Submission ID #: 61125
Scriptwriter Name: Bridget Colvin
Project Page Link: http://www.jove.com/files_upload.php?src=18643173

Title: Deciphering the Structural Effects of Activating EGFR Somatic Mutations with Molecular Dynamics Simulation

Authors and Affiliations: Mahlet Z. Tamirat[1], Kari J. Kurppa[2], Klaus Elenius[2,3,4], and Mark S. Johnson[1]

[1]Structural Bioinformatics Laboratory, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University
[2]Medicity Research Laboratories and Institute of Biomedicine, University of Turku
[3]Turku Bioscience Centre, University of Turku and Åbo Akademi University
[4]Department of Oncology and Radiotherapy, University of Turku and Turku University Hospital

Corresponding Author:
Mark S. Johnson
johnson4@abo.fi

Co-authors:
mtamirat@abo.fi
kjkurp@utu.fi
klaele@utu.fi

Author Questionnaire

1. Microscopy: Does your protocol involve video microscopy, such as filming a complex dissection or microinjection technique? **N**

2. Software: Does the part of your protocol being filmed demonstrate software usage? **Y**

Videographer: All screen capture files provided, do no not film screen shots

3. Filming location: Will the filming need to take place in multiple locations (greater than walking distance)? **N**

Script Length
Steps: **57**

Introduction

1. Introductory Interview Statements

REQUIRED:

1.1. **Mark Johnson**: Molecular dynamics simulation is a computational technique that can be used to explore molecular movements and can reveal conformational changes crucial to understanding biochemical and cellular function **[1]**.

1.1.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

REQUIRED:

1.2. **Mark Johnson**: Probing the range of dynamic motions accessible to a macromolecule is difficult. Using the results of molecular dynamics simulations in conjunction with experimental data allows assessment of their functional significance **[1]**.

1.2.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

OPTIONAL:

1.3. **Mahlet Tamirat**: We will demonstrate using molecular dynamics simulations how mutations observed in cancer patients affect EGFR tyrosine kinase conformations and ligand binding **[1]**.

1.3.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

Protocol

2. Structure Preparation

2.1. To prepare the wild-type apo active EGFR (E-G-F-R) kinase structure, open the Chimera visualization program **[1-TXT]** and, under the **File** menu, click **Fetch by ID [2]**.

2.1.1. WIDE: Talent opening Chimera, with monitor visible in frame **TEXT: Chimera: TEXT: https://www.cgl.ucsf.edu/chimera/; EGFR: epidermal growth factor receptor**

2.1.2. SCREEN: screenshot_1: 00:00-00:03

2.2. Select the Protein Data Bank database and specify the protein data bank 2GS2 (two-G-S-two) code **[1]**.

2.2.1. SCREEN: screenshot_1: 00:03-00:09

2.3. To build the missing structural 2GS2 elements, acquire these segments from other EGFR structures **[1]**.

2.3.1. SCREEN: screenshot_1: 00:15-00:23

2.4. To construct the five-residue glutamate 746 (seven-hundred-forty-six) to alanine 750 (seven-hundred-fifty) ELREA (el-rea) deletion in EGFR, click **Favourite**, **Sequence**, and **Show sequence** to open the wild-type 2GS2 sequence and, in the resulting sequence window, click **Edit** and **Add sequence** to select the deletion mutant FASTA (fasta) format sequence **[1-TXT]**.

2.4.1. SCREEN: screenshot_1: 00:25 - 00:44 *Video Editor: can speed up* **TEXT: ELREA: Glutamate-Leucine-Arginine-Glutamate-Alanine**

2.5. In the alignment window select the **Structure** and **Modeller homology**. In the pop-up window, specify the 2GS2 composite structure as the template and the mutant sequence as the query to be modeled **[1]**.

