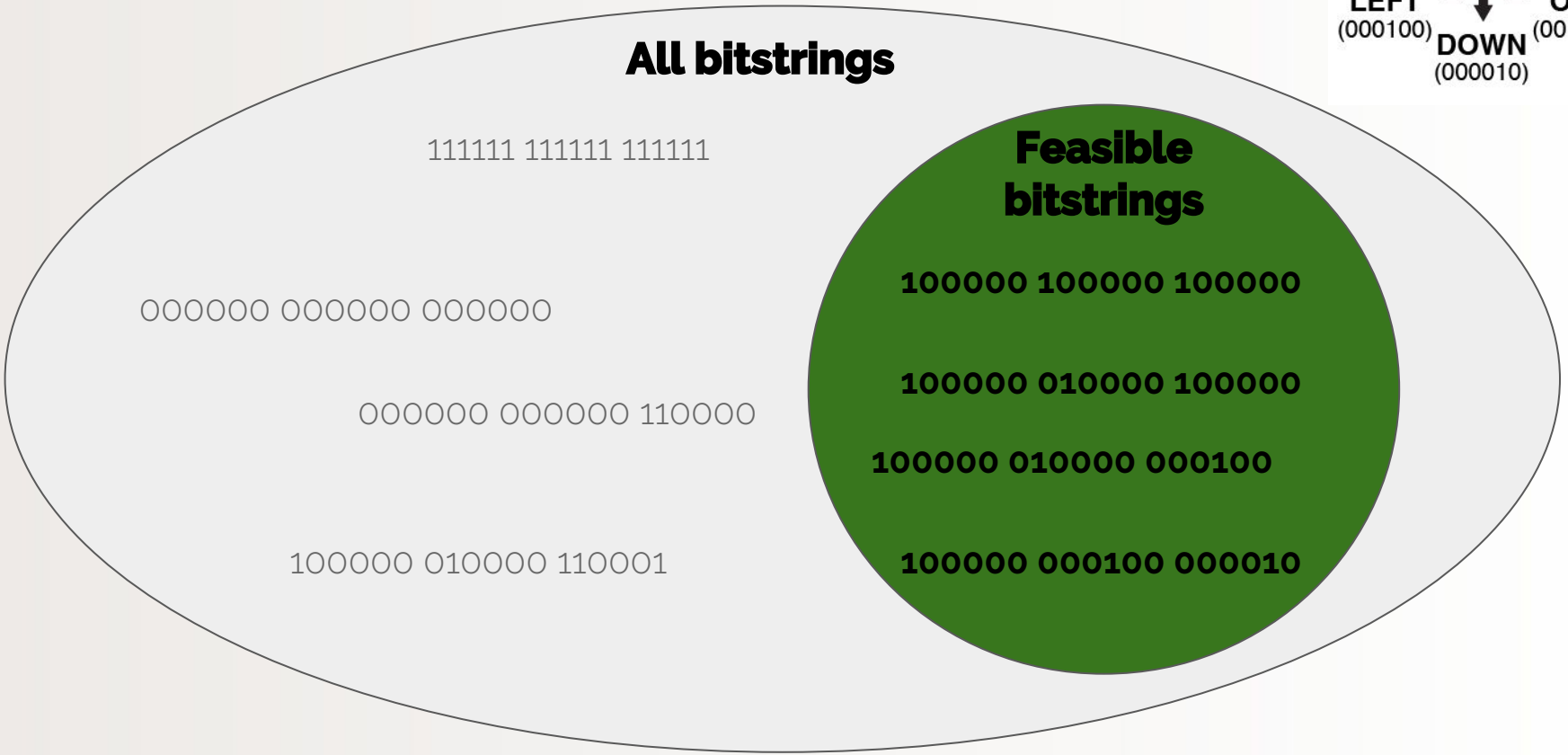
of valid solutions.




Babej, Fingerhuth & Ing (2018). PQ Whitepaper II (on arXiv soon)

# Feasible subspace




Hadfield et al. (2017). *Proceedings of the Second Int. Workshop on Post Moores Era Supercomputing* (pp. 15-21). ACM.

# Making use of the QAOA Mixer Hamiltonian

ProteinQure

**Traditional X mixer**

$$H_M = \sum_i X_i,$$

'scrambles all bits'

**Proposed SWAP mixer**

$$\mathrm{H}_M = \texttt{SWAP}_{simple} = \sum_{t=1}^{N-2} \mathrm{M}_t,$$

preserves Hamming weight!
> ensures that we stay in feasible subspace

where

$$\mathrm{M}_t = \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} \texttt{SWAP}_{i,j}.$$

and

$$\texttt{SWAP}_{i,j} = \frac{1}{2}(X_i X_j + Y_i Y_j).$$

Babej, Fingerhuth & Ing (2018). PQ Whitepaper II (on arXiv soon)

