

Supplementary Information

A multiscale modeling approach to investigate molecular mechanisms of pseudokinase activation and drug resistance in the HER3/ErbB3 receptor tyrosine kinase signaling network

Shannon E. Telesco, Andrew J. Shih, Fei Jia, and Ravi Radhakrishnan*

Department of Bioengineering, University of Pennsylvania

210 S. 33rd Street, 240 Skirkanich Hall, Philadelphia, PA 19104, USA

* *Corresponding Author Email: rradhak@seas.upenn.edu*

Homology modeling of the HER3 kinase domain

The protein sequence selected for alignment of the kinase domains included residues 678-957 (EGFR) and 683-962 (HER4); we opted to exclude the flexible C-tail from the alignment, as its sequence is highly variable among the ErbB kinases. A total of 50 models were generated from each of the templates (EGFR, HER4, and multiple templates) by satisfying a set of static and dynamic spatial restraints in MODELLER¹. These restraints are expressed in terms of a molecular probability density function, or objective function, which is optimized and applied in the ranking of the set of models constructed in MODELLER¹.

We then evaluated several stereochemical properties such as torsion angles and main-chain bond lengths for the most energetically favorable model from each template to identify any poorly-refined regions in the models. As the flexible A-loop (residues 833-855 in HER3) exhibited particularly unfavorable residue energies, we applied the loop-modeling algorithm in MODELLER² to remodel this sub-domain in each of our top models as well as in our HER3 crystal structure, which is missing nine residues in the A-loop. Nine residues representing the least energetically favorable amino acids were remodeled, including residues 842-850 (EGFR-based model), 840-848 (HER4-based model), 840-848 (MT model), and 845-853 (HER3 crystal structure), as the accuracy of loop modeling decreases with loop length². A total of 500 A-loop models were created, as this extent of conformational sampling for 9-residue loops correlates with maximal accuracy in the loop prediction². We then selected the top A-loop model from each template based on a combination of objective function score and stereochemical quality.

The Ramachandran plots for the top structures reveal that the majority of residues in the HER3 models lie within the most favored regions of phi-psi space (92.1%, 93.8%, 94.6% and 90.5% for the models based on EGFR, HER4, MT, and the HER3 crystal structure, respectively, see Fig. S1). PROCHECK³ was used to calculate the residue-by-residue G-factor, which provides a measure of the deviation of a given stereochemical property from its standard distribution, computed from a database of high-resolution protein structures. Specifically, the overall G-factors plotted in Fig. S2 average the contributions from the side chain torsion angle G-factors and the G-factors for main-chain bond lengths and bond angles, where darkened bars indicate low-probability conformations. Fig. S2 displays the improvement in the G-factors in each model following A-loop remodeling, particularly for the side chain torsion angles. For both the Ramachandran and G-factor analyses, the MT model appeared to produce the best results.

The stereochemical quality of each model was further evaluated using the Discrete Optimized Protein Energy (DOPE) method⁴, which is an atomic distance-dependent statistical potential optimized for model assessment in MODELLER. The DOPE score profiles for the leading models exhibit a significant energetic improvement in the A-loop region as compared to the original structures (Fig. S3), especially for the MT model. It is also apparent that the MT model is biased toward the HER4 template, particularly for residues 775-825, where the DOPE profiles for the MT- and HER4-based models decrease in energy, in contrast to the EGFR-based model, which exhibits more energetic peaks in this region. A combination of DOPE energy, objective function score and stereochemical quality were considered in order to determine the most energetically favorable HER3 models derived from each template. Furthermore, the RMSD among the top model A-loops was computed, as minimal variation among the low energy conformations correlates with a more pronounced free energy minimum and a higher level of accuracy in the best structural prediction². The superposition of the top 10 models from each template resulted in a dominant cluster of conformations, increasing our confidence in the reliability of the top structures. The top 10 MT models exhibited the smallest RMSD, or minimal variation. Our results emphasize the importance of selecting the best available template for homology modeling of even highly related proteins, and indicate that the application of multiple templates in the sequence alignment may improve the quality of homology models in certain cases⁵.

References

1. Sali, A., and T. L. Blundell. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779-815.
2. Fiser, A., R. K. Do, and A. Sali. (2000). Modeling of loops in protein structures. *Protein Sci* 9:1753-1773.
3. Laskowski, R. A., M.W. MacArthur, D.S. Moss, and J.M. Thornton. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283-291.
4. Shen, M. Y., and A. Sali. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15:2507-2524.
5. Larsson, P., B. Wallner, E. Lindahl, and A. Elofsson. (2008). Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Sci* 17:990-1002.
6. Schoeberl, B., E. A. Pace, J. B. Fitzgerald, B. D. Harms, L. Xu, L. Nie, B. Linggi, A. Kalra, V. Paragas, R. Bukhalid, V. Grantcharova, N. Kohli, K. A. West, M. Leszczyniecka, M. J. Feldhaus, A. J. Kudla, and U. B. Nielsen. (2009). Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Sci Signal* 2:ra31.
7. Shi, F., S. E. Telesco, Y. Liu, R. Radhakrishnan, and M. A. Lemmon. ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc Natl Acad Sci U S A* 107:7692-7697.