2.5.1. SCREEN: screenshot_1: 00:46-01:01 *Video Editor: can speed up*

2.6. Then select a mutant model from the resulting models based on the zDOPE score and visual inspection **[1]**.

2.6.1. SCREEN: screenshot_1: 01:02-01:10

2.7. To prepare the wild-type apo (A-poh) inactive EGFR kinase structure, open protein data bank structure 2GS7 and add the missing segments from other EGFR structures, modeling the deletion mutant form as demonstrated **[1]**.

2.7.1. SCREEN: screenshot_1: 01:11-01:24

2.8. To prepare the ATP-bound wild-type active EGFR kinase structure, use protein data bank structure 2ITX as the principal structure, building the missing segments using other EGFR structures and modeling the deletion mutant form using the modeler as demonstrated **[1]**.

2.8.1. SCREEN: screenshot_1: 01:25-01:38

2.9. To construct the wild-type EGFR asymmetric dimer structure, open 2GS2 in Chimera and click **Tools**, **Higher-Order Structure**, and **Unit Cell** to convert the structure to the biological assembly that contains the activator and receiver kinases in the asymmetric arrangement **[1]**.

2.9.1. SCREEN: screenshot_1: 01:40-01:57 *Video Editor: can speed up*

2.10. Select the 2GS2 structure and enter **Make copies [1]**. Then select and save a single asymmetric dimer from the multiple copies of the dimer resulting from the symmetry operations **[2]**.

2.10.1. SCREEN: screenshot_1: 01:58-02:00
2.10.2. SCREEN: screenshot_1: 02:05-02:15

2.11. To build the alanine 702 (seven-hundred-two) valine mutant, select **Tools**, **Structure editing**, and **Rotamers** to replace alanine 702 with valine **[1]**.

2.11.1. SCREEN: screenshot_1: 02:17-02:35 *Video Editor: can speed up*

2.12. Open the structures in **Maestro [1]** and click the **protein preparation wizard** button **[2]**.

2.12.1. Talent opening structure, with monitor visible in frame

2.12.2. SCREEN: screenshot_2: 00:27-00:30

2.13. Then select **add hydrogen atoms** and **Fill in missing side-chain atoms** and click **Preprocess [1]**.

2.13.1. SCREEN: screenshot_2: 00:33-00:44

2.14. To determine the protonation states of ionizable residues at pH 7.0, click **Refine** and **Use PROPKA** (prop-kah) to optimize the orientation of asparagine, glutamine and histidine residues for hydrogen bonding **[1]**.

2.14.1. SCREEN: screenshot_2: 00:52-01:08 *Video Editor: please speed up and exclude 01:00-01:06*

2.15. Then minimize the structure **[1]**.

2.15.1. SCREEN: screenshot_2: 01:13-01:18

3. System Setup

3.1. To set up the simulation system, open the **Leap** program **[1]** and import the amber ff14SB (F-F-fourteen-S-B) force field and TIP3P (T-I-P-three-P) water molecules. For the ATP-bound systems, import parameters for ATP and load the structure **[2]**.

3.1.1. WIDE: Talent opening program
3.1.2. SCREEN: screenshot_3: 00:10-01:10 *Video Editor: please speed up*

3.2. Solvate the structure in an octahedral box with explicit TIP3P water molecules that extends 10 angstroms in all directions from the surface atoms of the protein **[1]**.

3.2.1. SCREEN: screenshot_3: 01:14-01:24 *Video Editor: please speed up*

3.3. Check the built system and add the necessary ions to neutralize it **[1]**.

3.3.1. SCREEN: screenshot_3: 01:26-02:01 *Video Editor: please speed up and exclude 01:30-01:40 and 01:56-02:00*

3.4. To sufficiently model biomolecular systems, add additional sodium and chloride atoms to the simulation box to bring the system salt concentration to 0.15 Molar. Then generate and save the topology and coordinate files of the system to serve as inputs for the subsequent production simulation **[1]**.

3.4.1. SCREEN: screenshot_3: 02:09-02:57 *Video Editor: please speed up and exclude 02:18-02:22, 02:32-02:38*

4. Molecular dynamics simulation

4.1. Using Amber, initially energy minimize **[1]** the simulation system to circumvent any unfavorable configurations **[2]**.