# Making use of the QAOA Mixer Hamiltonian

ProteinQure

**Short range overlaps**

$$H_M = \texttt{SWAP}_{short} = \sum_{t=1}^{N-2} \sum_{k \neq k_n}^{n-1} M_{t,k,k_n},$$

where

$$M_{t,k,k_n} = \frac{1}{4} \texttt{SWAP}(q_{m,k}, q_{m,k_n}) \prod_{j=0}^{1} (\mathbb{1} + Z_{m+(-1)^j, \bar{k}_n})$$

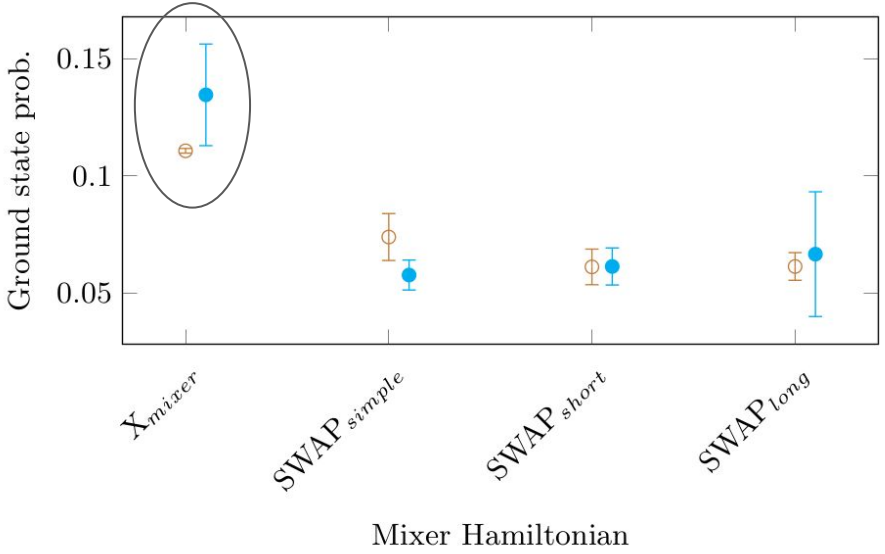
**Long range overlaps**

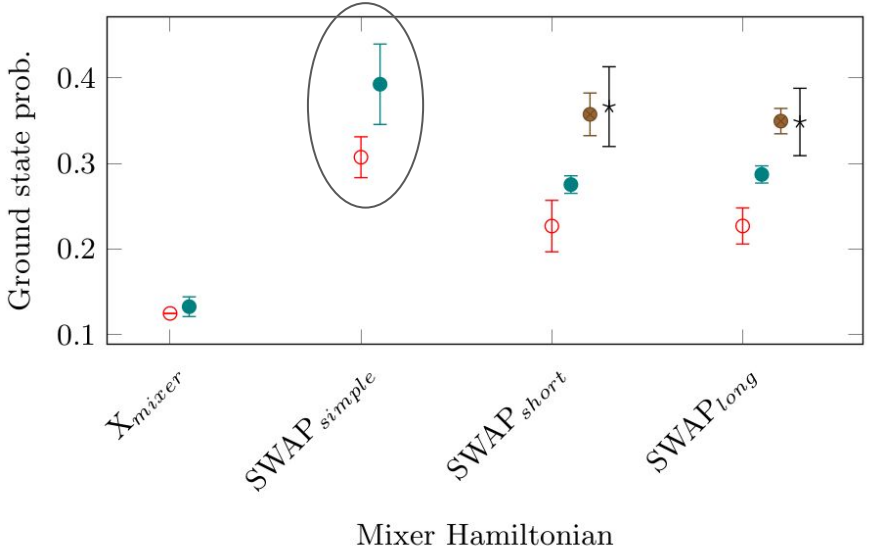
$$M_{t,k,k_n} \prod_{i=0}^{N-5} \left(1 + (D_{i,t+1} - 1)[(i-t) \bmod 2]\right).$$

Babej, Fingerhuth & Ing (2018). PQ Whitepaper II (on arXiv soon)

