

Calculating transfer entropy from variance-covariance matrices provides insight into allosteric communication in ERK2

*Luisa Garcia Michel, Clara Keirns, Benjamin Ahlbrecht, and Daniel A. Barr**

Department of Chemistry, University of Mary, Bismarck, ND 58504

*corresponding author: dabarr@umary.edu

Abstract

Transfer entropy methods provide an approach to understanding asymmetric information flow in coupled systems, with particular application to understanding allosteric interactions in biomolecular systems. Transfer entropy analysis holds the potential to reveal pathways or networks of residues that are coupled in their information flow and thus give new insights into folding and binding dynamics. Most current methods for calculating transfer entropy require very long simulations and almost equally long calculations of joint probability histograms to compute the information transfer that make these methods either functionally intractable or statistically unreliable. Available approximate methods based on graph and network theory approaches are rapid but lose sensitivity to the chemical nature of the biomolecules and thus are not applicable in mutation studies. We show that reliable estimates of the transfer entropy can be obtained from the variance-covariance matrix of atomic fluctuations, which converges quickly and retains sensitivity to the full chemical profile of the biomolecular system. We validate our method on ERK2, a well-studied kinase involved in the MAPK signaling cascade for which considerable computational, experimental, and mutation data are available. We present the results of transfer entropy analysis on data obtained from molecular dynamics simulations of wild type active and inactive ERK2, along with mutants Q103A, I84A, L73P, and G83A. We show that our method is consistent with the results of computational and experimental studies on ERK2, and we provide a method for interpreting networks of interconnected residues in the protein from a perspective of allosteric coupling. We introduce new insights about possible allosteric activity of the extreme N-terminal region of the kinase, which to date has been under-explored in the literature and may provide an important new direction for kinase studies. We also describe evidence that suggests activation may occur by different paths or routes in different mutants. Our results highlight systematic advantages and disadvantages of each method for calculating transfer entropy and show the important role of transfer entropy analysis for understanding allosteric behavior in biomolecular systems.

Introduction

Transfer entropy was first proposed by Thomas Schreiber in 2000 as an asymmetric approach to understanding information flow in coupled systems.¹ Since that original paper, transfer entropy methods have been used to describe thermodynamic systems,² to suggest or infer causal relationships among data sets,^{3–5} and to measure information flow in biological,⁶ neurological,⁷ social,⁸ and economic systems.⁹ While transfer entropy analyses hold great promise for extracting significant information from complex data sets, there remain important concerns with both methodology and interpretation, particularly in biomolecular systems.

One of the greatest challenges with current transfer entropy approaches is that they require extremely large data sets to populate sparse multidimensional probability histograms; in theory, it is desirable to have three times as many samples as possible state configurations, though this is typically unrealistic in practice.¹⁰ In order to ameliorate the impacts of (inevitable) undersampling, several methods are used to ensure that the relationships extracted from the data are statistically significant. In one approach, the time series can be shuffled many times and the (average) transfer entropy calculated from these shuffled time series is subtracted from the original transfer entropy to obtain the effective transfer entropy.¹¹ In other methods, random statistical noise is introduced into the calculations to determine if the calculated transfer entropy rises above the level of background noise.¹² In either case, the methods used to avoid undersampling may result in missing key relationships in the data, chiefly because the choice of method requires making assumptions about the underlying structure of the data, which is often not known *a priori*.¹²

In biomolecular systems particularly, these limitations on transfer entropy methodologies have important consequences. In order to gather sufficient statistics to satisfactorily populate the multidimensional histograms required for reliable probability analysis, very long simulations (often microsecond timescale) are required. Even with the increase in available computational tools and GPU speedups, such simulations are still not feasible for most proteins as a matter of routine. The other currently-available option is to use approximate methods derived from graph theory and the Gaussian Network Model (GNM) to model the dynamics of the protein and estimate the transfer entropy.^{5,6,13} One significant drawback of these approximate methods is that, because there is no sensitivity to side chain identity or motions, there is limited utility for these methods in mutation studies, which are often of great interest in biomolecular systems.^{5,14}

Here we demonstrate a method that balances these two considerations: one that is fast enough for practical use but sensitive enough to the chemical nature of the protein to be capable of studying the effects of amino acid mutations on the information transfer and allosteric behavior of the protein. In this method, the transfer entropy is extracted from the variance-covariance matrix derived from short simulations (which converge on a timescale of 5-20ns¹⁵) and computed using the formalism of the dynamic Gaussian Network Model (dGNM), since the covariance matrix is known to be inversely proportional to the contact (Kirchoff) matrix used as the basis for the GNM method.⁵ Because our method is based on the covariance matrix of protein fluctuations (for which there are robust methods to calculate the convergence¹⁶), we avoid the sampling and convergence issues that plague many current methods, as well as the need for trajectory shuffling or the introduction of statistical noise.