4.1.1. WIDE: Talent at computer, with monitor visible in frame
4.1.2. SCREEN: screenshot_4: 00:00-00:04

4.2. In the minimization input file, adjust the **maximum cycle** variable for the total minimization cycle and the **number of cycles** to indicate the number of cycles for the steepest descent algorithm **[1]**.

4.2.1. SCREEN: screenshot_4: 00:05-00:16 *Video Editor: please speed up*

4.3. Use the **restraint_weight** variable to apply the restraint force on the solute atoms specified by the **restraint mask** parameter **[1]**.

4.3.1. SCREEN: screenshot_4: 00:12-00:16 *Video Editor: please speed up*

4.4. Carry out the minimization in multiple steps, gradually lowering the restraint applied on solute atoms from 25 to 0 kilocalorie-per-mole-Ångstrom-square. Then use the command to run the minimization **[1]**.

4.4.1. SCREEN: screenshot_4: 00:17-01:01 *Video Editor: please speed up*

4.5. Heat the system for 100 picoseconds from 0 to 300 Kelvin and use the commands to set a 10 kilocalorie-molar/square-angstrom restraint on solute atoms. Then use the command to carry out the heating **[1]**.

4.5.1. SCREEN: screenshot_5: 00:00-00:49 *Video Editor: please speed up between 00:00-00:43*

4.6. Equilibrate the system for 900 picoseconds under an isothermal-isobaric ensemble and set a 9-angstrom distance cutoff for long range electrostatic interactions. Gradually lower the solute atom restraint to 0.1 kilo-calorie-per-mol-Ångstrom-square and finalize the equilibration with an unrestrained 5-nanosecond simulation **[1]**.

4.6.1. SCREEN: screenshot_6: 00:00-01:08 *Video Editor: please speed up*

4.7. Run the equilibration with the command as indicated **[1]**.

4.7.1. SCREEN: screenshot_6: 01:09-1:14

4.8. Adjust the text to allow the production simulation to be carried out for 100 nanoseconds, with the conformations being saved every 10 picoseconds **[1]**.

4.8.1. SCREEN: screenshot_7: 00:00-00:23

4.9. Then run the simulation with the command as indicated **[1]**.

4.9.1. SCREEN: screenshot_7: 00:24-00:30 *Video Editor: please speed up*

5. Visual Inspection Analysis

5.1. To visualize the conformations sampled during the wild-type and mutant EGFR kinase simulations, open the amber topology files and the corresponding trajectory files in **Visual Molecular Dynamics [1]** and, using convenient secondary structure representations, analyze the overall structural dynamics of the proteins from the recorded trajectory **[2-TXT]**.

5.1.1. WIDE: Talent opening file(s), with monitor visible in frame
5.1.2. SCREEN: screenshot_8: 00:26-00:28 **TEXT: https://www.ks.uiuc.edu/Research/vmd/**

5.2. Then view specific interactions between atoms and residues of interest, such as the catalytically essential lysine 745-(seven-hundred-forty-five) glutamate 762 (seven-hundred-sixty-two) salt bridge **[1]**.

5.2.1. SCREEN: screenshot_8: 00:29-00:33

5.3. Alternatively, save multiple conformations sampled during the simulation in Protein Data Bank format and open the conformations in Chimera **[1]**.

5.3.1. SCREEN: screenshot_9: 00:00-00:06

5.4. Use the **MatchMaker** option to superimpose the structures on the initial or median structure and display the median structure in solid and the rest of the aligned structures in faded white to allow visualization of the recorded structural movements with more clarity **[1]**.

