Supplemental Methods –

PDB structure and model preparation

Protein structures of ubiquitin, TATA binding protein (TBP) (PDB: 1cdw)(1), BRAF kinase bound to the drug sorafenib (PDB: 1uwh)(2), SARS-CoV-2 receptor binding domain bound to angiotensin converting enzyme 2 (ACE2)(PDB:6m17)(3), and the allosteric RTX cysteine protease domain of the Vibrio cholera toxin (PDB:3eeb)(4) were obtained from the Protein Data Bank (PDB). After downloading the structures from the PDB database, any crystallographic reflections, ions, and other solvents used in the crystallization process were removed. Any missing loop structures in the protein structures were inferred using the MODELLER homology modelling server in UCSF Chimera. pdb4amber (AmberTools20) was employed to add hydrogen atoms (i.e. reduce the structure) and remove crystallographic waters.

Molecular dynamic simulation protocols

For each molecular dynamic comparison (monomer vs. dimer, wildtype vs. mutant protease; protease bound and unbound to drug), accelerated molecular dynamic (MD) simulations were performed (5). MD simulation protocol was followed as previously described, with slight modifications(6–9, 9–12). In brief, for each MD comparison, large replicate sets of accelerated MD simulation were prepared and then conducted using the particle mesh Ewald method implemented on NVIDIA graphical processor units by pmemd.cuda running Amber20 (13–17) and/or Langevin integration using mixed precision implementation via OpenMM (18). Alternative settings for OpenMM integrators such as Langevin, Verlet and aMD (19, 20) were also compared (see Supplemental Figures 8-9). The MD simulations were done on a high performance computing workstation mounting dual Nvidia 2080Ti graphics processor units. All comparative MD analysis via our ATOMDANCE was based upon 100 randomly resampled windows collected on of 10ns of accelerated MD in each comparative state, e.g., monomer vs. dimer, wildtype vs. mutant, protease bound to drug vs. protease unbound to drug). Explicitly solvated protein systems were first prepared using teLeap (AmberTools 20), using ff14SSB protein force field, in conjunction with modified GAFF2 small molecule force field (21, 22). Solvation was generated using the Tip3p water model in a 12nm octahedral water box. Charge neutralization was performed using Na+ and Cl- ions using the AmberTools22 teLeap program. Force field modifications for the small molecule ligands were generated using scaled quantum mechanical optimization via the sqm version 17 program in antechamber/AmberTools22 (23). For each MD comparison, an energy minimization was first performed, then heated to 300K for 300 pico seconds, followed by 10 ns of equilibration, and then finally a replicate set of 100 MD production run was created for each comparative state. Each MD production run was simulation for 1 ns of time. All simulations were regulated using the Andersen thermostat at 300k and 1atm (24). Root mean square atom fluctuations were conducted in CPPTRAJ using the atomicfluct command (25). All molecular color-mapping of our results were conducted in UCSF ChimeraX (26, 27). Any x-ray crystal protein structures requiring missing loop refinement were corrected using MODELLER prior to preparation for molecular dynamics simulations (28).

Additional information about the ATOMDANCE software suite for comparative molecular dynamics simulations

ATOMDANCE is available at GitHub/GitHub page

https://github.com/gbabbitt/ATOMDANCE-comparative-protein-dynamics

https://gbabbitt.github.io/ATOMDANCE-comparative-protein-dynamics/

Examples presented in this manuscript were generated from structure, topology, and trajectory files deposited here

https://zenodo.org/record/7679282#.Y_wIK9LMJ9A

DOI 10.5281/zenodo.7679282

makeMovie.py

A supplemental python GUI program for making molecular dynamics movies that are weighted by the normalized MMD in atom fluctuation between two functional states. The program first creates a multi-frame PDB file representing the true dynamics of the protein system, then it creates a multi-frame PDB file where the noise in the trajectories is dampened or amplified according to MMD. This creates a purely visual effect in a color-mapped movie of protein motion that demonstrates what the MMD filter captures. We have demonstrated this in examples of both dampening of atom motion during TATA binding protein interaction with DNA and with amplification of motion in the activation loop of BRAF kinase during cancer drug binding. https://people.rit.edu/gabsbi/img/videos/MMDmovie.mp4

MDgui.py

We also provide a full python GUI for running MD simulations using open source AmberTools and openMM. The user can generate MD trajectory and topology files using any software they prefer. Other options include Amber (licensed), NAMD/QwikMD (free), CHARMM (licensed), or openMM (free). The cpptraj software available on GitHub or in AmberTools can be used to convert common trajectory file formats to the binary format (.nc) used by ATOMDANCE. We also offer a useful python+perl-based GUI for licensed versions of Amber available here. https://gbabbitt.github.io/amberMDgui/

Note on the naming of things:
DROIDS – acronym for Detecting Relative Outlier Impacts in Dynamics Simulations

maxDemon – abbreviated from Maxwell's Demon, a 19th century thought experiment connecting the concepts of information and entropy in thermodynamics involving a mythical demon watching/assessing the motion of every atom in a system.

ChoreoGraph – evokes a notion of when motions of atoms at amino acids site 'move together' in a coordinated manner, in much the same way dancers may move together in choreography.