We apply our method to the well-studied extra-cellular signal-regulated kinase-2 (ERK2, shown in Figure 1) and a series of known ERK2 mutants^{16,17} and demonstrate that our method is consistent with important available experimental and theoretical data.¹⁸⁻²⁰ Regions and mutated residues of interest are shown in Figure 1. Three structural changes occur upon activation^{16,21} (induced by phosphorylation at T183 and Y185): helix C (residues 60-75) directs closing of the N-domain over the C-domain, the activation lip (residues 170-185) refolds, and loop L16 (329-335) folds to expose a hydrophobic leucine zipper motif that provides the dimerization interface for the activated kinase. The D site (residues 108, 113-123, 155-162, 315-320) and F site (residues 197-200, 229-235, 257-262) include direct substrate binding specificity.^{18,22} The DFG site (165-

167) is a conserved kinase site that is targeted by type II inhibitors but plays an unclear role in ERK2.^{20,22} Three autoactivating ERK2 mutants in this study are the gatekeeper residue Q103A, along with I84A and L73P.¹⁷ A fourth mutant, G83A, has been shown to prevent activation of the kinase by inhibiting domain closure.¹⁶

We note that, in some kinases at least, the binding effects of inhibitors cannot be determined solely from structural data, a phenomenon believed to be due to conformational trapping or averaging in the X-ray structures; in fact, inhibitors that bind in the same mode at the same site have been shown to have opposite effects on the ERK2.²⁰ Studies sensitive to the dynamic nature of the structural ensemble are required to elucidate such effects. By comparing our method in context with a variety of approaches for the study of allosteric communication in proteins,²³ we are able to provide new insights about the strengths and weaknesses of various methods and shed new light on the role of the N-terminal beta-sheet region in the communication pathways of ERK2.

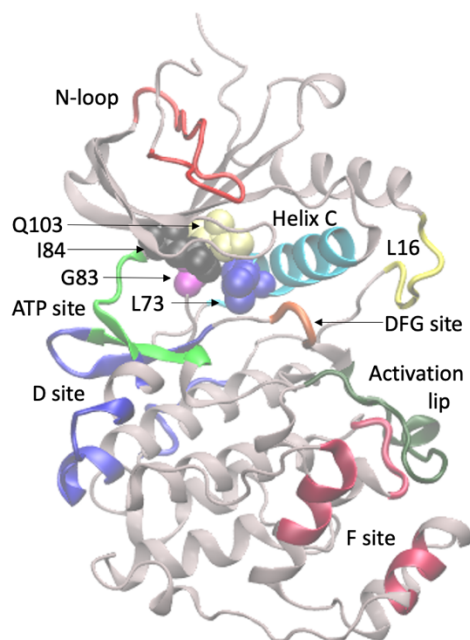


Figure 1: Structure of inactive ERK2 showing important binding sites and allosteric regions (see text) in colored cartoons and mutated residues Q103, I84, L73, and G83 in space-filling spheres.

Methods

We base our calculation of the transfer entropy on approach developed in the dynamic Gaussian Network Method,^{5,6,13} which calculates the transfer entropy according to the equations:¹³

$$\langle \Delta R_i(t) \Delta R_j(t + \tau) \rangle = \sum_k A_{ij}(k) \exp\{-\lambda \tau / \tau_0\} \quad (1)$$