5.4.1. SCREEN: screenshot_9: 00:10-00:40 *Video Editor: please speed up*

6. Root-Mean Square Deviation (RMSD) and Root-Mean Square Fluctuation (RMSF) Analysis

6.1. To analyze the global stability of the wild-type and mutant EGFRs and to examine the flexibility of the different structural units **[1]**, import the amber topology and corresponding trajectory files. In the root-mean square deviation input file, indicate the backbone atoms of the initial structure as reference for the root-mean square fitting **[2]**.

6.1.1. WIDE: Talent importing file, with monitor visible in frame
6.1.2. SCREEN: screenshot_10: 00:03-00:48 *Video Editor: please speed up*

6.2. In the root-mean square fluctuation input file, indicate the C-alpha atoms of the initial structure as reference for the root-mean square fitting **[3]**.

6.2.1. SCREEN: screenshot_10: 00:51-01:17 *Video Editor: please speed up*

6.3. Then run the analysis with the Cpptraj program and plot the output data **[1-TXT]**.

6.3.1. SCREEN: screenshot_10: 01:18-01:38 *Video Editor: please exclude 01:22-01:25*
TEXT: https://amber-md.github.io/cpptraj/CPPTRAJ.xhtml

6.4. Alternatively, to align the conformational ensembles and to color each residue based on the C-alpha atom root-mean square deviation, open the conformations in Chimera and align them with the **Matchmaker** option. Select **Tools**, **Depiction**, and **Render by attribute [1]**.

6.4.1. SCREEN: screenshot_11: 00:00-00:06

6.5. Select residues of the conformational ensembles and set the **C-alpha root-mean square deviation** as the attributes. Then click **OK [1]**.

6.5.1. SCREEN: screenshot_11: 00:07-00:30

6.6. The chain trace of the conformations will be colored blue, white, or red, reflecting the regions of high, medium, and low structural stability, respectively **[1]**.

6.6.1. SCREEN: screenshot_11: 00:31-00:36

7. Hydrogen Bond Analysis

7.1. To analyze the hydrogen bond interactions between ATP and both wild-type and deletion EGFRs **[1]**, prepare a Cpptraj script to carry out this task **[2]**.

7.1.1. WIDE: Talent inputting script, with monitor visible in frame
7.1.2. SCREEN: screenshot_12: 00:00-00:03

7.2. Specify the analysis for the intermolecular hydrogen bonds only with the **nointramol** variable **[1]** and define a hydrogen bond with a donor-acceptor distance of less than or equal to 3.5 ångstroms and a bond angle of greater than or equal to 135 degrees **[2]**.

7.2.1. SCREEN: screenshot_12: 00:04-00:09 *Video Editor: please speed up*
7.2.2. SCREEN: screenshot_12: 00:10-00:16 *Video Editor: please speed up*

7.3. To assess the intramolecular interactions, for example, between the catalytically important lysine and glutamate residues, specify lysine as the hydrogen donor and glutamate as the acceptor residues **[1]** and run the script as indicated to allow analysis of the result **[2].**

7.3.1. SCREEN: screenshot_12: 00:19-00:32 *Video Editor: can speed up*
7.3.2. SCREEN: screenshot_12: 00:33-00:49 *Video Editor: can speed up*

8. Free Energy Calculations

8.1. To compute the binding free energies between ATP and both wild-type and deletion EGFRs **[1]**, prepare the gas phase ligand, receptor, and ligand-receptor complex protein data bank files in the leap program and set ATP as the ligand and EGFR as the receptor **[2]**.

8.1.1. WIDE: Talent at computer, opening file(s), with monitor visible in frame
8.1.2. SCREEN: screenshot_13: 00:00-00:13

8.2. Set the generalized born radii value to m-bond-i-2. Then generate the amber topology and coordinate the files for the gas phase protein data bank files **[1]**.

8.2.1. SCREEN: screenshot_13: 00:14-00:25 *Video Editor: please speed up*

8.3. Similarly, to calculate binding free energies between the activator and receiver kinases of wild-type and alanine 702 valine EGFRs, specify the receiver kinase as the ligand and the activator kinase as the receptor and save the corresponding topology and coordinate files **[1]**.