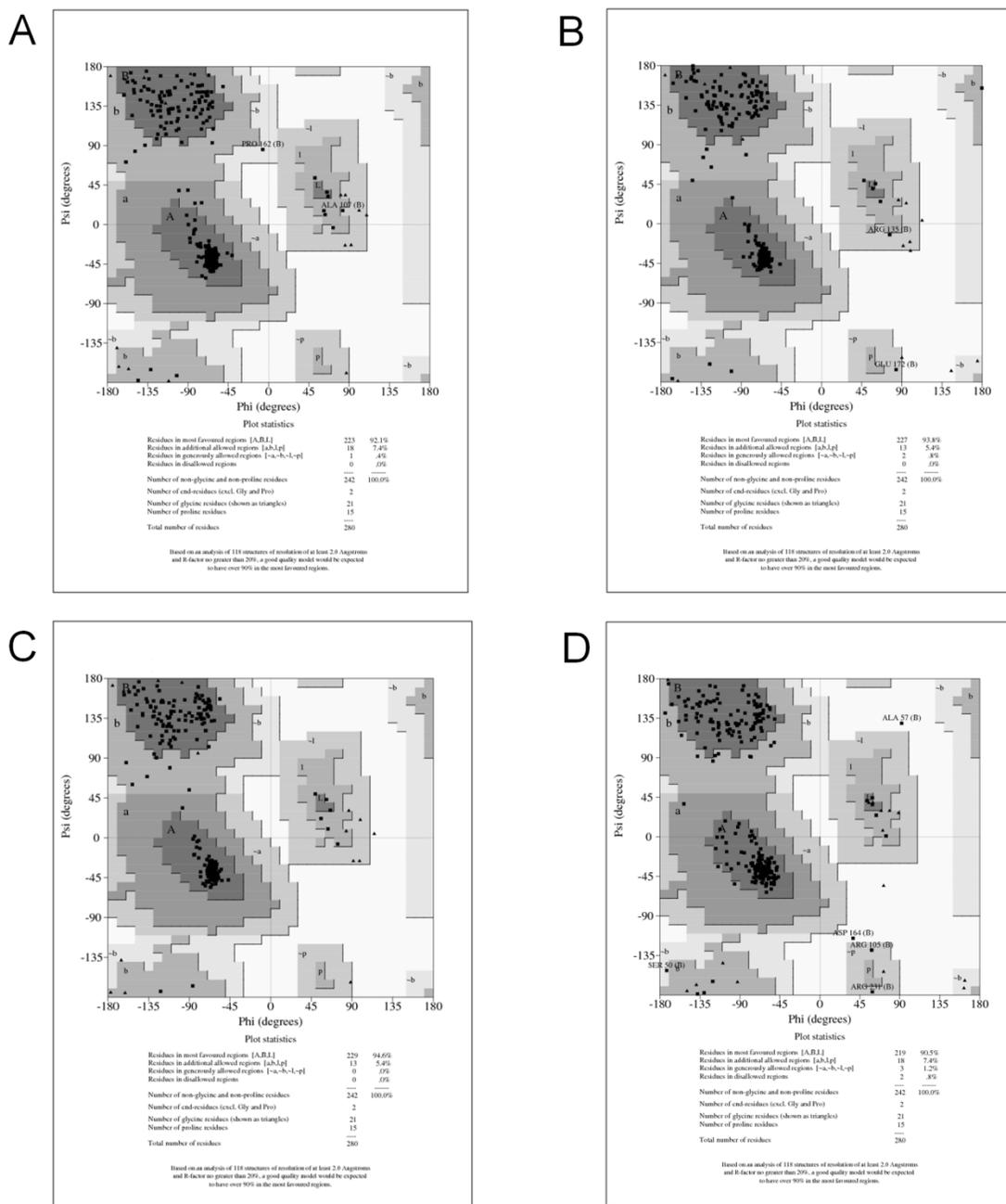


Figure S1. Ramachandran plots of the top HER3 structures modeled on (A) the EGFR template, (B) the HER4 template, (C) Multiple templates and (D) Loop-modeled HER3 crystal structure. All templates produced high-quality models with at least 90% of residues lying in the most favored regions of phi-psi space.

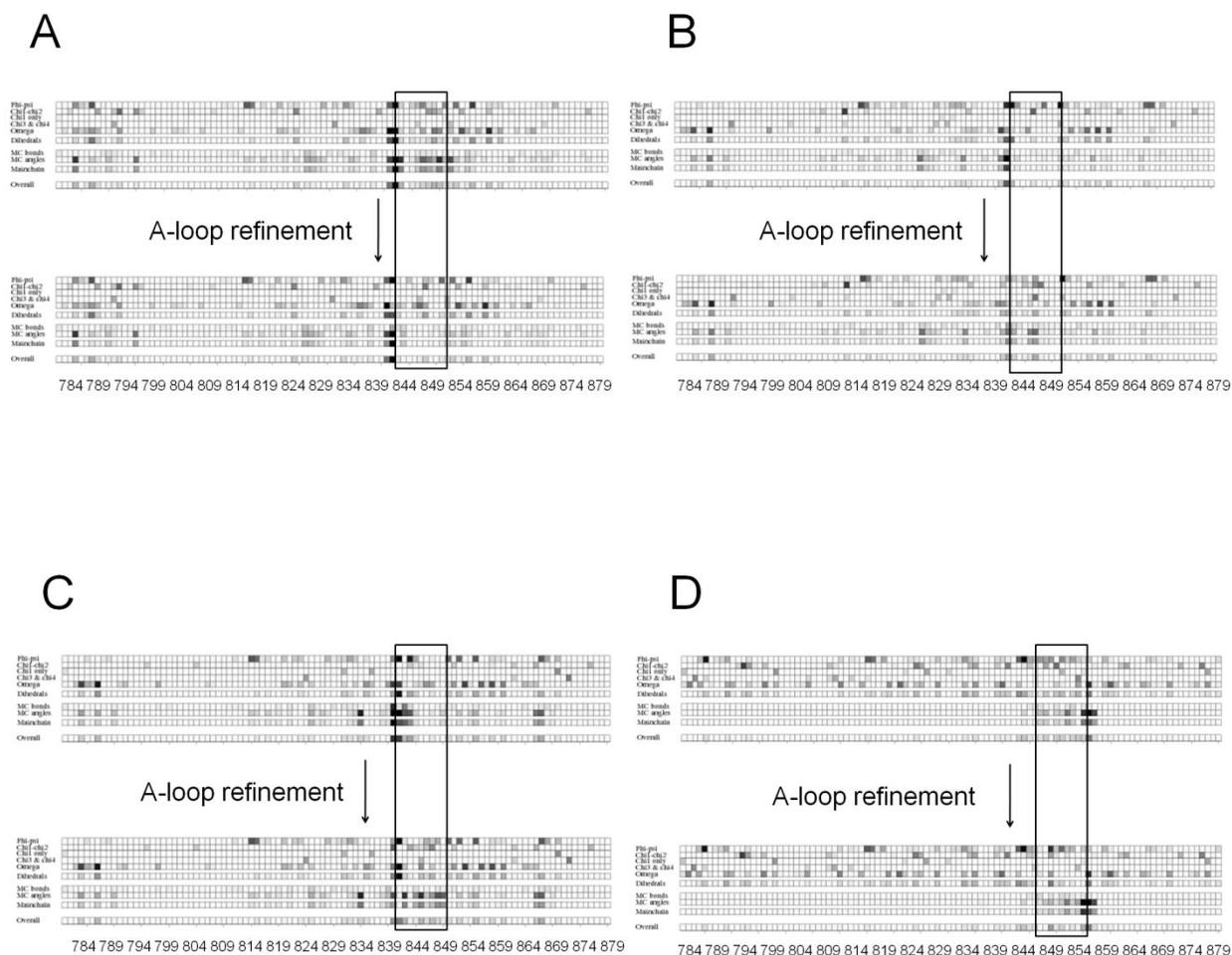


Figure S2. G-factor plots for the top HER3 models constructed from each ErbB template before and after A-loop refinement. Plots are shown for the HER3 structures modeled on (A) the EGFR template, (B) the HER4 template, (C) Multiple templates and (D) the loop-modeled HER3 crystal structure. The G-factors provide a measure of the deviation of a given stereochemical property from its standard distribution, computed from a database of high-resolution protein structures, where darkened bars indicate low-probability conformations. The A-loop region is boxed to highlight the improvement in G-factor scores after A-loop refinement.

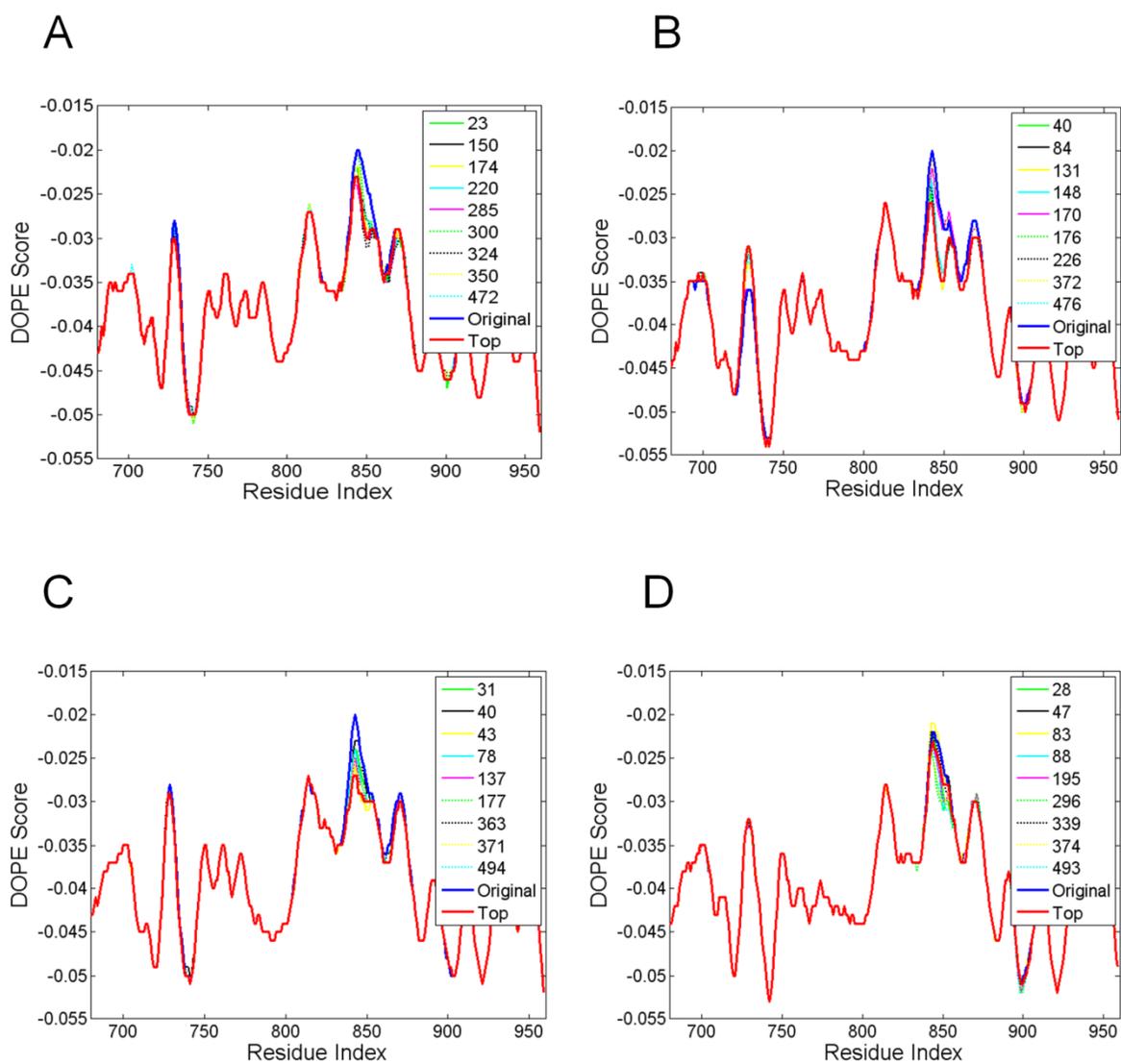


Figure S3. The Discrete Optimized Protein Energy (DOPE) plots for the top 10 refined A-loop models based on each ErbB template. The DOPE profiles are shown for the HER3 models based on (A) the EGFR template, (B) the HER4 template, (C) Multiple templates and (D) the loop-modeled HER3 crystal structure. The original, unrefined model is highlighted in blue and the top-scoring refined model is highlighted in red. The top models derived from each template are of similar quality, as illustrated by the improvement in the DOPE score in the A-loop region.

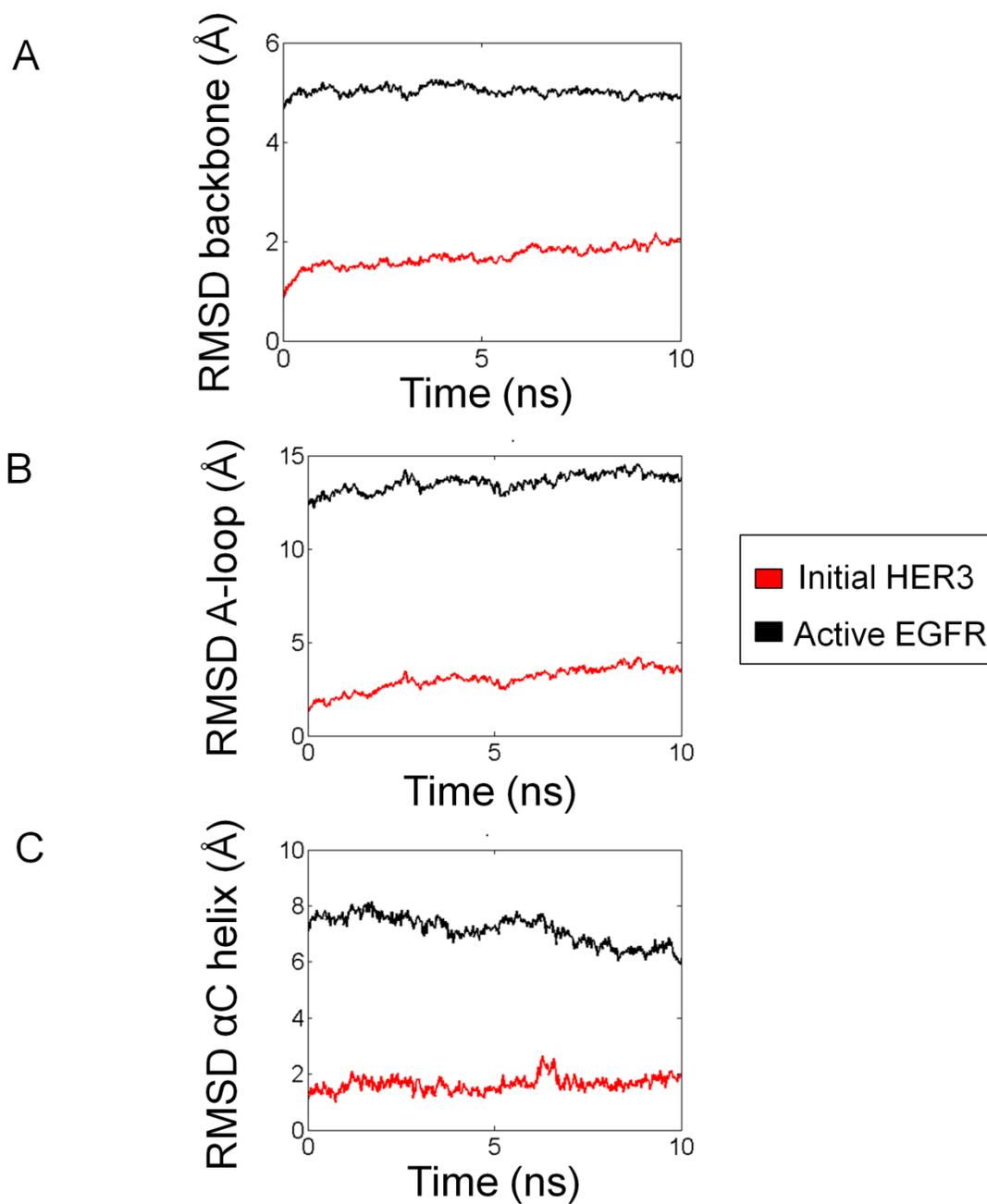


Figure S4. Molecular dynamics time course plots of the RMSD for (A) the backbone atoms of the HER3 kinase, (B) the backbone atoms of the A-loop and (C) the backbone atoms of the α C helix. The RMSD is plotted in reference to the initial (unsimulated) structure (*red*) as well as the active EGFR structure (*black*), for reference.

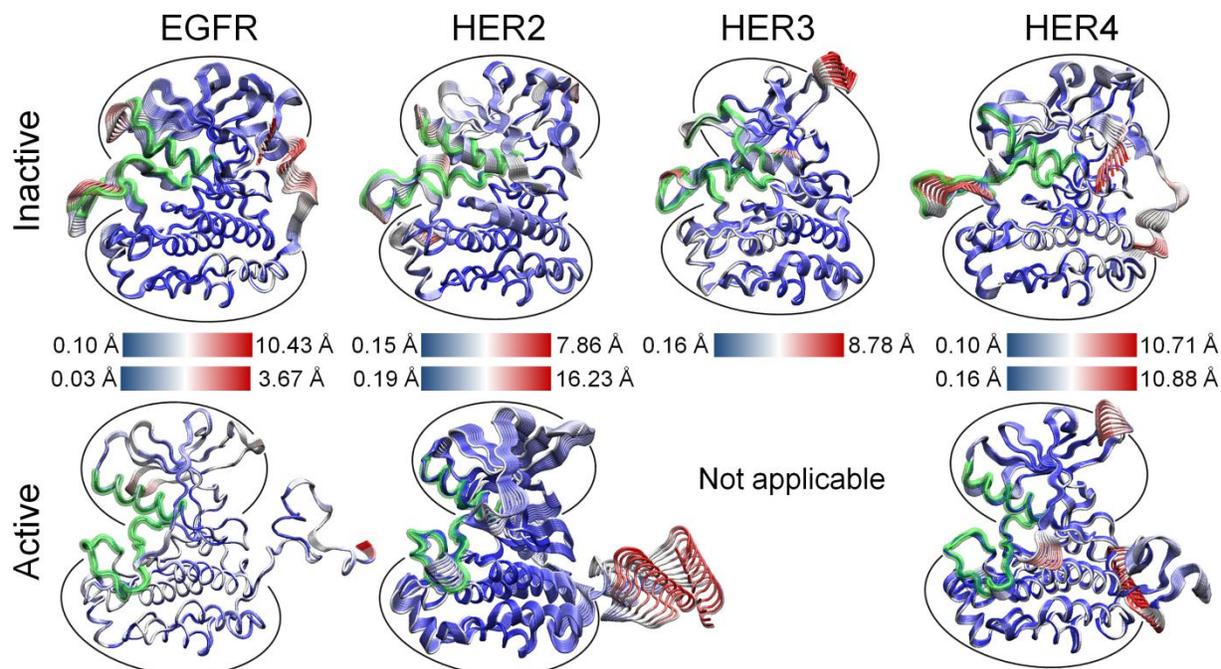


Figure S5. Motion along the first principal component of the MD trajectory is illustrated for the complete HER3 kinase and compared to the active and inactive conformations of EGFR, HER2 and HER4. The structures are color-coded according to the RMSD, where red regions indicate large-amplitude fluctuations and blue regions indicate small-amplitude fluctuations. The A-loop and α C helix are highlighted in green for structural reference. Overall, the global motions are conserved across the ErbB family members.

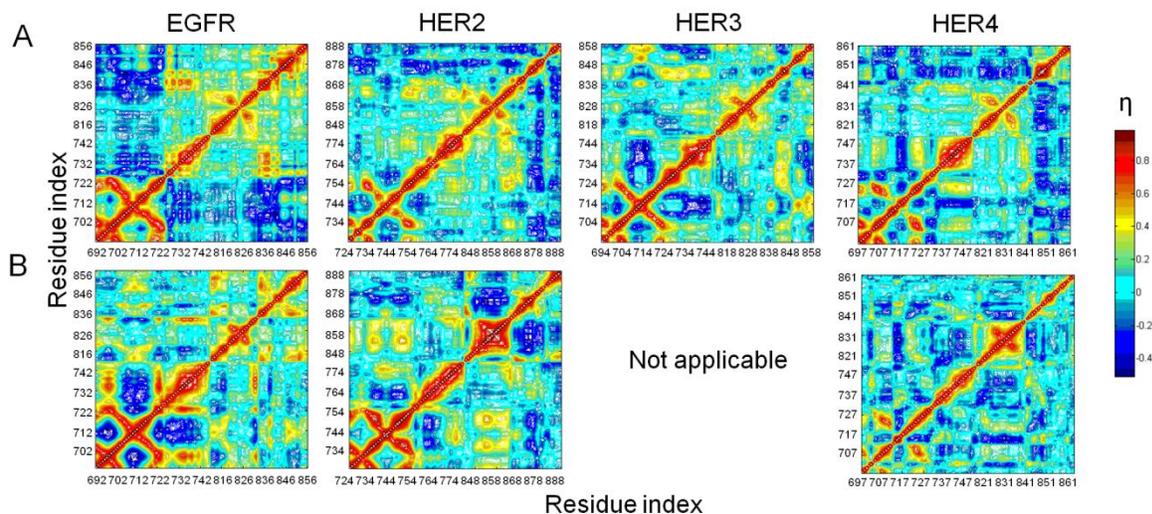


Figure S6. Normalized PCA cross-correlation matrices for vector displacements of atoms for the ErbB kinase systems (EGFR, HER2, HER3, HER4) in (A) the inactive state and (B) the active state. Correlated fluctuations of the C_{α} atoms in the active site region of the kinases (A-loop, C-loop, N-loop and α C helix, *ie*, residues 694-748 and 813-858 in HER3) are colored according to their degree of correlation as quantified by the normalized correlation coefficient, η . Residue pairs with a high degree of correlated motion are shown in orange and red, anticorrelated residue pairs are shown in dark blue, and weakly or uncorrelated residue pairs are shown in green and cyan.

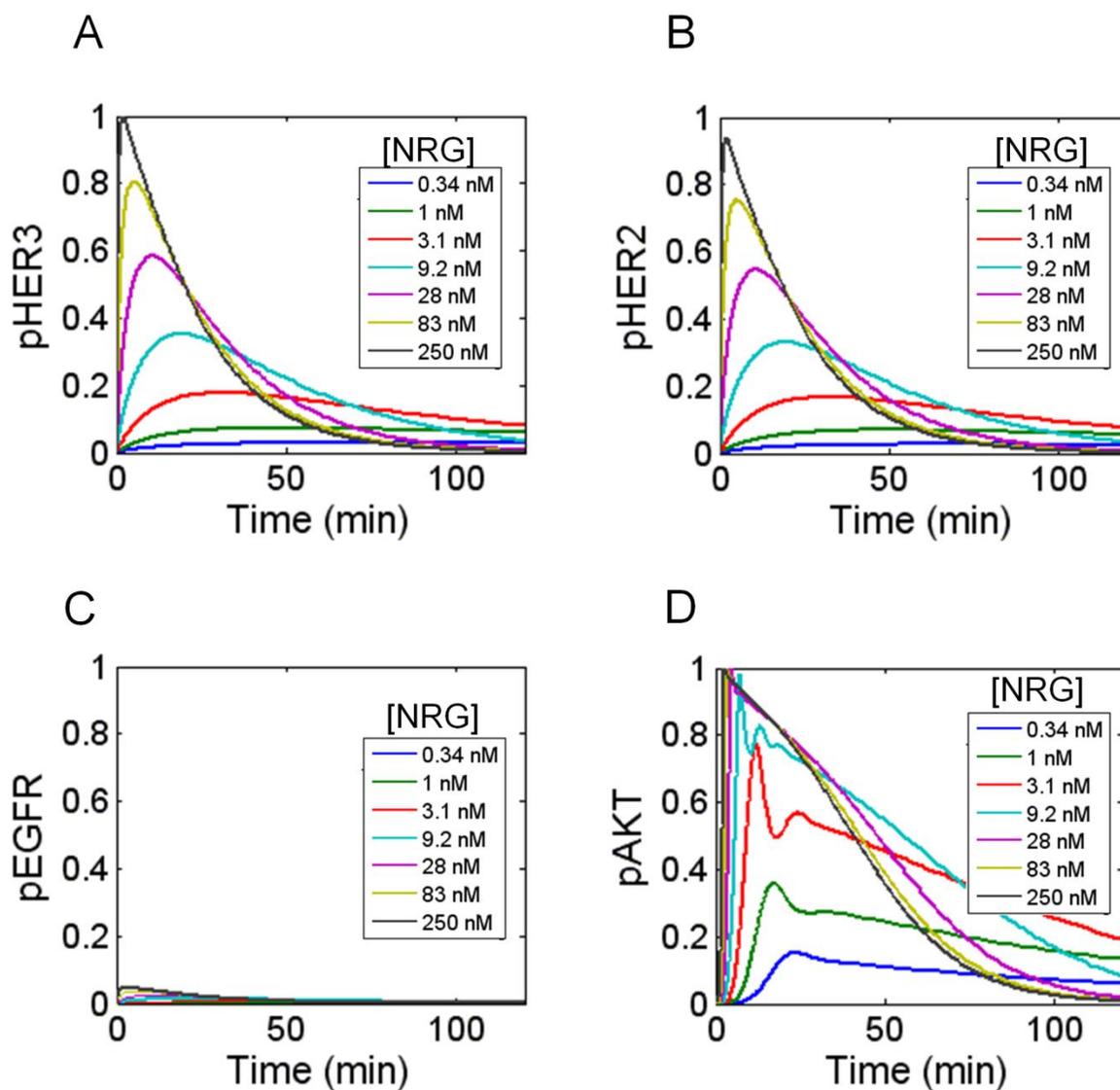


Figure S7. Time course plots for (A) pHER3, (B) pHER2, (C) pEGFR and (D) pAKT for various concentrations of the HER3 ligand NRG-1 β . For each phosphorylated ErbB species, data was normalized to the maximum pHER3 signal observed, to facilitate comparison of the RTK activation levels. For pAKT, data was normalized to the maximum pAKT signal observed.

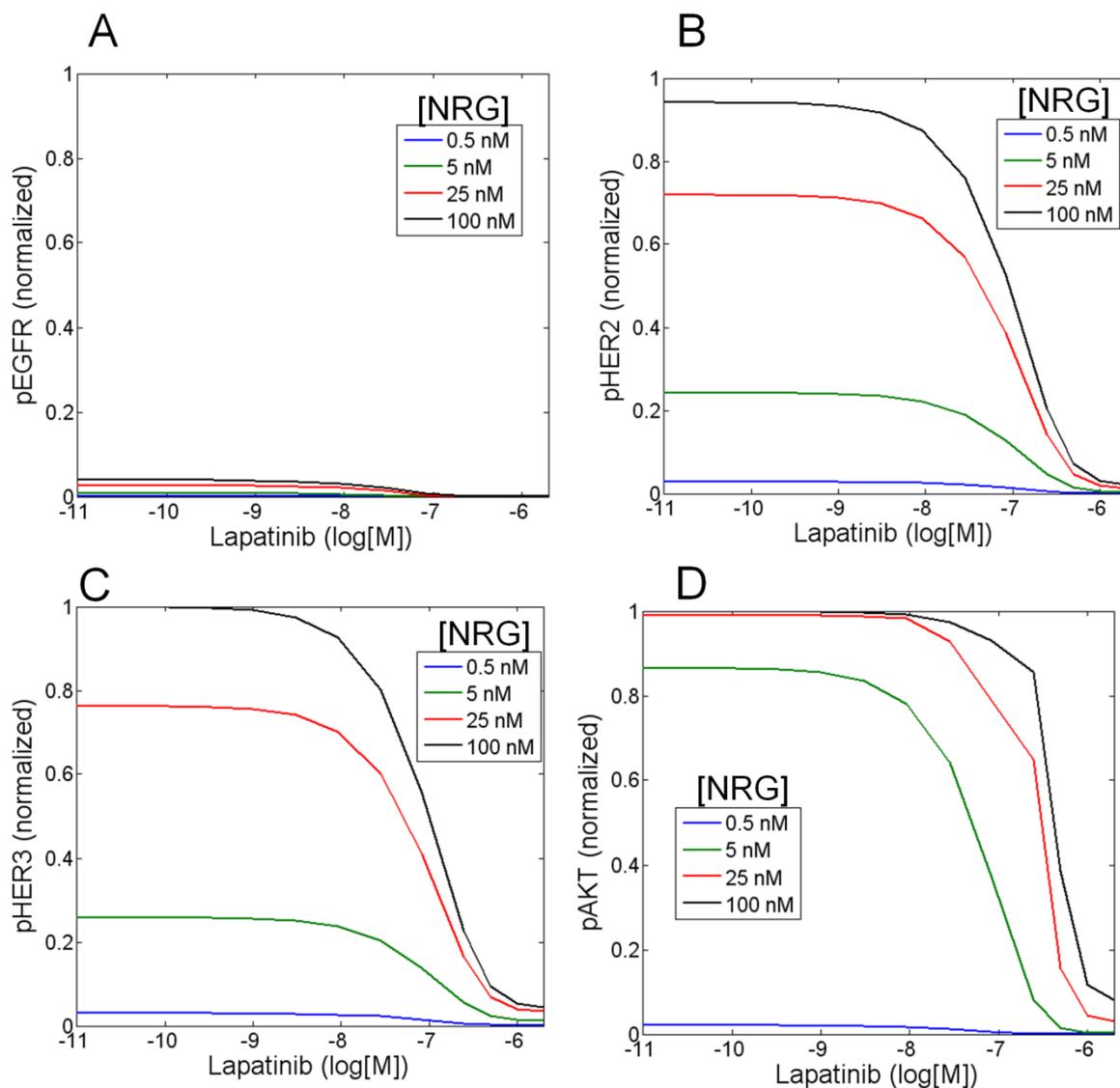


Figure S8. Dose-response curves of lapatinib treatment in the HER3 signaling model. The response to the TKI was computed following a 30 minute pre-incubation with lapatinib and 10 min stimulation with increasing concentrations of NRG1-β. Results for pEGFR, pHER2 and pHER3 were normalized to the no-inhibitor control value for 100 nM pHER3 to facilitate comparison of the profiles for the three ErbB kinases. Results for pAKT were normalized to the no-inhibitor control value for 100 nM pAKT.

Table S1. Kinetic parameters for the HER3 signaling model

All parameters are derived from Reference ⁶, except for k30_weak, which is derived from Reference ⁷. First- and second-order rate constants are given in units of sec⁻¹ and molecules⁻¹ sec⁻¹, respectively. The parameter names refer to the SBML model, which is provided as a Supplementary file.

| Description | Value | Name in SBML model |
|---------------------------------------------------|----------------------|--------------------|
| HER3-NRG binding | 5×10^{-11} | kf1 |
| HER3-NRG dissociation | 0.001 | kr1 |
| HER3-2 binding to NRG | 5×10^{-11} | kf2 |
| HER3-2 dissociation from NRG | 0.001 | kr2 |
| HER3-NRG binding to HER2 or HER3 | 3×10^{-6} | kf7 |
| HER3-NRG dissociation from HER2 | 0.001 | kr7 |
| HER3-NRG binding to EGFR | 3×10^{-8} | kf10 |
| Constitutive dimerization | 4.2×10^{-9} | kf12 |
| Constitutive dimerization | 0.001 | kr12 |
| Phosphorylation of ErbB dimers | 1 | kf30 |
| Phosphorylation of HER3 homodimers | 0.001 | kf30_weak |
| Binding of ErbB phosphatase to ErbB dimers | 5×10^{-6} | kf38 |
| Dissociation of ErbB phosphatase from ErbB dimers | 0.1 | kr38 |
| ErbB receptor dephosphorylation | 1 | kf45 |
| PI3K binding to HER3 dimers | 3×10^{-6} | kf52 |
| PI3K dissociation from HER3 dimers | 0.1 | kr52 |
| PI3K binding to non-HER3 dimers | 7.5×10^{-7} | kf54 |
| PI3K dissociation from non-HER3 dimers | 0.1 | kr54 |
| PIP2 binding to HER3 dimers | 5×10^{-6} | kf59 |
| PIP2 dissociation from HER3 dimers | 0.1 | kr59 |
| PIP2 binding to non-HER3 dimers | 5×10^{-7} | kf61 |
| PIP2 dissociation from non-HER3 dimers | 0.1 | kr61 |
| PIP3 activation by HER3 dimers | 0.2 | kf66 |
| PIP3 activation by non-HER3 dimers | 0.013 | kf68 |
| PIP3-PTEN binding | 5×10^{-6} | kf73 |
| PIP3-PTEN dissociation | 0.1 | kr73 |
| PIP3 inactivation | 0.1 | kf74 |
| PIP3-AKT or PIP3-pAKT binding | 2.6×10^{-4} | kf75 |
| PIP3-AKT or PIP3-pAKT dissociation | 0.1 | kr75 |
| PDK1 binding to PIP3-AKT | 6.7×10^{-5} | kf76 |
| PDK1 dissociation from PIP3-AKT | 0.1 | kr76 |
| Phosphorylation of AKT | 1 | kf77 |
| PDK1-PIP3 dissociation | 0.2 | kf78 |
| Phosphorylation of pAKT | 1 | kf81 |
| PP2A-pAKT binding | 1.7×10^{-6} | kf82 |
| PP2A-pAKT dissociation | 0.1 | kr82 |
| AKT dephosphorylation | 1.5 | kf83 |
| ppAKT-PP2Aoff binding | 8.3×10^{-9} | kf86 |
| ppAKT-PP2Aoff dissociation | 0.5 | kr86 |
| Activation of PP2Aoff | 0.1 | kf87 |
| Internalization of ligand-bound monomers | 0.1 | kf88 |
| Recycling of ligand-bound monomers | 0.005 | kr88 |

| | | |
|----------------------------------------------------|------------------------|--------|
| Internalization of HER3-2, HER2-2, HER3-3 dimers | 0.005 | kf93 |
| Recycling of HER3-2, HER2-2, HER3-3 dimers | 0.005 | kr93 |
| Internalization of HER3-1 dimers | 0.005 | kf98 |
| Recycling of HER3-1 dimers | 0.005 | kr98 |
| NRG endosomal binding | 0.038 | kf127 |
| Ligand degradation | 0.002 | kf185 |
| Degradation of ligand-bound monomers | 0.002 | kf187 |
| Degradation of HER2 dimers and HER3-3 dimers | 0.002 | kf192 |
| Degradation of EGFR heterodimers | 0.002 | kf201 |
| Lapatinib binding to EGFR or EGFR-HER3 | 6.4×10^{-12} | kflap1 |
| Lapatinib dissociation from EGFR or EGFR-HER3 | 3.83×10^{-5} | krlap1 |
| Lapatinib binding to HER2 or HER2-3 | 1.5×10^{-12} | kflap2 |
| Lapatinib dissociation from HER2 or HER2-3 | 3.83×10^{-5} | krlap2 |
| Lapatinib binding to EGFR-EGFR or EGFR-HER2 dimers | 1.28×10^{-11} | kflap3 |
| Lapatinib binding to HER2 homodimers | 3×10^{-12} | kflap4 |