# Simulation results

ProteinQure



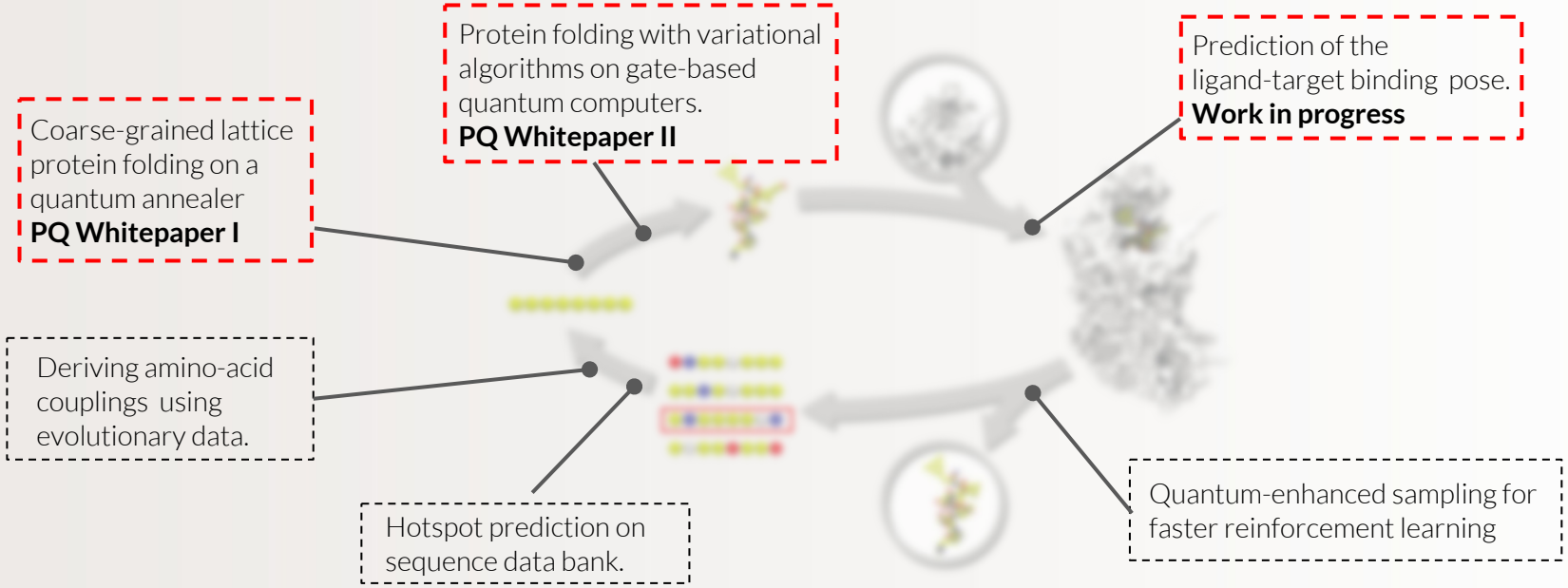

Initial state:
uniform superpos. over **all** bitstrings

Initial state:
uniform superpos. over **all feasible** bitstrings

Babej, Fingerhuth & Ing (2018). PQ Whitepaper II (on arXiv soon)

# Plenty of hard optimization problems in protein drug design

ProteinQure




**Hybrid methods will lead to speedups and the ability to scale to new problem sizes in the future.**

# Interested in collaborating?

ProteinQure

**Reach out to us if you're interested in:**

- Quantum algorithm development
- Hybrid quantum/classical algorithms on NISQ devices
- Life science applications of quantum computing
- Programming on all the different hardware platforms

# ProteinQure

## Computational drug discovery employing machine learning and quantum computing

## Contact us!

mark@proteinqure.com

Or visit our website @ www.proteinqure.com

# Deeper circuits depths are needed

| $H_M$ | Qubits | CNOT depth | Total depth | Qubits | CNOT depth on 19Q-A | Total depth on 19Q-A |
|---|---|---|---|---|---|---|
| $X_{mixer}$ | 6 | 104 | 116 | 11 | 243 | 608 |
| $\texttt{SWAP}_{simple}$ | 6 | 116 | 140 | 10 | 265 | 651 |
| $\texttt{SWAP}_{short}$ | 6 | 132 | 186 | 11 | 392 | 921 |
| $\texttt{SWAP}_{long}$ | 6 | 132 | 186 | 11 | 392 | 916 |
| $X_{mixer} + I$ | 6 | 111 | 146 | 11 | 329 | 820 |
| $\texttt{SWAP}_{simple} + I$ | 6 | 123 | 170 | 11 | 342 | 842 |
| $\texttt{SWAP}_{short} + I$ | 6 | 146 | 220 | 11 | 348 | 846 |
| $\texttt{SWAP}_{long} + I$ | 6 | 146 | 220 | 11 | 348 | 833 |

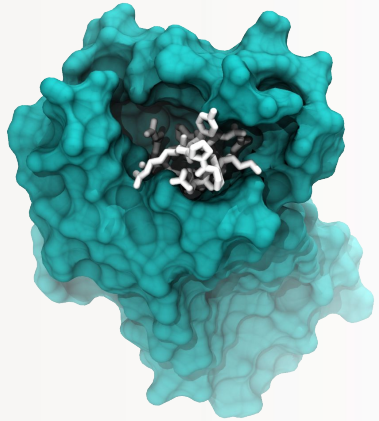
Realistic CNOT depths right now:

- Rigetti: 10.
- Google: 40.

Folded the mini-protein PSVK using Rigetti's 19Q-Acorn by breaking the problem into subproblems.

Babej, Fingerhuth & Ing (2018). PQ Whitepaper II (on arXiv soon)

# ProteinQure's Story & Roadmap



| Sep 2017 | Oct 2017 | Dec 2017 | Jan - May 2018 | June 2018 |
|---|---|---|---|---|
| Ideation of ProteinQure | Incorporation of ProteinQure Inc.<br><br>US$ 80k pre-seed | New record in lattice protein folding on quantum computer: 8 amino acids in 3D | Quantum binding<br><br>QAOA research<br><br>Classical tech dev | US$ 1.5M seed round<br><br>> 2 years of runway to build the product + 8 employees |