ATOMDANCE – an homage to a song composition by Icelandic singer Bjork Guomundsdottir from her 2015 album Vulnicura (One Little Indian Records)

Supplemental File – video overview with dynamics of DNA-bound TATA binding protein and sorafenib drug-bound B-Raf kinase domain weighted in accordance with maximum mean discrepancy in atom fluctuation. https://people.rit.edu/gabsbi/img/videos/MMDmovie.mp4

References

1. Nikolov, D.B., H. Chen, E.D. Halay, A. Hoffman, R.G. Roeder, and S.K. Burley. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc. Natl. Acad. Sci. U. S. A.* 93:4862–4867.

2. Wan, P.T.C., M.J. Garnett, S.M. Roe, S. Lee, D. Niculescu-Duvaz, V.M. Good, C.M. Jones, C.J. Marshall, C.J. Springer, D. Barford, R. Marais, and Cancer Genome Project. 2004. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell*. 116:855–867.

3. Yan, R., Y. Zhang, Y. Li, L. Xia, Y. Guo, and Q. Zhou. 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*. 367:1444–1448.

4. Lupardus, P.J., A. Shen, M. Bogyo, and K.C. Garcia. 2008. Small molecule-induced allosteric activation of the Vibrio cholerae RTX cysteine protease domain. *Science*. 322:265–268.

5. Case, D.A., T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R.J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.

6. Babbitt, G.A., E.P. Fokoue, H.R. Srivastava, B. Callahan, and M. Rajendran. 2022. Statistical machine learning for comparative protein dynamics with the DROIDS/maxDemon software pipeline. *STAR Protoc.* 3:101194.

7. Babbitt, G.A., M.L. Lynch, M. McCoy, E.P. Fokoue, and A.O. Hudson. 2022. Function and evolution of B-Raf loop dynamics relevant to cancer recurrence under drug inhibition. *J. Biomol. Struct. Dyn.* 40:468–483.

8. Babbitt, G.A., M.L. Lynch, M. McCoy, E.P. Fokoue, and A.O. Hudson. 2022. Function and evolution of B-Raf loop dynamics relevant to cancer recurrence under drug inhibition. *J. Biomol. Struct. Dyn.* 40:468–483.

9. Babbitt, G.A., E.P. Fokoue, J.R. Evans, K.I. Diller, and L.E. Adams. 2020. DROIDS 3.0—Detecting Genetic and Drug Class Variant Impact on Conserved Protein Binding Dynamics. *Biophys. J.* 118:541–551.

10. Babbitt, G.A., J.S. Mortensen, E.E. Coppola, L.E. Adams, and J.K. Liao. 2018. DROIDS 1.20: A GUI-Based Pipeline for GPU-Accelerated Comparative Protein Dynamics. *Biophys. J.* 114:1009–1017.

11. Rajendran, M., M.C. Ferran, and G.A. Babbitt. 2022. Identifying vaccine escape sites via statistical comparisons of short-term molecular dynamics. *Biophys. Rep.* 2:100056.

12. Rajendran, M., and G.A. Babbitt. 2022. Persistent cross-species SARS-CoV-2 variant infectivity predicted via comparative molecular dynamics simulation. *R. Soc. Open Sci.* 9:220600.

13. Case, D.A., T.E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K.M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, and R.J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.

14. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

15. Ewald, P.P. 1921. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* 369:253–287.

16. Pierce, L.C.T., R. Salomon-Ferrer, C. Augusto F. de Oliveira, J.A. McCammon, and R.C. Walker. 2012. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J. Chem. Theory Comput.* 8:2997–3002.

17. Salomon-Ferrer, R., A.W. Götz, D. Poole, S. Le Grand, and R.C. Walker. 2013. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* 9:3878–3888.

18. Eastman, P., J. Swails, J.D. Chodera, R.T. McGibbon, Y. Zhao, K.A. Beauchamp, L.-P. Wang, A.C. Simmonett, M.P. Harrigan, C.D. Stern, R.P. Wiewiora, B.R. Brooks, and V.S. Pande. 2017. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Comput. Biol.* 13:e1005659.

19. Hamelberg, D., C.A.F. de Oliveira, and J.A. McCammon. 2007. Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. *J. Chem. Phys.* 127:155102.

20. Hamelberg, D., J. Mongan, and J.A. McCammon. 2004. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* 120:11919–11929.

21. Maier, J.A., C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, and C. Simmerling. 2015. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 11:3696–3713.

22. Wang, J., R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25:1157–1174.

23. Walker, R.C., M.F. Crowley, and D.A. Case. 2008. The implementation of a fast and accurate QM/MM potential method in Amber. *J. Comput. Chem.* 29:1019–1031.

24. Andersen, H.C. 1980. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* 72:2384–2393.

25. Roe, D.R., and T.E.I. Cheatham. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 9:3084–3095.

26. Goddard, T.D., C.C. Huang, E.C. Meng, E.F. Pettersen, G.S. Couch, J.H. Morris, and T.E. Ferrin. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci. Publ. Protein Soc.* 27:14–25.

27. Pettersen, E.F., T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, and T.E. Ferrin. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* 30:70–82.

28. Sali, A., and T.L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.