8.3.1. SCREEN: screenshot_13: 00:27-00:41

8.4. Prepare a molecular mechanics generalized Born surface area input file and set the **igb value** to 2 and the **saltcon** to 0.1 **[1]**.

8.4.1. SCREEN: screenshot_13: 00:43-00:52

8.5. Then, using the MMPBSA.py (M-M-P-B-S-A dot pie) script available in Amber, enter the command as indicated to execute the binding energy calculations and analyse the output data **[1-TXT]**.

8.5.1. SCREEN: screenshot_13: 00:55-01:00 **TEXT: http://ambermd.org/**

Results

9. Results: Representative Molecular EGFR Somatic Mutation Dynamic Simulation

9.1. During this representative 100-nanosecond simulation **[1]**, the alanine 702 valine mutant showed increased conformational stability of the juxtamemebrane B segment, likely due to tighter hydrophobic interactions **[2]**, as compared to the wild type EGFR **[3]**.

9.1.1. LAB MEDIA: Figure 4A
9.1.2. LAB MEDIA: Figure 4A *Video Editor: please emphasize red data line*
9.1.3. LAB MEDIA: Figure 4A *Video Editor: please emphasize blue data line*

9.2. The alanine 702 valine mutant also exhibited a lower free energy of binding between the activator and receiver kinases **[1]** relative to the wild type EGFR **[2]**, representing more favorable dimer interactions that maintain the active EGFR kinase conformation **[3]**.

9.2.1. LAB MEDIA: Figure 6A *Video Editor: please emphasize red data line*
9.2.2. LAB MEDIA: Figure 6A *Video Editor: please emphasize blue data line*
9.2.3. LAB MEDIA: Figure 6A

9.3. Simulation of the deletion mutation resulted in lower C-alpha atom fluctuations of the functionally key alpha-C helix **[1]** as compared to the wild-type EGFR **[2]**, prolonging the time of the EGFR active state **[3].**

9.3.1. LAB MEDIA: Figure 4B *Video Editor: please emphasize yellow data line*
9.3.2. LAB MEDIA: Figure 4B *Video Editor: please emphasize blue data line*
9.3.3. LAB MEDIA: Figure 4B

9.4. The deletion mutation also resulted in frequent formation of hydrogen bonds between the side-chain polar atoms of lysine 745 (seven-hundred-forty-five) and glutamate 762 (seven-hundred-sixty-two) **[1]**, a key interaction for EGFR enzymatic activity compared to the wild-type EGFR **[2].**

9.4.1. LAB MEDIA: Figure 5A *Video Editor: please emphasize yellow data line*
9.4.2. LAB MEDIA: Figure 5A *Video Editor: please emphasize blue data line*

9.5. Additionally, the number of hydrogen bonds between ATP and EGFR were greater for the deletion mutant **[1]** than for the wild-type EGFR **[2].**

9.5.1. LAB MEDIA: Figure 5C *Video Editor: please emphasize yellow data bars*
9.5.2. LAB MEDIA: Figure 5C *Video Editor: please emphasize blue data bars*

9.6. The deletion mutation also resulted in an inward movement of the EGFR inactive state alpha-C helix, a structural change expected during the transition to the active state **[1]**.

9.6.1. LAB MEDIA: Figure 7B

9.7. In contrast, the alpha-C helix of wild-type inactive EGFR maintained its initial conformation **[1]**.

9.7.1. LAB MEDIA: Figure 7A

Conclusion

10. Conclusion Interview Statements

10.1. **Mahlet Tamirat**: To assess the impact of the mutations, it is critical to select the appropriate conformational states of the biologically relevant structures and to properly prepare and equilibrate these structures **[1]**.

10.1.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera

10.2. **Mahlet Tamirat**: It is important to combine molecular dynamic simulations with experimental studies, as the synergy between these techniques benefits the interpretation of the results and can inform additional wet-lab experiments **[1]**.

10.2.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera