From Quarks to Drugs

Pietro Faccioli

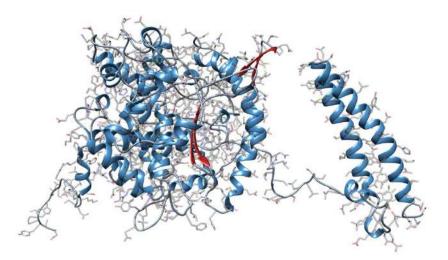


università degli studi di trento

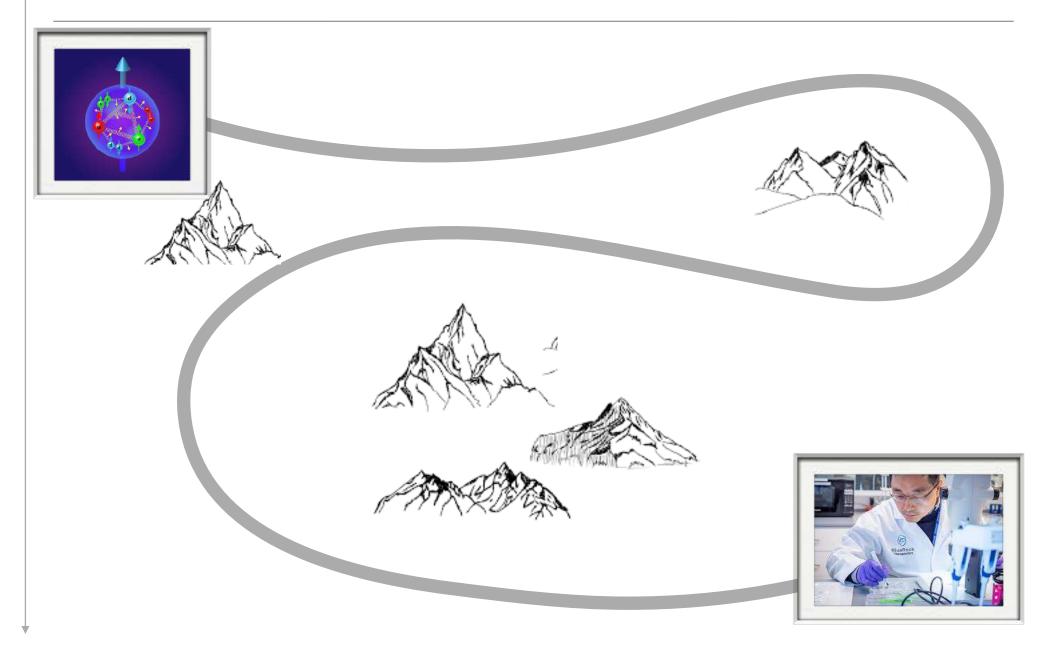
Dipartimento di Fisica



Trento Institute for Fundamental Physics and Applications



A SCIENTIFIC JOURNEY



Fundamental







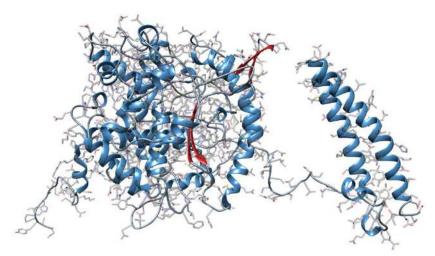
Biology



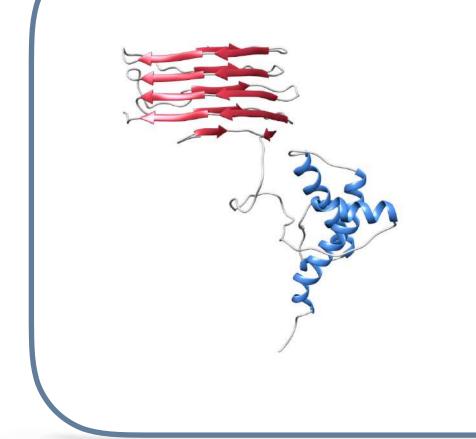




Prologue: proteins are complex many-body systems.



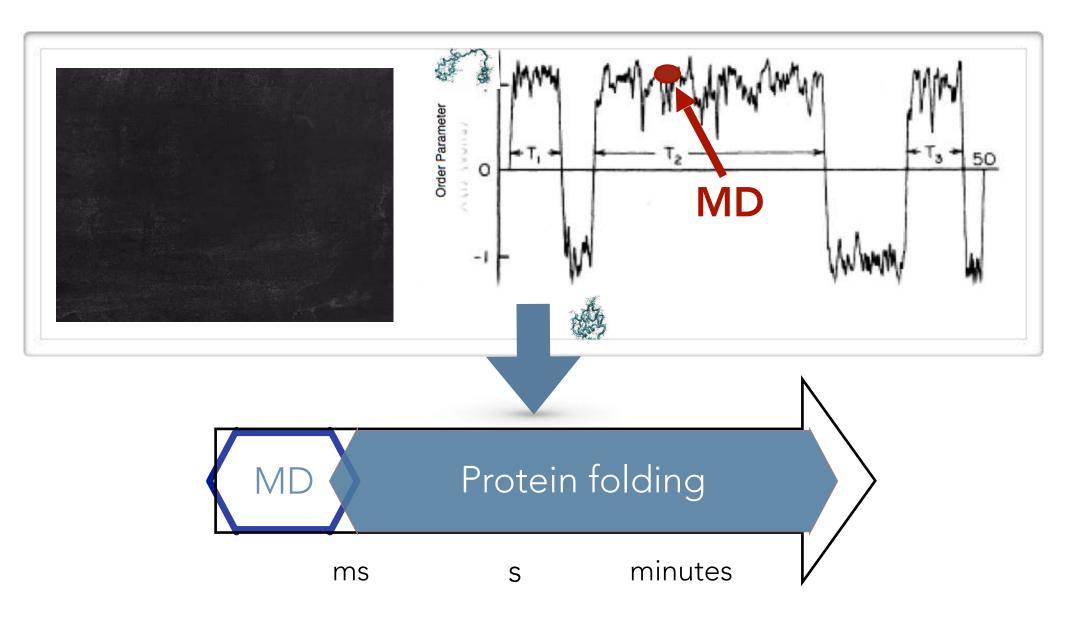
REDUCTIONIST'S APPROACH TO MOLECULAR BIOLOGY



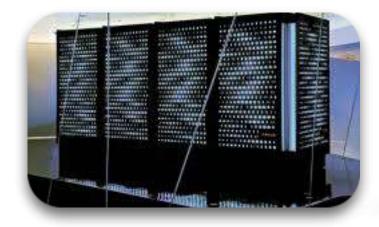
Challenge:

Integrate ~10⁶ coupled Newton-type equations looking for **extremely rare events**

RARE EVENT PROBLEMS

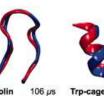


MD YIELDS CORRECT PROTEIN NATIVE STATES



Anton supercomputer (DES Research)

MD











Chignolin 106 µs cin025 1.0 Å 0.6 µs

208 µs BBA 2JOF 1.4 Å 14 µs 1FME 1.6 Å 18 µs

325 µs Villin 2F4K 1.3 Å 2.8 us







WW domain 1137 us 2F21 1.2 Å 21 US

NTL9 2936 µs 2HBA 0.5 Å 29 µs BBL 429 μs 2WXC 4.8 Å 29 μs

429 µs Protein B 1PRB 3.3 Å 3.9 µs











Homeodomain 327 µs Protein G 1154 µs 2P6J 3.6 Å 3.1 µs 1M/O 1.2 Å 65 µs

α3D 707 μs 2A3D 3.1 Å 27 μs

A-repressor 643 μs 1LMB 1.8 Å 49 µs



Atomic-Level Characterization of the Structural Dynamics of Proteins David E. Shaw, *et al. Science* **330**, 341 (2010); DOI: 10.1126/science.1187409

How Fast-Folding Proteins Fold

Kresten Lindorff-Larsen, 1*+ Stefano Piana, 1*+ Ron O. Dror, 1 David E. Shaw1,2+

ZOOLOGY OF ENHANCED SAMPLING METHODS

Markov State Models (Folding@Home), Milestoning, Transition Path Sampling, Transition Interface Sampling, Forward Flux Sampling, Temperature Accelerated Molecular Dynamics, Metadynamics, Umbrella Sampling, Blue Moon Sampling, String Method,Stochastic Difference, ... [and counting]

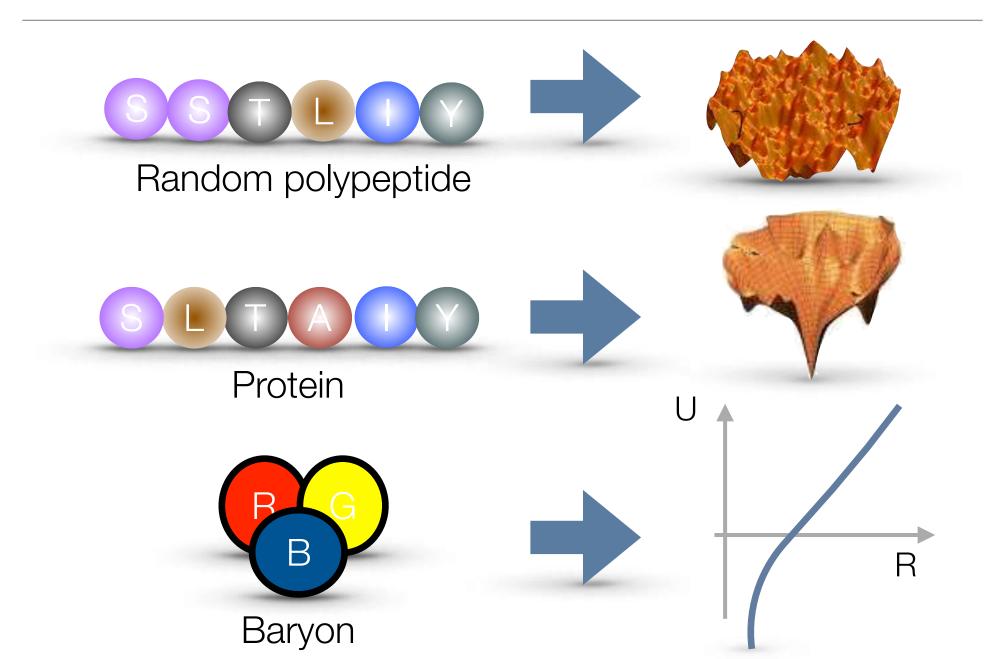
They are **all too computationally demanding** for many biologically relevant problems.



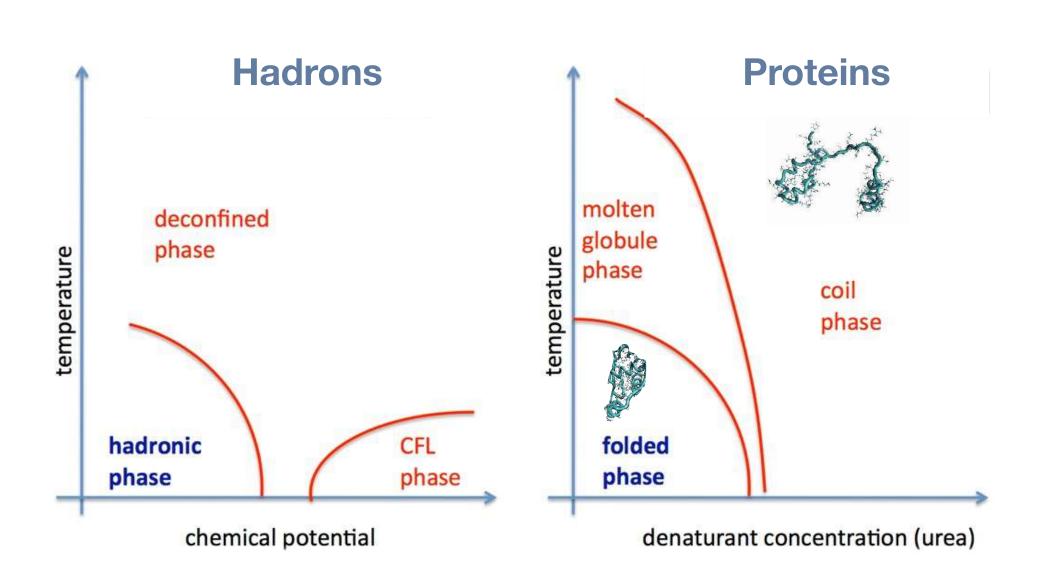




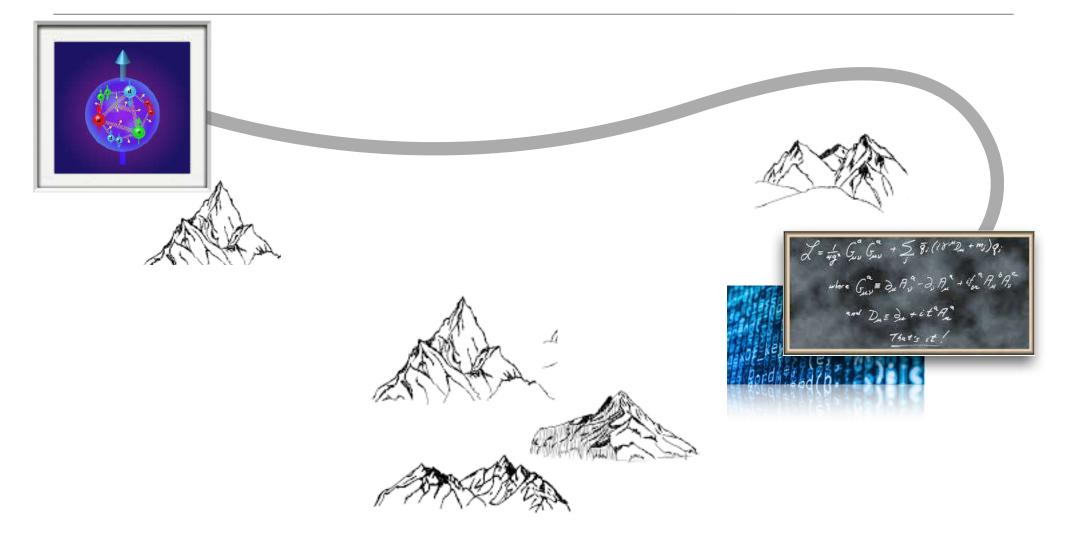
PROTEINS AND HADRONS ARE VERY SPECIAL



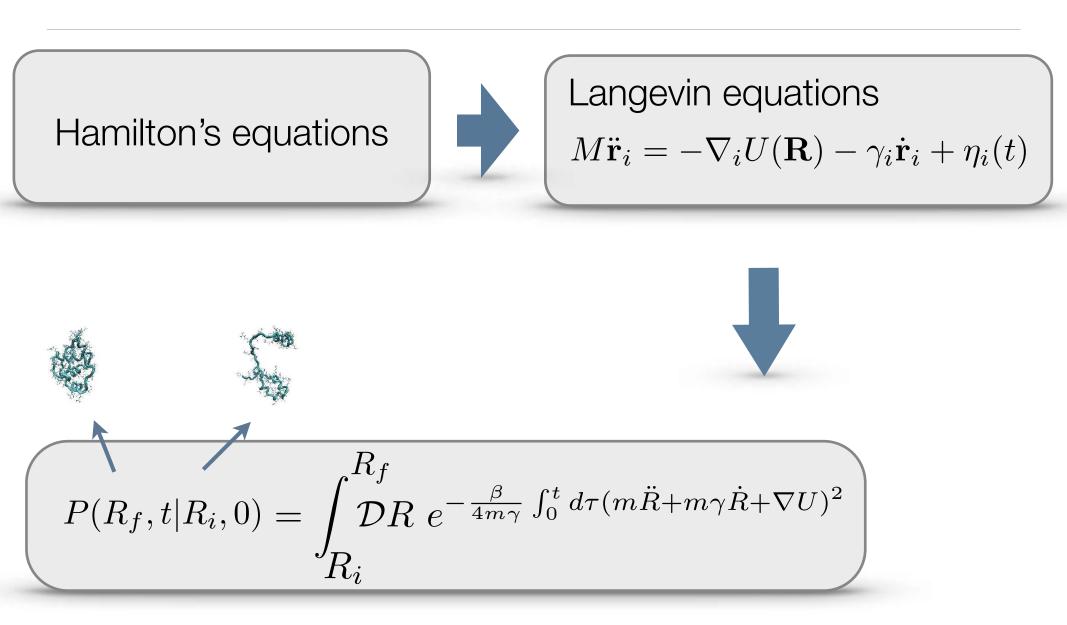
PHASE DIAGRAM



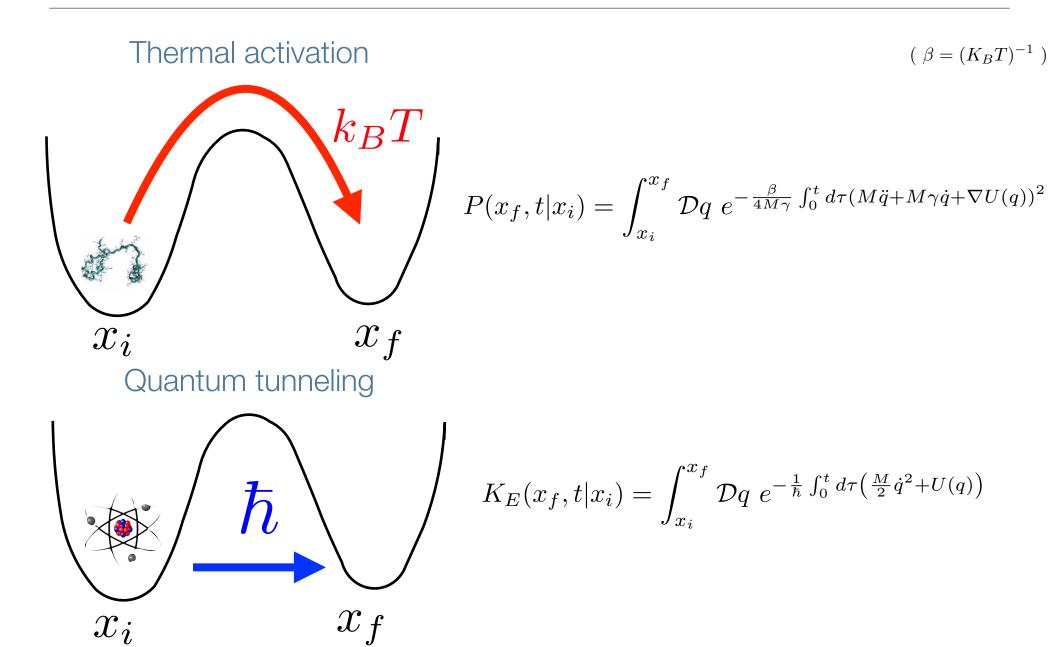
PHASE 1: MATHEMATICAL FORMALISM & HIGH PERFORMANCE COMPUTING



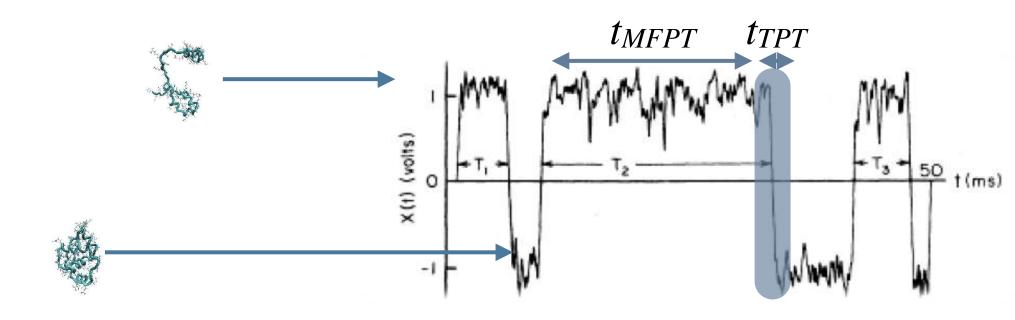
PATH INTEGRAL REPRESENTATION



A USEFUL ANALOGY

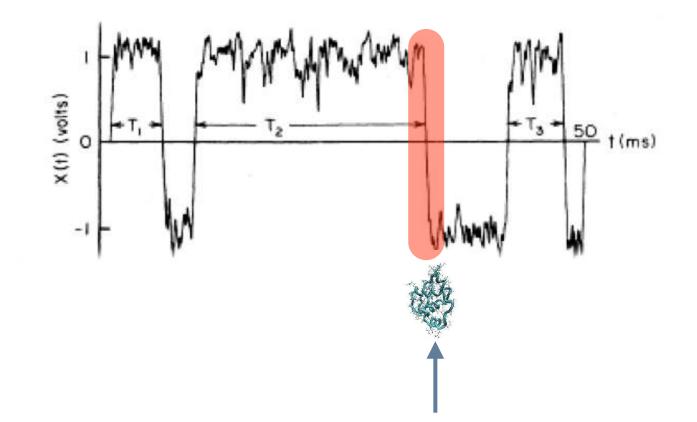






$$t_{TPT} \sim \tau_0 \log \left[\log \left(\frac{t_{MFPT}}{\tau_0} \right) \right]$$

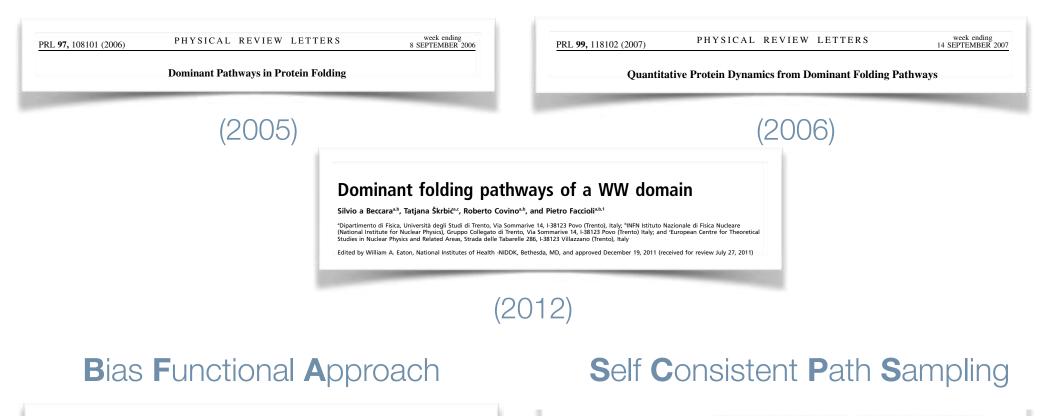
IS THIS A "FREE LUNCH"?



All atom 3D structure of the native state **are given in input,** not predicted

VARIATIONAL APPROACHES TO TRANSITION PATH SAMPLING

Dominant Reaction Pathways



week endin

6 MARCH 20

PRL 114, 098103 (2015)

PHYSICAL REVIEW LETTERS

Variational Scheme to Compute Protein Reaction Pathways Using Atomistic Force Fields with Explicit Solvent THE JOURNAL OF CHEMICAL PHYSICS 147, 064108 (2017)

Self-consistent calculation of protein folding pathways

S. Orioli. S. a Beccara. and P. Facciolia)

FULLY EXPLOITING THEORETICAL PHYSICS TOOLS

PHYSICAL REVIEW LETT

(3)

(4)

δ

ōX.

This equation s

that for which I

force is least. In derived in the emphasize that

solvent-induced Let us now e history-depend

the ratchet-and

formalism

attempts

terms of s

z. Conver

spontaneo

(1) with a

To defi

-41

0

 $z_m(t)$ den

time t (we

obeys the

Let us

rithm developed in

072336-4 Bartolucci, Orioli, and Faccioli

between the Gibbs distribution and the SCR estimate forwardand backward-committors, as in Eq. (A3). Introducing the distribution

$$P^{(P)}(x,t) \equiv \int dx_i \ P^{(P)}(x,t|x_i,0) \ \rho_0(x_i),$$

the density in Eq. (22) reads

$$m_{SCR}(x) = \frac{1}{t_f - \tau_0} \int_{\tau_0}^{t_f} dt \ Q^{(R)}(x, t_f - t) P^{(P)}(x, t).$$

Using the detailed balance condition, we find $P^{(P)} = e^{-\beta U(x)} \frac{1}{Z_R} Q^{(P)}(x, t)$. Then, inserting this result into Eq. we find

$$m_{SCR}(x) = \frac{e^{-\beta U(x)}}{Z_R(t_f - \tau_0)} \int_{\tau_0}^{t_f} dt Q^{(R)}(x, t_f - t) Q^{(P)}(x, t).$$

Finally, recalling that $Q^{(R)}(x, t)$ and $Q^{(P)}(x, t)$ are n time-independent in the SCR and using Eqs. (17) and we recover a fundamental result of TPT [cf. Eq. (A Appendix A],

$$m_{SCR}(x) \propto e^{-\beta U(x)} q_{SCR}^+(x) (1 - q_{SCR}^+(x)).$$

Within the same framework, it is possible to d the reactive current in the SCR in complete analogy Eq. (22),

$$\begin{split} J^{i}_{SCR}(x) &= \frac{-D}{t_f - \tau_0} \int_{\tau_0}^{t_f} dt Q^{(R)}(x, t_f - t) \\ &\times (\vec{\nabla} - \overleftarrow{\nabla} + \beta \nabla U(x)) \, P^{(P)}(x, t). \end{split}$$

$$\begin{aligned} V_{eff}^{R}(\mathbf{X}) &\simeq \frac{D_{0}\left(1-b\right)}{\pi b\Omega} \nabla^{2} V_{eff}(\mathbf{X}) \\ &+ \frac{1}{2} \left(\frac{D_{0}(1-b)}{\pi b\Omega}\right)^{2} \nabla^{4} V_{eff}(\mathbf{X}) \\ &+ \frac{1}{6} \left(\frac{D_{0}\left(1-b\right)}{\pi b\Omega}\right)^{3} \nabla^{6} V_{eff}(\mathbf{X}) - \frac{D_{0}^{2}(1-b^{3})}{3\pi \left(b\Omega\right)^{3}} \left(\partial_{i}\partial_{j}V_{eff}(\mathbf{X})\right)^{2}. \end{aligned}$$
(24)

(72)

PRL 114, 098103 (2015)

 $\mathcal{P}_{\text{bias}}[X] = \int \mathcal{D}Y e^{-S_{\text{bias}}[X,Y] - U(X_i,Y_i)/k_{\delta}T}$

 $S_{\text{bias}} \equiv \frac{1}{4k_{\pi}T} \int_{0}^{T} d\mathbf{r} \left[\sum_{\mathbf{\hat{\gamma}}:m_{i}}^{N} (m_{i} \mathbf{\hat{x}}_{i} + m_{i} \mathbf{\hat{\gamma}}_{i} \mathbf{\hat{x}}_{i} + \nabla_{i} U - \mathbf{F}_{i}^{\text{bias}})^{2} \right]$

The Onsager-Machlup functional $S_{OM}[X, Y]$ entering

Let us now return to the problem of computing the

 $\frac{\delta}{s_{Y}} \left[\mathcal{P}_{\text{bias}}[X] \langle e^{-(S_{\text{OM}}[X,Y] - S_{\text{bias}}[X,Y;t])} \rangle_{\text{bias}} \right] = 0. \quad (5)$

reaction pathways in the unbiased Langevin dynamics

[Eq. (1)]. Using the standard reweighting trick we can

We now introduce our main approximation, by restrict-

ing the search for the optimum path X(r) within an

ensemble of trajectories generated by integrating the biased

Langevin equation. By definition, these paths have a large

statistical weight in the biased dynamics; i.e., they lie in the

functional vicinity of some path $\tilde{X}(\tau)$ which satisfies

 $(\delta/\delta \bar{X})\mathcal{P}[\bar{X}] = 0$. Thus, the typical biased paths approx-

where the functional $S_{\text{bias}}[X, Y]$ is defined as

 $+\sum_{i=1}^{N'}\frac{1}{\gamma_{i}m_{i}}(m_{j}\dot{\mathbf{y}}_{i}+m_{j}\gamma_{j}\dot{\mathbf{y}}_{j}+\nabla_{j}U)^{2}\Big].$

Eq. (2) is recovered, setting $\mathbf{F}_{i}^{\text{bias}} = 0$ in Eq. (4).

write the variational condition $(\delta/\delta X)\mathcal{P}[X] = 0$ as

Note that the first line is the leading order term (i.e. L = 1), while the second and third lines display the order L = 2 and L = 3 corrections, respectively.

We emphasize that the result of the EST construction is a new expression for the *same* path integral (15), in which the UV cutoff been lowered from Ω to $b\Omega$. Equivalently, the path integral is discretized according to a larger elementary time step, $\Delta t \rightarrow \Delta t/b$:

$$Z^{\Delta t}(t) \equiv \oint_{\Delta t} \mathcal{D}\mathbf{X} \ e^{-S_{eff}[\mathbf{X}]} \propto \oint_{\Delta t/b} \mathcal{D}\mathbf{X} \ e^{-S_{eff}[\mathbf{X}] - \int_0^t d\tau \ V_{eff}^R[\mathbf{X}(\tau)]} \equiv Z_{EST}^{\Delta t/b}(t)$$
(25)

In these expressions, the symbol $f_{\Delta t}$ denotes the fact that the path integral is discretized according to an elementary time step Δt and we have suppressed the subscript "<", in the paths. It can be shown that the proportionality factor between $Z^{\Delta t}(t)$ and $\frac{\Delta t}{t}$

$$(v_2) \quad v_2 \quad = \quad (v_2) \quad v_2 \quad v_$$

FIG. 3: Diagrammatic representation of the local time-derivative expansion of a non-local diagram -Eq. (49)-. Solid lines are fast-mode propagators, while dashed lines represent a single time derivative acting on the corresponding vertex function.

Notice that each term in the perturbative expansion (35) generates a new vertex, with an increasing power of the $x_s(\tau)$ field. The couplings to the fast modes depend implicitly on the time τ_s through the slow modes $x_s(\tau)$. By Witk theorem coch term in the arrive (20) can be noticed to a feature mode with vertexer given by (26) and

by the most in, the relation in the set
$$S(0)$$
 can be related to a regiminal graph with vertices green by (60) and propagators given by — see appendix A — $\frac{2}{2}$

$$\langle x_{>}^{i}(\tau_{1}) \ x_{>}^{j}(\tau_{2}) \rangle_{0} = \sum_{|\omega_{m}|,|\omega_{n}| \in S_{k}} G_{>}^{0\ ij}(\omega_{n},\omega_{m}) \ e^{i(\omega_{m}\tau_{1}+\omega_{n}\tau_{2})} = \sum_{|\omega_{n}| \in S_{k}} \delta_{ij} \ \frac{2}{\beta\ \gamma\ t\ \omega_{n}^{2}} \ e^{i\omega_{m}(\tau_{2}-\tau_{1})}. \tag{37}$$

The expansion (34) can be re-organized as the exponent of the sum performed over only connected diagrams:

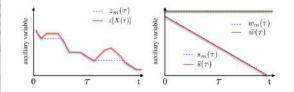
$$e^{-\beta S_{\geq}[x < (\tau)]} = e^{\sum (\text{all connected diagrams})}$$
. (38)

Hence, the path integral (26) for the slow modes can be given the following exact diagrammatic representation

$$Z(t) \equiv \oint \mathcal{D}x_{\leq} e^{-\beta S_{eff}[x_{\leq}(t)]] + \sum (\text{all connected diagrams})}.$$
 (39)

J. Chem. Phys. 147, 064108 (2017)

FIG. 1. Illustrative representation of



the dynamics of the auxiliary variables introduced in the path integral representation of rMD (left panel) and in the derivation self-consistent path sampling algorithm (right panel).

III. SELF-CONSISTENT PATH SAMPLING

of such a variable is frozen any time z_m becomes smaller than z(X) and any time the collective coordinate z(X) is increasing. Its time derivative is otherwise set equal to z(X). Therefore, by choosing the initial conditions $z_m(0) = z(X(0))$, $z_m(\tau)$ is identically set equal to the minimum value attained by the collective coordinate z until time τ (see left panel of Fig. 1). The functional S = U(X - L) in the suprement of Eq. (8) with

The functional $S_{rMD}[X, z_m]$ in the exponent of Eq. (8) coincides with an OM action with the addition of the unphysical biasing force \mathbf{F}_{i*}

$$S_{rMD} = \sum_{i=1}^{N} \Gamma_i \int_0^t d\tau \left[m_i \ddot{\mathbf{x}}_i + m_i \gamma_i \dot{\mathbf{x}}_i + \nabla_i U - \mathbf{F}_i \right]^2, \quad (9)$$

In Eq. (8), $\Phi[z_m, X]$ denotes a Jacobian factor that needs to be introduced in order to ensure that the statistical weight of the paths is not affected by the measure of the $\int Dz_m$ integral, i.e.

$$\int \mathcal{D} z_m \Phi[z_m, X] \,\delta\left[z_m(\tau) - \int_0^\tau d\tau' \dot{z}[X(\tau')]\theta(-\dot{z}[X(\tau')])\right]$$

Let us now introduce our new algorithm, which provides major improvement with respect to the rMD and BF schemes discussed in Sec. II A. Indeed, it follows directly from the unbiased Langevin equation and allows us to remove the systematic errors associated to the choice of biasing coordinate.

Our starting point is path integral representation of the *unbiased* Langevin dynamics (2). We introduce two dumb auxiliary variables $w_m(\tau)$ and $s_m(\tau)$ into this path integral by means of appropriate functional Dirac deltas,

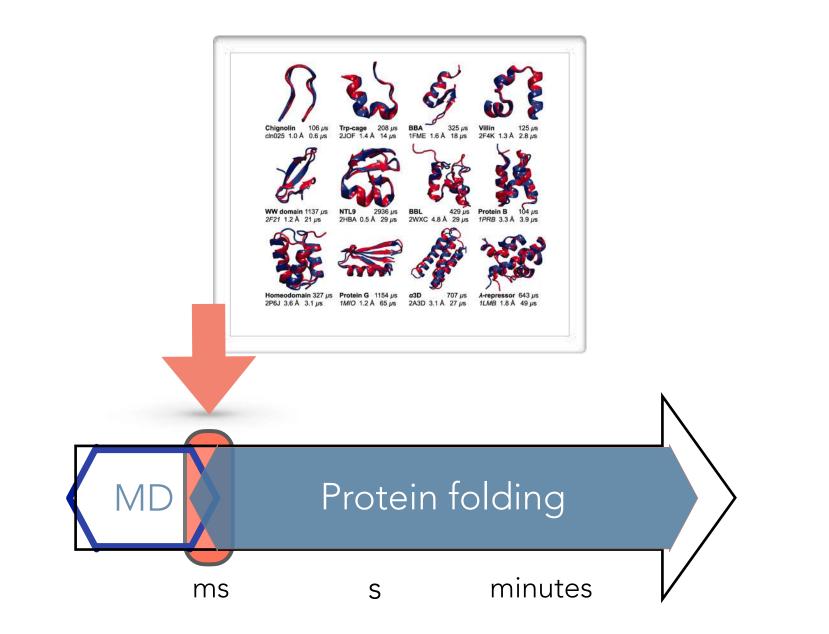
$$p(X_N, t|X_U) = \int_{X_U}^{X_N} DX \cdot e^{-\tilde{S}[X]} \int_{\mathbb{R}(0)} Ds_m \int_{\tilde{w}(0)} Dw_m$$

$$\cdot \delta \left[w_m(\tau) - \int_0^{\tau} d\tau' \dot{\tilde{w}}(\tau') \, \theta(-\tilde{\tilde{w}}(\tau')) \, \theta(w_m(\tau') - \tilde{w}(\tau')) \right]$$

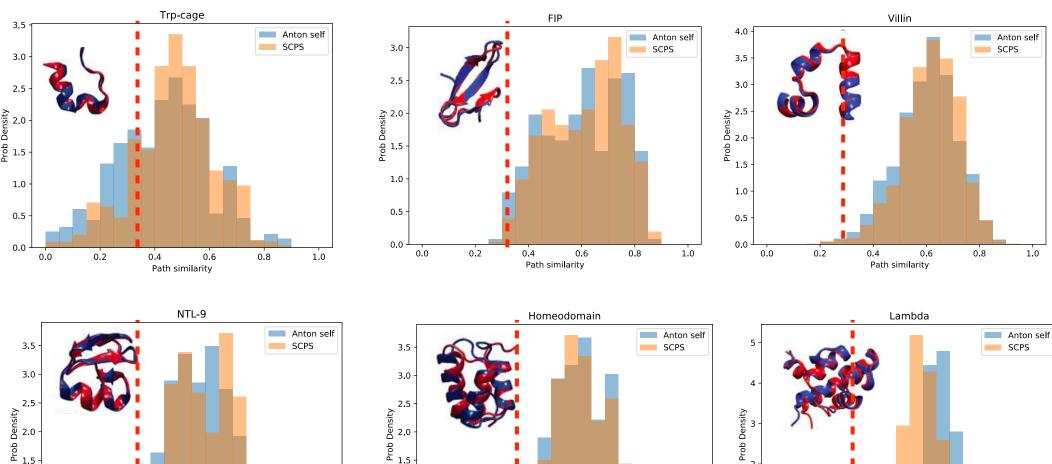
$$\cdot \delta \left[s_m(\tau) - \int_0^{\tau} d\tau' \dot{\tilde{s}}(\tau') \, \theta(-\tilde{\tilde{s}}(\tau')) \, \theta(s_m(\tau') - \tilde{s}(\tau')) \right], \quad (12)$$

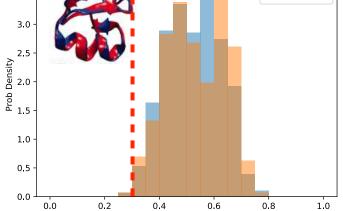
where $\tilde{s}(\tau)$ and $\tilde{w}(\tau)$ are two external time-dependent functions to be defined below. In analogy with the path integral repre-

VALIDATING SCPS AGAINST MD

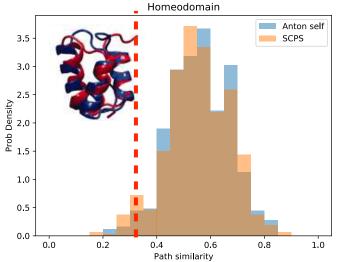


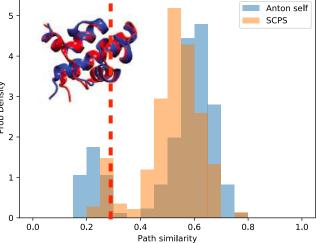
VALIDATING SCPS AGAINST MD



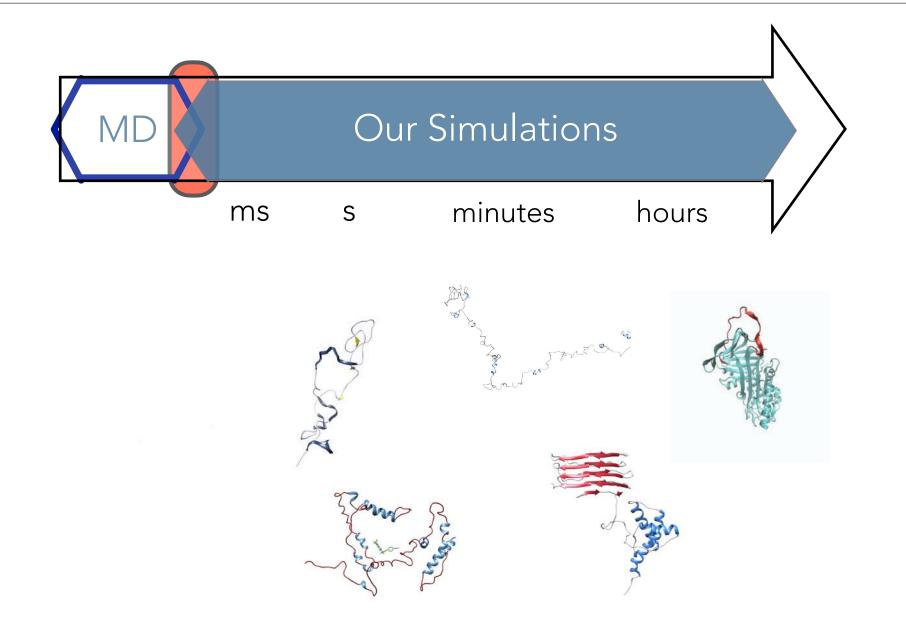


Path similarity

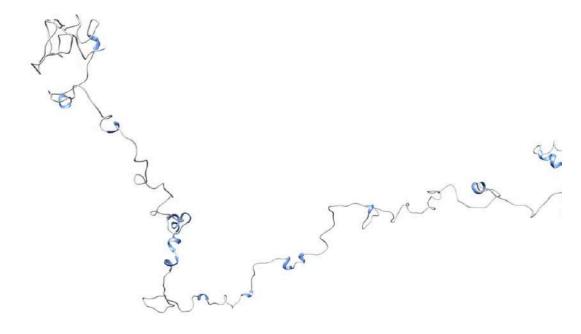


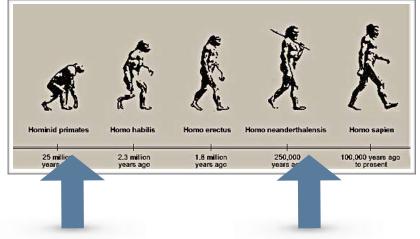


VENTURING INTO THE BIO-ZONE



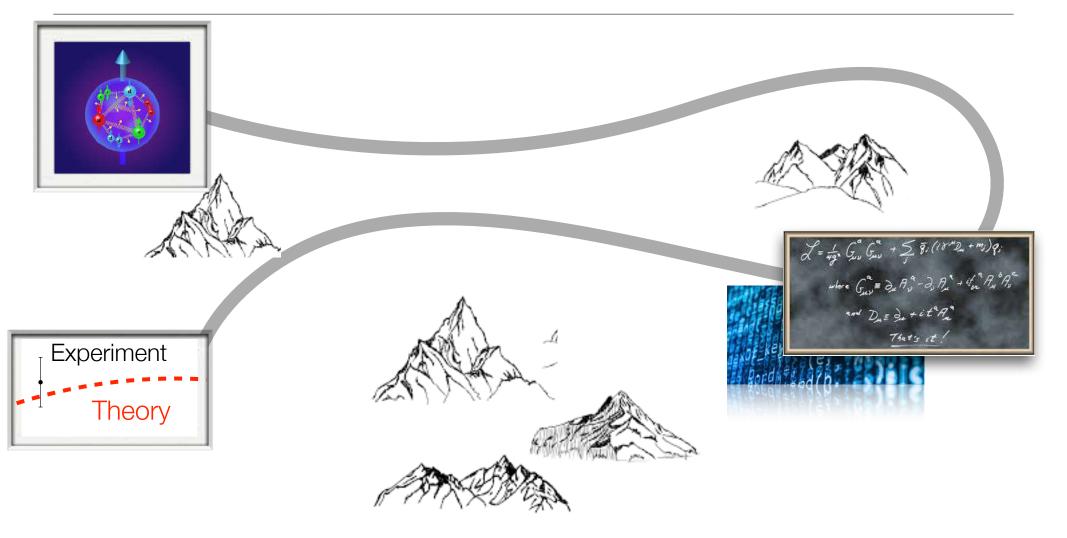
HUGE COMPUTATIONAL GAIN





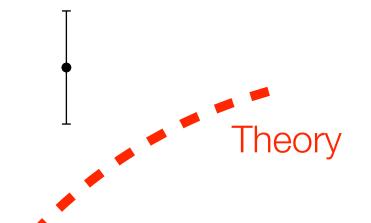
Using top allpurpose supercomputers Using top special-purpose supercomputer

PHASE 2: VALIDATION



VALIDATION AGAINST EXPERIMENT

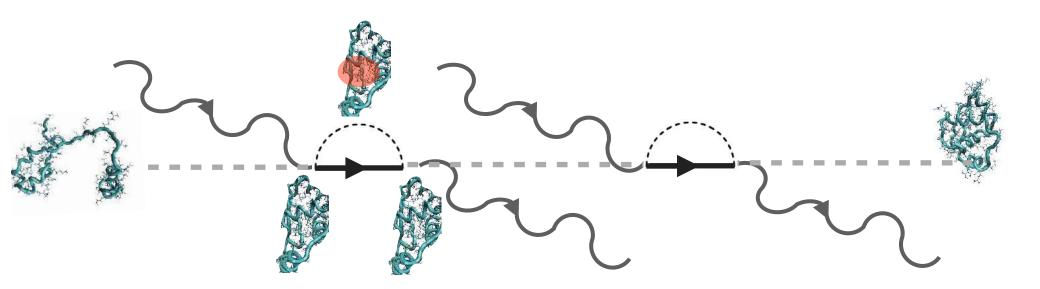
Experiment



Challenge:

Most available techniques provide only indirect probes, we seek for **direct validation**

TIME-DEPENDENT LINEAR SPECTROSCOPY

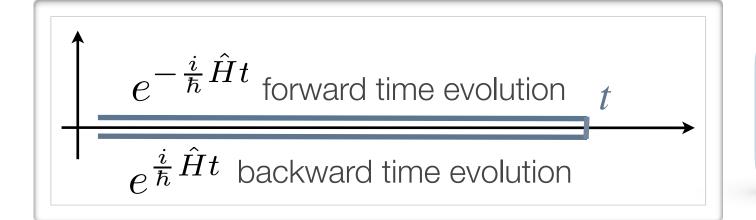


Ground stateOne exciton

Challenge:

Need a theory for **non-equilibrium dynamics** of **quantum** electronic excitations in conformationally evolving proteins

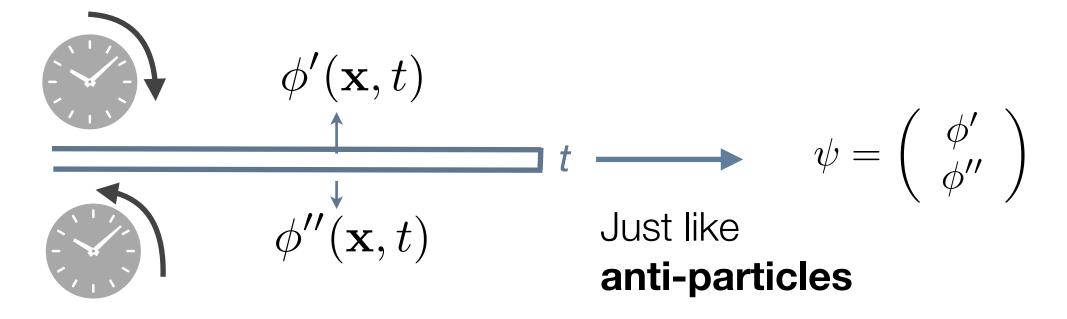
$$\hat{\rho}(t) = e^{\frac{i}{\hbar}\hat{H}t} \hat{\rho}(0) e^{-\frac{i}{\hbar}\hat{H}t}$$



multiple time directions...

USE QUANTUM FIELD THEORY!

Using QFT we get rid of the multiple time issue:

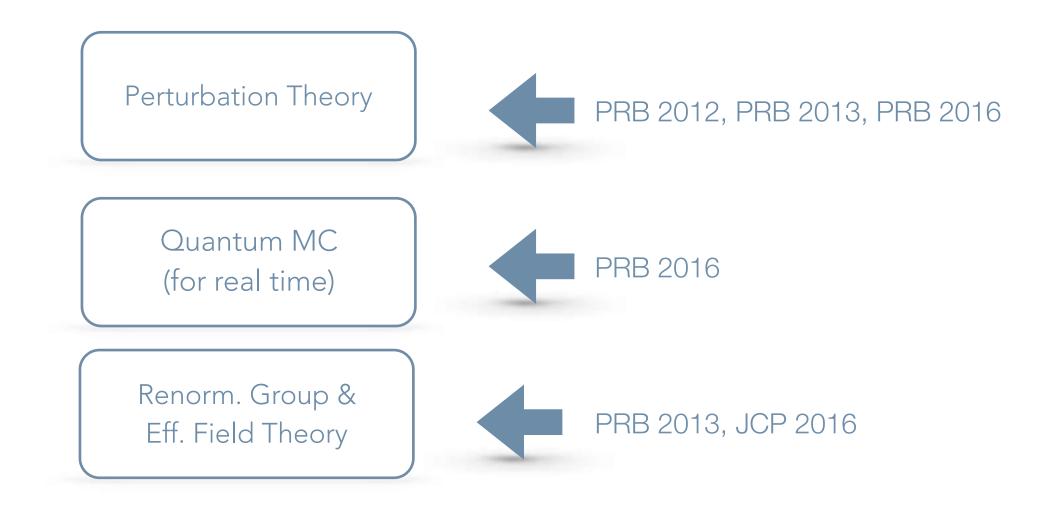


One "relativistic" field doublet but just one time

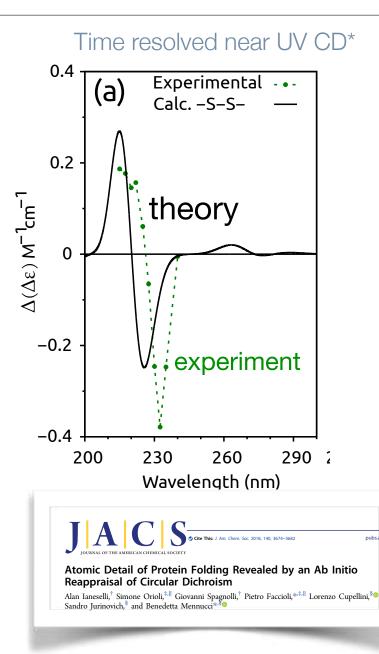
MOLECULAR QUANTUM FIELD THEORY*

P. Faccioli & E. Schneider (2013-2016)

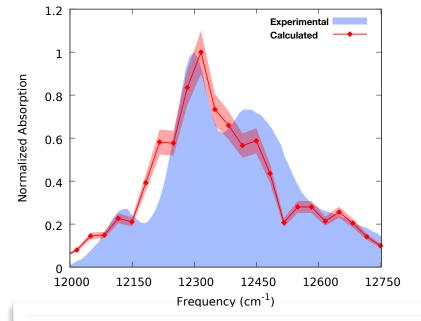
SOLVING MQFT: AN ARSENAL OF METHODS



EXAMPLES OF DIRECT COMPARISON WITH EXPERIMENTS



Linear absorption spectrum



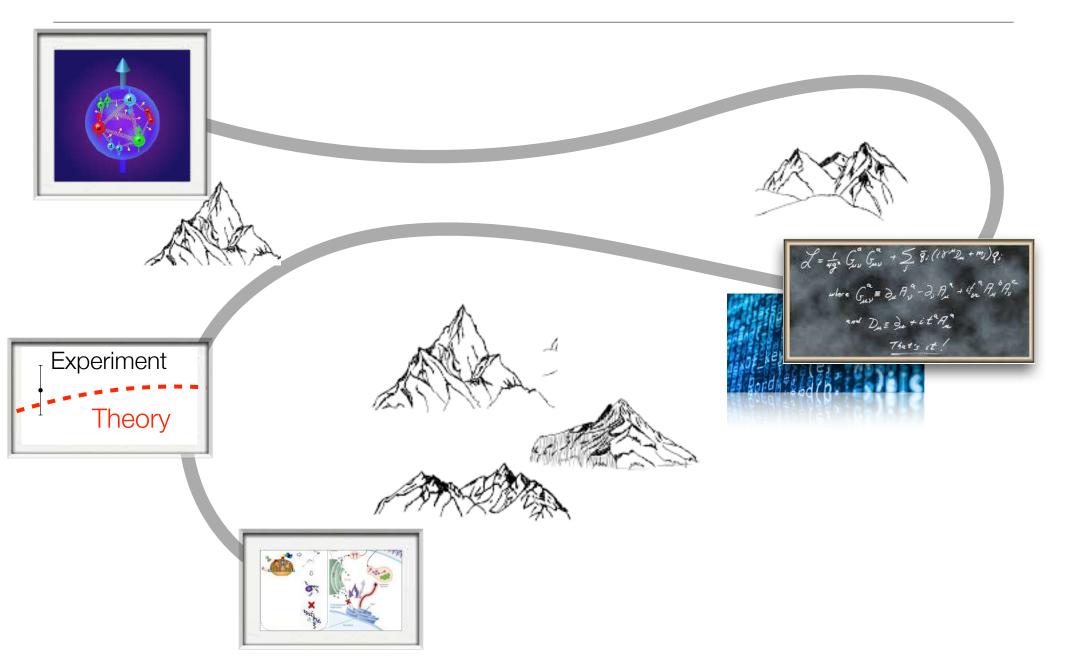
Microscopic Calculation of Absorption Spectra of Macromolecules: an Analytic Approach

Matteo Carli Physics Department of Trento University, Via Sommarive 14, Povo (Trento), 38123, Italy and Scuola Internazionale Superiore di Studi Avanzati (SISSA), via Bonomea 265, Trieste 34136, Italy

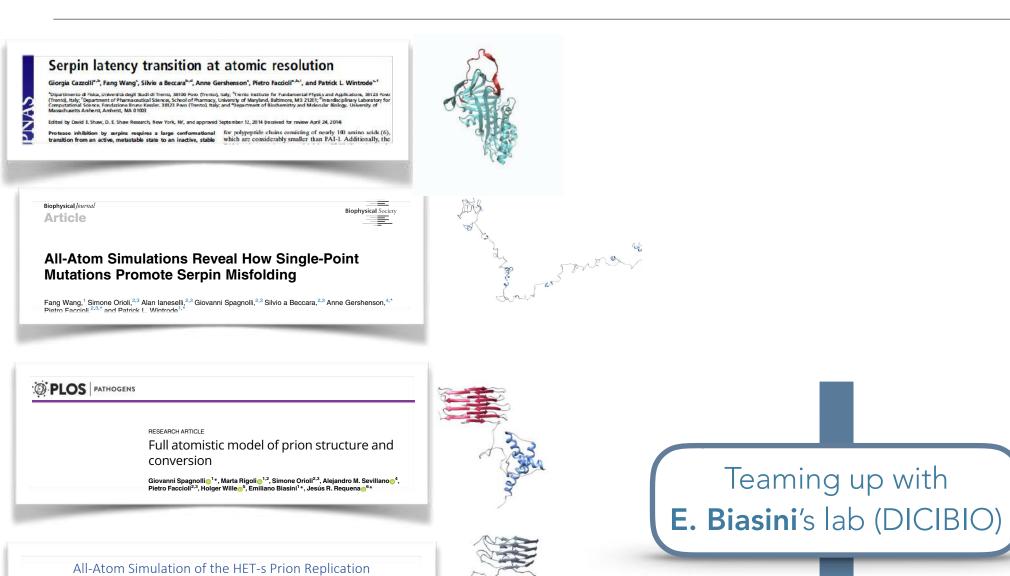
Michele Turelli and Pietro Faccioli* Physics Department of Trento University, Via Sommarive 14, Rovo (Trento), 38123, Italy and Trento Institute for Fundamental Physics and Applications (INPN-TIFPA), Via Sommarive 23, Povo (Trento), 38123, Italy

