RealQM vs. DFT for Protein Folding

Why density-functional theory cannot dynamically simulate folding, and RealQM can

Claes Johnson[1]

[1]KTH Royal Institute of Technology, Stockholm, Sweden , `claesjohnson@gmail.com`

July 2, 2026

Developed in collaboration with Claude (Anthropic).

Abstract

Protein folding is a **dynamical** process — a chain moving from an extended state to its native fold over microseconds, in a system of 10^3–10^5 atoms. We argue that **density-functional theory (DFT) cannot follow this trajectory**: its cost scales as N^3 in the electron number and it is locked to femtosecond steps, putting a microsecond fold of a solvated protein at $\sim 10^{13}$ core-hours ($\sim$millennia) — 10^9–10^{12} beyond feasibility. **RealQM can**: its cost is set by a *fixed spatial grid* (200^3–1000^3 cells), *independent of the atom count* (a 200^3 grid handles $\sim$1,000 atoms, a 1000^3 grid $\sim$100,000), and it runs large-step (Brownian) dynamics on the GPU — with folding trajectories already demonstrated (chignolin, crambin, GB1, solvated α-helix, coiled coil). Biases may be used to guide the fold; that helps RealQM and does nothing for DFT, whose disqualifying cost is *per step.*

1 The problem: folding is dynamics, at scale

Folding is not a single minimisation but a **time-evolution** through a funnel of conformations. Timescale: microseconds (small fast-folders) to milliseconds. Size: $\sim 10^3$ atoms bare, but 10^4–10^5 atoms ($\sim 10^5$ electrons) with explicit solvent. What is required is to follow the *whole trajectory*, not to score a static structure — so any quantum method must run **many time steps on a large system**. This is where DFT and RealQM diverge.

2 Why DFT cannot

- $O(N^3)$ **in electrons.** Each self-consistent solve scales as the cube of the electron number (orthogonalisation / diagonalisation); a solvated protein ($\sim 10^5$ e) is $\sim 10^6\times$ a 1,000-atom benchmark.
- **Femtosecond steps.** Ab-initio MD is locked to $\sim$1 fs by nuclear vibrations, so $1\,\mu\text{s} = 10^9$ steps, each a full electronic solve.
- **Bottom line.** For $\sim 5\,\mu$s of a solvated protein: 5×10^9 steps $\times$ $\sim 10^6$ core-seconds/step $\approx$ $\mathbf{10^{13}}$ **core-hours** $\approx$ $\sim$1,000+ **years on a million-core machine** — for one trajectory.

In practice DFT-AIMD reaches $\sim$100 ps for $\sim$1,000 atoms. Folding needs $\sim 10^6\times$ longer *and* $\sim 30\times$ larger ($\times N^3 = 10^4\times$ per step) — a combined $\sim\mathbf{10^9}$**–**$\mathbf{10^{12}}$ beyond feasibility, a wall rather than a hardware wait. **Biases do not help DFT**: restraints cut how many steps you must *sample*, but the disqualifying cost is the *per-step* N^3 electronic solve, paid every step.

3 Why RealQM can

RealQM is quantum mechanics as real-space continuum mechanics on a grid, with a different cost structure:

- **Cost set by the grid, not by N.** All atoms share *one fixed spatial grid*; per-step cost scales with the number of grid cells (multigrid), *independent of atom count.* A 200^3 grid handles $\sim$1,000 atoms; a 1000^3 grid $\sim$100,000 atoms — at the same per-cell cost. Adding atoms does not blow the cost up as N^3 does.
- **Large-step dynamics.** Overdamped / Brownian dynamics replaces fs-locked Newtonian AIMD, reaching folding timescales in orders of magnitude fewer steps.
- **GPU-native.** The real-space grid maps directly onto GPU hardware.
- **Demonstrated.** Folding trajectories already run — chignolin, crambin, GB1, a solvated α-helix, a coiled coil.
- **Biases permitted.** As in classical MD (native-contact biasing, restraints, enhanced sampling), biases may guide the fold; this helps RealQM and, again, does nothing for DFT.

4 Side by side

	DFT (ab-initio MD)	**RealQM**
cost per step	$O(N^3)$ in electrons	$O(\text{grid})$ — fixed by resolution
capacity	$\sim$1,000 atoms	200^3: 1,000; 1000^3: 100,000 atoms
time stepping	fs-locked, $10^9/\mu$s	Brownian / large-step
reach today	$\sim$100 ps / 1,000 atoms	full folding trajectories
$\sim 5\,\mu$s fold	$\sim 10^{13}$ core-hr (millennia)	days–weeks on GPUs
biases	don't help	permitted, guide the fold
verdict	**impossible**	**feasible**

The gap between the verdicts is $\sim 10^6$–10^{10} in compute: the difference between "not in any foreseeable machine" and "runs on a GPU cluster."

5 Honest status

Beyond the capability gap, RealQM has a real advantage on the **physics of the dynamics** that should be stated plainly: it computes the **forces on the atoms directly from the electron charge potentials** — the actual Coulomb forces that drive the motion — rather than as **gradients of an energy surface**, as in the standard energy-first approach. Dynamics *is* force; RealQM supplies force physically and directly. On this basis RealQM is **stronger, not weaker** — the force computation is a strength, not a weak point.

The one genuinely open question is quantitative **fidelity of the folded state** — whether the physical force field carries the chain to the correct native structure across the whole funnel, and how far biases (native-contact restraints, enhanced sampling, standard tools in classical MD) are used. That is ordinary validation, not a defect of the method: the physical, charge-based forces are a reason for confidence, and biases help RealQM while doing nothing for DFT, which remains disqualified by cost.