* with B. Mennucci's Lab (U. Pisa)

PHASE 3: EXPLOITATION IN MOLECULAR BIOLOGY



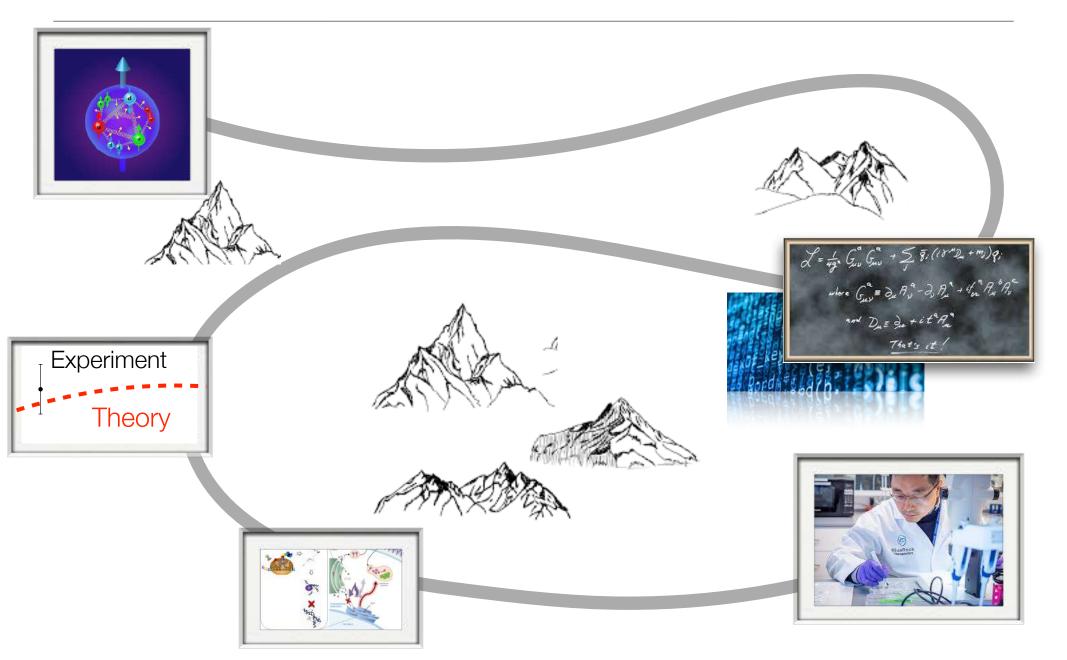
EXPLORING BIOLOGICAL PROCESSES



Luca Terruzzi^{1,2}*, Giovanni Spagnolli^{2,3}*[#], Alberto Boldrini^{1,2}, Jesús R. Requena⁴, Emiliano Biasini^{2,3#} and Pietro

Faccioli5,6#

PHASE 4: PHARMACOLOGICAL RESEARCH



ROLE OF PROTEIN INACTIVATION

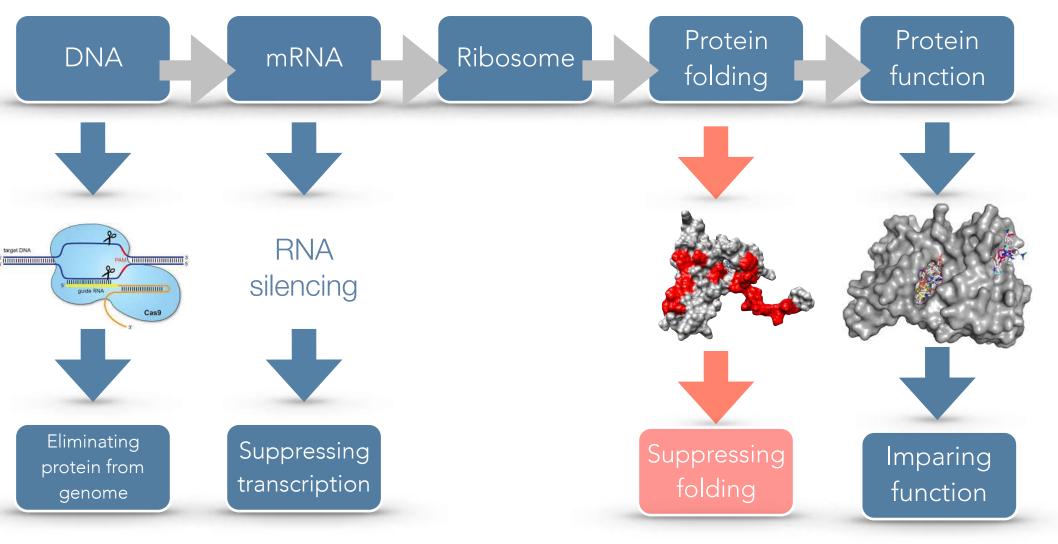
MOST OF BIOLOGICAL FUNCTIONS IN CELLS ARE CARRIED OUT BY **PROTEINS**



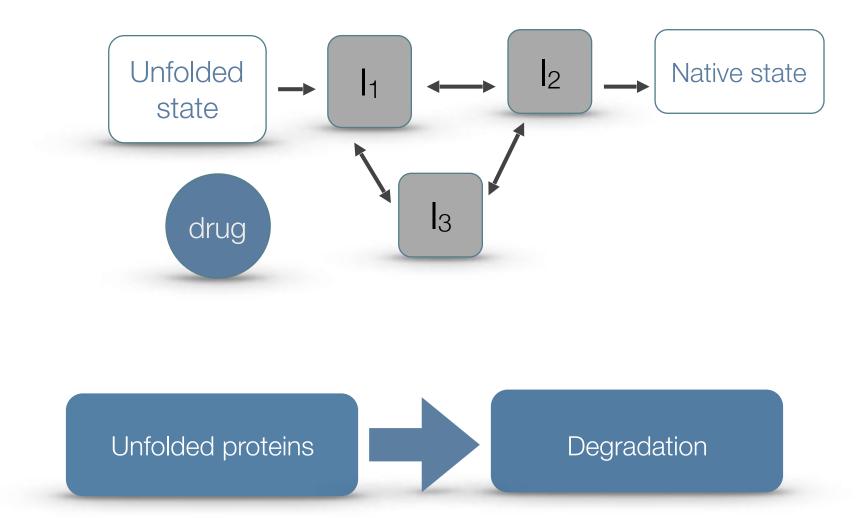
MOST OF MEDICINAL CHEMISTRY IS BASED ON INHIBITING BIOLOGICAL FUNCTIONS OF PROTEINS

PHARMACOLOGICAL PROTEIN INACTIVATION BY FOLDING INTERMEDIATE TARGETING

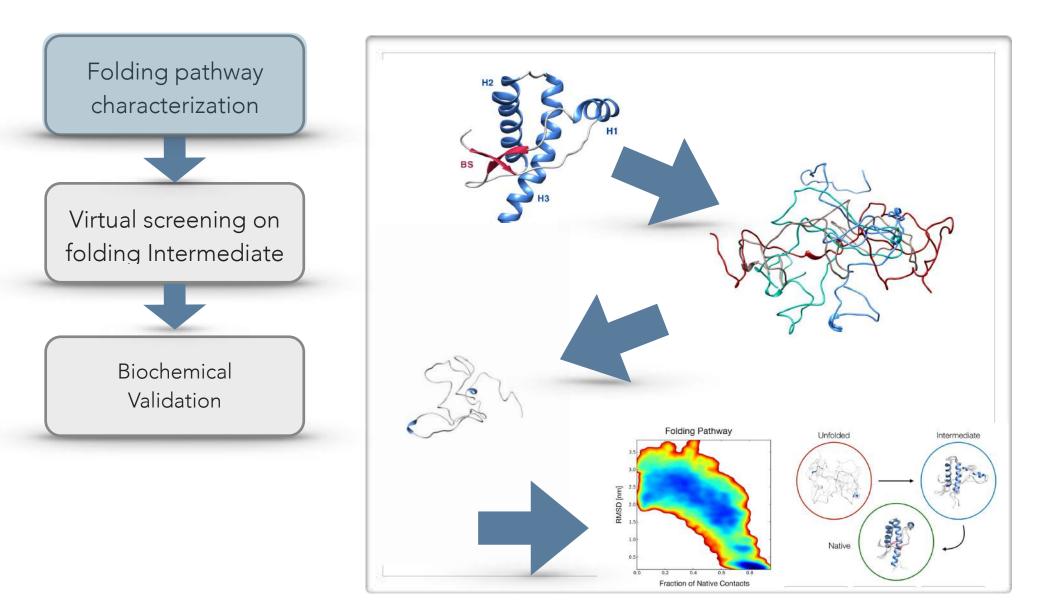
patent file # 102018000007535 (with E. Biasini)



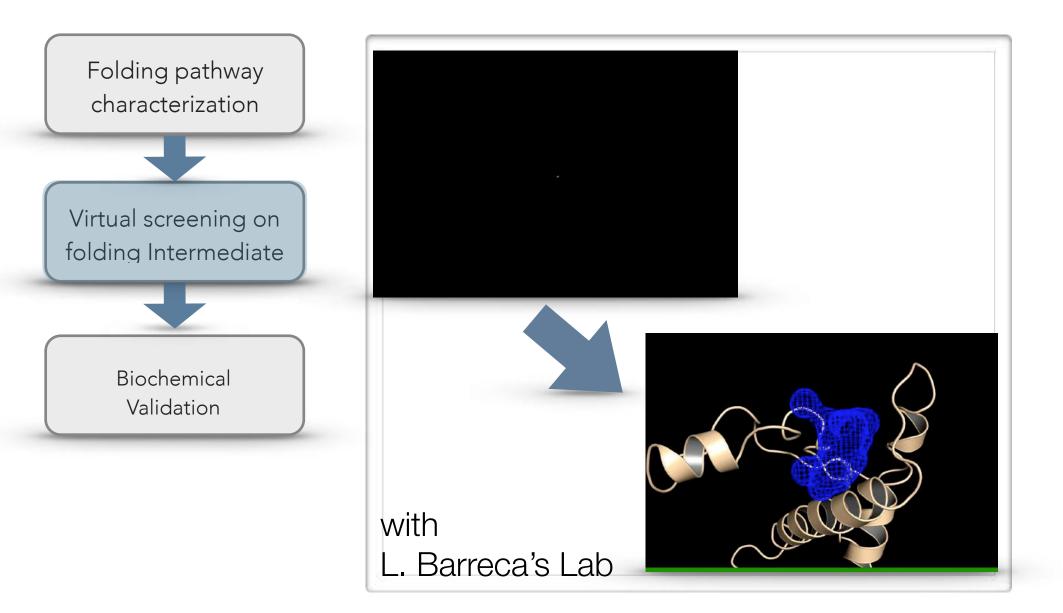
PHARMACOLOGICAL PROTEIN INACTIVATION BY FOLDING INTERMEDIATE TARGETING



PPI-FIT PIPELINE

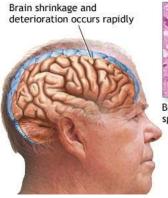


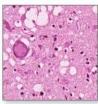
PPI-FIT PIPELINE



DRUGGING THE UNDRUGGABLE

Inactivation of Cellular Prion protein





Brain section showing spongiform pathology characteristic of Creutzfeldt-Jakob

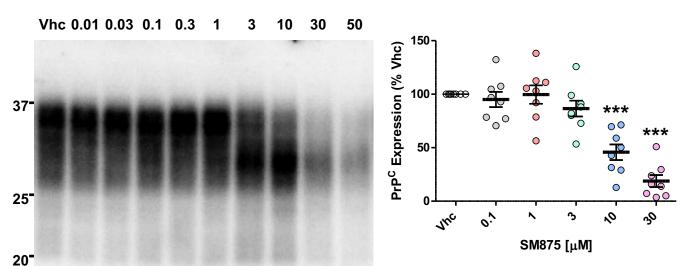
Check for updates

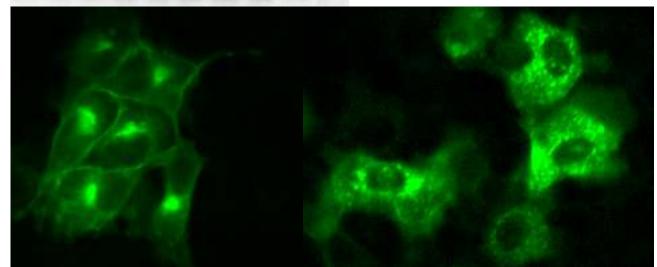
COMMUNICATIONS BIOLOGY

ARTICLE https://doi.org/10.1038/s42003-020-01585-x

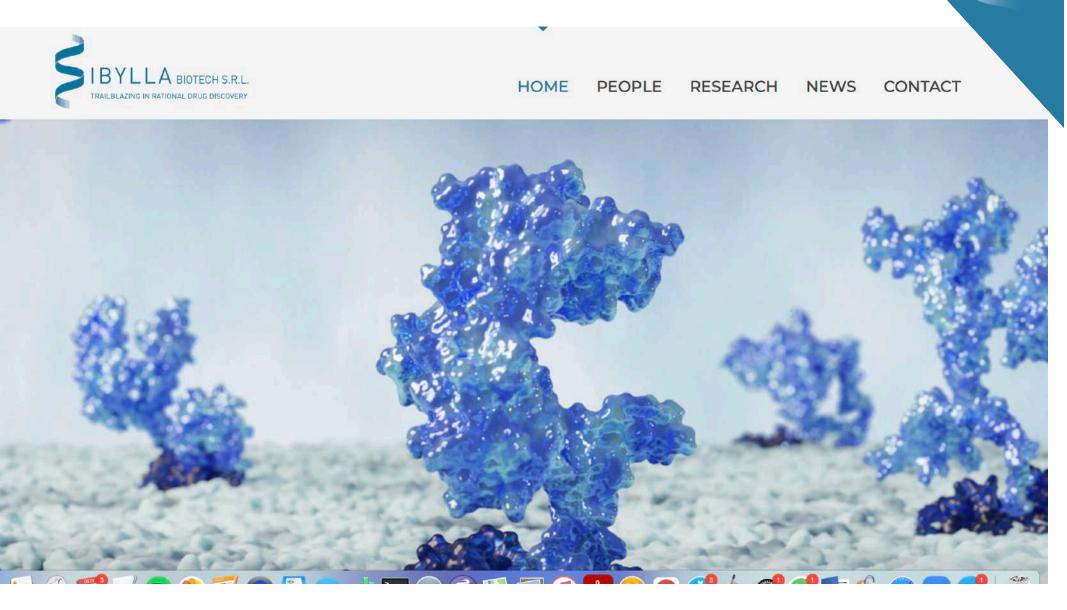
Pharmacological inactivation of the prion protein by targeting a folding intermediate

OPEN





Technology Transfer Initiative



Joining Forces against COVID-19



SARS-CoV-2 Replication

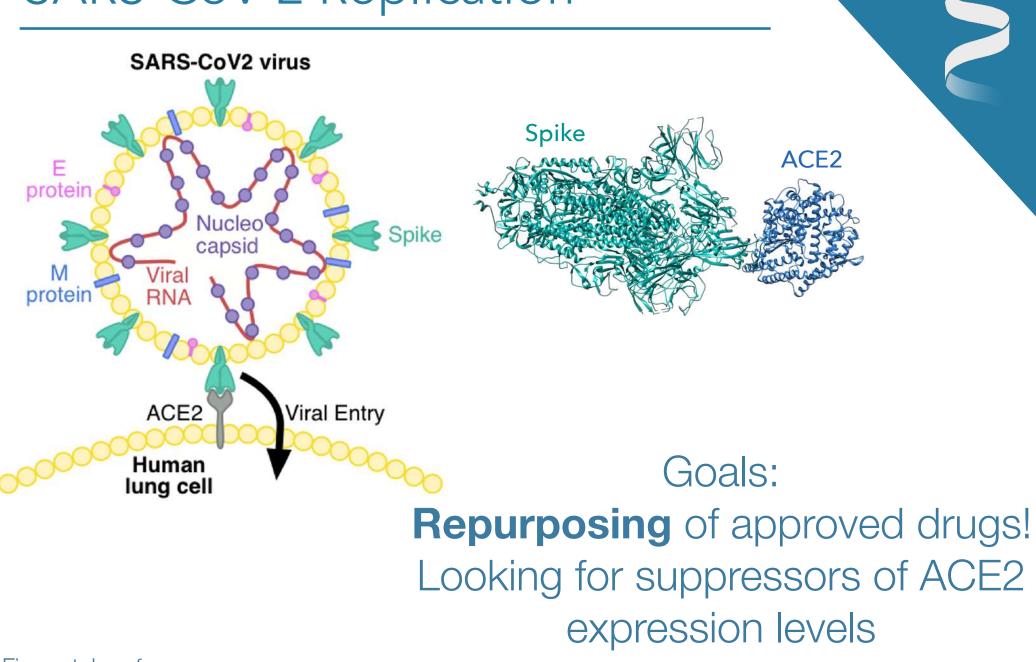
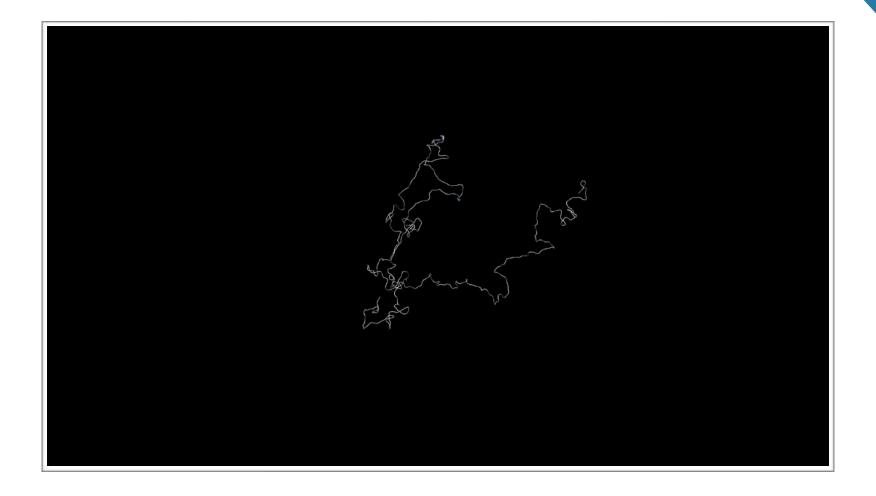
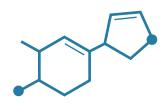


Figure taken from:

https://theconversation.com/where-are-we-at-with-developing-a-vaccine-for-coronavirus-134784

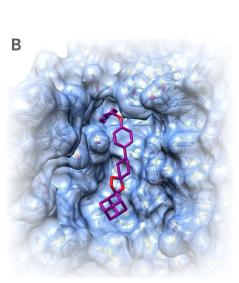
PPI-FIT ON ACE2

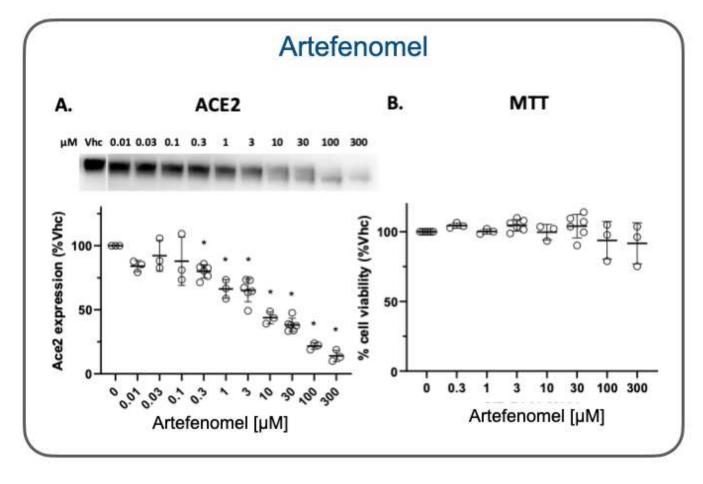




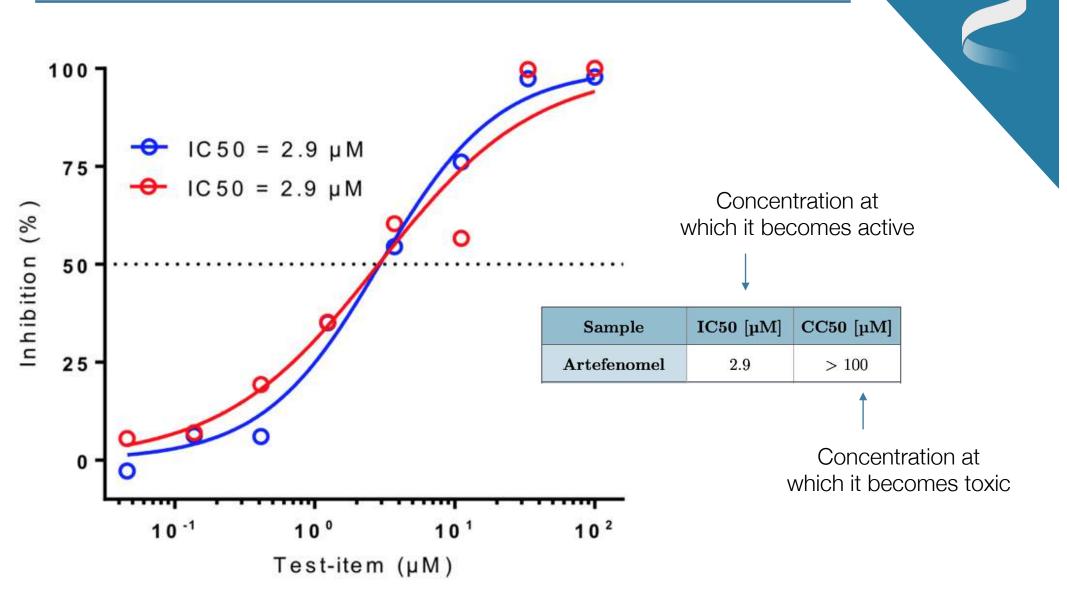
Out of 9000 candidates, we found 35 molecules binding in-silico the intermediate. Validation experiments on cellular bio-assays are ongoing.

DOSE-DEPENDENT RESPONSE





ANTI-VIRAL ACTIVITY AGAINST LIVE SARS-COV2

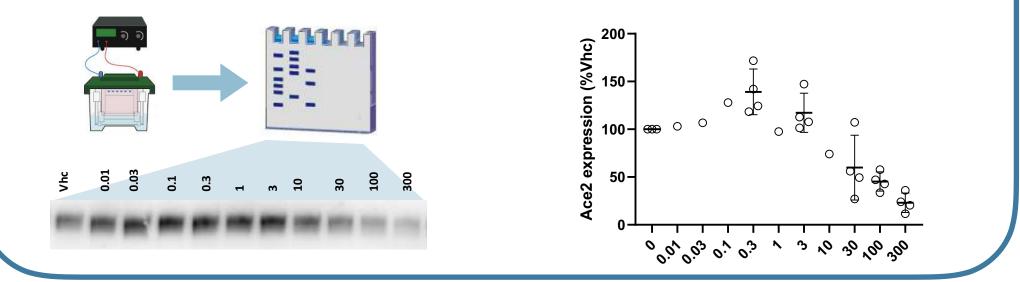


This value is *in principle* compatible with the maximum tolerated dose in humans. More to follow...

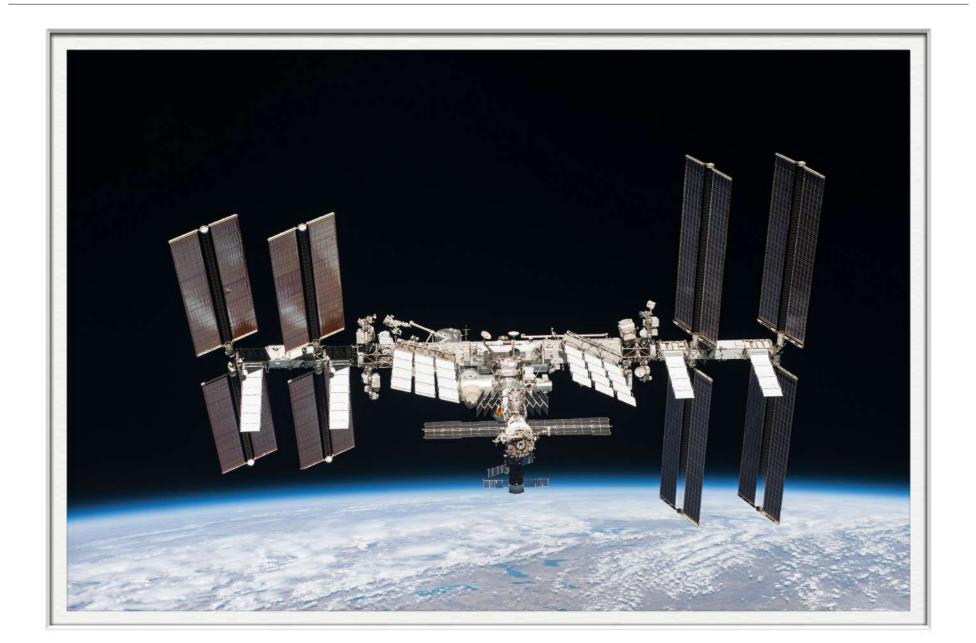
BREAKING NEWS!! (17/05/2020)

So far, Sibylla Biotech has tested 14 virtual hits

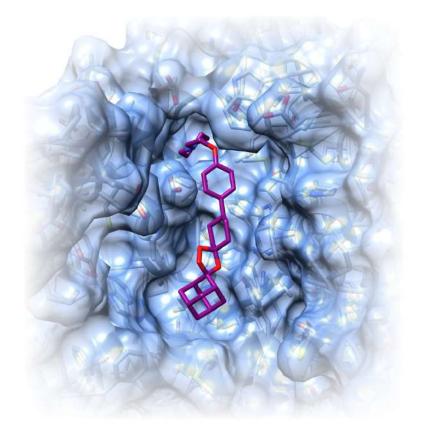
ONE DISPLAYS A **PROMINENT EFFECT** WITH CLEAR **DOSE-RESPONS**E RELATIONSHIP AND VERY **LOW TOXICITY**



SPACE IS THE NEXT FRONTIER!



A MAIN LIMITING FACTOR



Impossible to crystallize folding intermediates on Earth



Microgravity conditions may provide the solution! **Molecular Biology PRIN** (functional role of folding intermediates?)

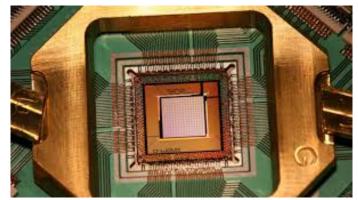
ZEFIR MISSION*



(* PRELIMINARY NAME)



Quantum Computing + AI



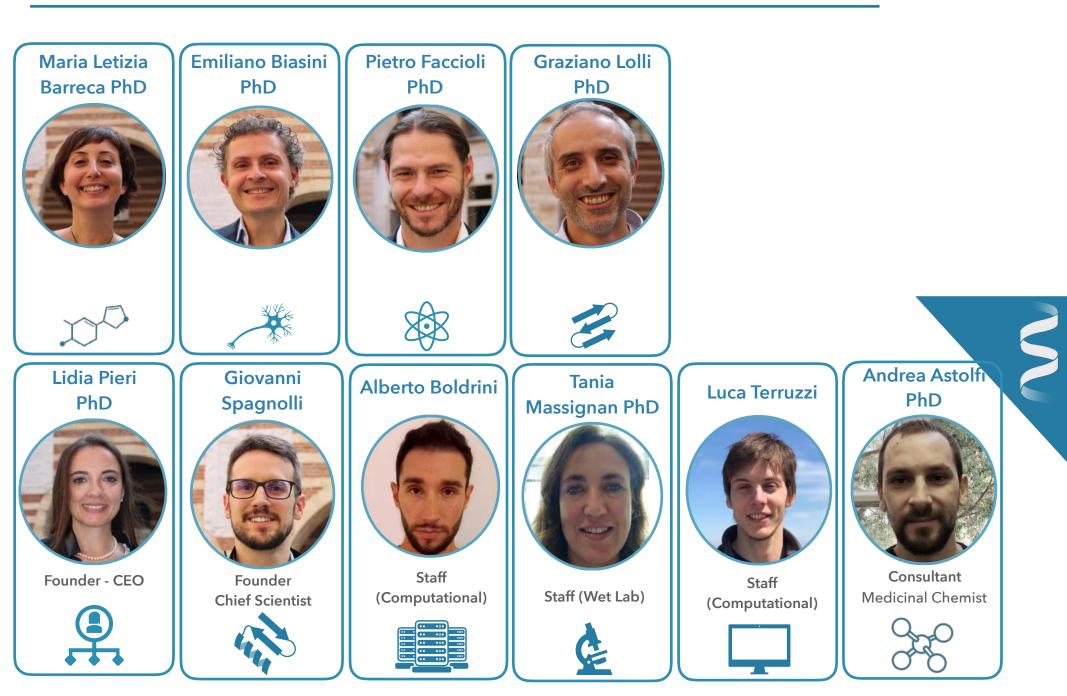
PHYSICAL REVIEW LETTERS VOL.XX, 000000 (XXXX)

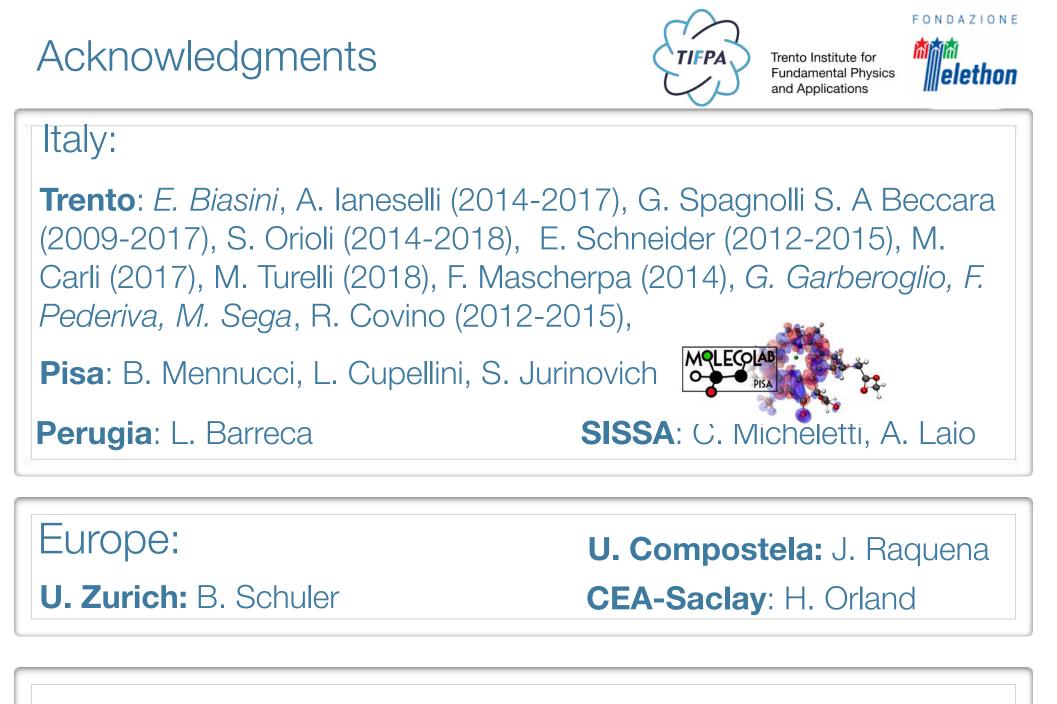
Dominant Reaction Pathways by Quantum Computing

Philipp Hauke[•], ¹ Giovanni Mattiotti[•], ² and Pietro Faccioli^{2,3} ¹INO-CNR BEC Center and Department of Physics, University of Trento, Via Sommarive 14, I-38123 Trento, Italy ²Department of Physics, University of Trento, Via Sommarive 14, I-38123 Trento, Italy ³INFN-TIFPA, Via Sommarive 14, I-38123 Trento, Italy

(Received 27 July 2020; accepted 18 December 2020)

People





USA: U. Maryland: P. Wintrode

U. Mass.: A. Gershenson