$$A_{ij}(k) = \lambda_k^{-1} u_i^{(k)} u_j^{(k)} \quad (2)$$

$$\begin{aligned}
T_{i \rightarrow j}(\tau) = & \frac{1}{2} \ln \left(\left(\sum_k A_{jj}(k) \right)^2 - \left(\sum_k A_{jj}(k) \exp\{-\lambda_k \tau / \tau_0\} \right)^2 \right) \\
& - \frac{1}{2} \ln \left[\left(\sum_k A_{ii}(k) \right) \left(\sum_k A_{jj}(k) \right)^2 \right. \\
& + 2 \left(\sum_k A_{jj}(k) \right) \sum_k A_{jj}(k) \exp\{-\lambda_k \tau / \tau_0\} \sum_k A_{ij}(k) \exp\{-\lambda_k \tau / \tau_0\} \\
& - \left\{ \left(\sum_k A_{ij}(k) \exp\{-\lambda_k \tau / \tau_0\} \right)^2 + \left(\sum_k A_{ij}(k) \right)^2 \right\} \left(\sum_k A_{jj}(k) \right) \\
& - \left. \left(\sum_k A_{jj}(k) \exp\{-\lambda_k \tau / \tau_0\} \right)^2 \left(\sum_k A_{jj}(k) \right) \right] - \frac{1}{2} \ln \left[\left(\sum_k A_{ij}(k) \right) \right] \\
& + \frac{1}{2} \ln \left(\left(\sum_k A_{jj}(k) \right) \left(\sum_k A_{ij}(k) \right) - \left(\sum_k A_{jj}(k)^2 \right) \right)
\end{aligned} \tag{3}$$

Where λ and u are the eigenvalues and eigenvectors, respectively, that result from diagonalization of the Kirchhoff matrix that defines the protein network of contacts held together by springs. The time delay required for evaluating transfer entropy is introduced by the τ parameter which acts as an effective lag time operating on the eigenvectors.¹³ In our method, we replace the Kirchhoff matrix with the variance-covariance matrix of fluctuations calculated according to the equation:²⁴

$$C_{ij} = \frac{\langle x_i x_j \rangle}{\sqrt{\langle x_i^2 \rangle \langle x_j^2 \rangle}} = \frac{\langle \Delta R_i \Delta R_j \rangle}{\sqrt{\langle \Delta R_i^2 \rangle \langle \Delta R_j^2 \rangle}}$$

In order to meaningfully compare transfer entropy data from the different methods, we compute the normalized directional transfer entropy:

$$NTE_{i \rightarrow j}(\tau) = \frac{T_{i \rightarrow j}}{S(\Delta R_j(t + \tau) \mid \Delta R_j(t))} \in [0,1]$$

$$DNTE_{i \rightarrow j} = NTE_{i \rightarrow j} - NTE_{j \rightarrow i} \in [0,1]$$

Intuitively, the normalized transfer entropy (NTE) tells us the proportion of uncertainty reduced regarding future fluctuations of residue j given the past fluctuations of residue i and j . An $NTE=1$ implies that i perfectly predicts j . An $NTE=0$ implies that i is statistically independent of

j or that the future values of j are perfectly predicted by its own past. That is, all the information regarding future fluctuations of j is already encoded within its own past, so i has no uncertainty to reduce. The normalized directional transfer entropy (DNTE) provides the magnitude and direction of information flow. A DNTE of 1 implies the flow of information is entirely dominated by residue i , while a DNTE of -1 would imply the converse. In the case where DNTE is 0, analysis of the individual normalized transfer entropies is required to alleviate whether the information flow is bidirectional or absent.

We note two fundamental assumptions in the dGNM approach. The first, that the protein can be accurately represented as a network of spring contacts, does not apply to our method since we derive our input from the covariance matrix built from molecular dynamics simulations, allowing us to bypass the graph or network theory assumptions altogether. The second assumption treats the fluctuations of protein atoms as harmonic and following a Gaussian distribution, which is rigorously true for the GNM method, since by definition the Gaussian Network Model starts with a low-energy structure and explores only harmonic fluctuations in the very local phase space around that configuration.²⁵ The Gaussian assumption is true for our method in the limit of the quasi-harmonic approximation^{26,27} which is reliable when the protein only visits a single state or energy well during the simulation, but can fail dramatically if multiple states or energy minima are visited, especially if the full configurational space of each state is not completely explored during the simulation.²⁸

Simulation data was obtained from previously reported work on ERK2.¹⁶ 20ns simulations of ERK2 in the inactive state²⁹ and active state²¹ as well as the mutants Q103A, I84A, L73P, and G83A were used as the basis for analysis by both the covariance-based and the full histogram transfer entropy methods. All simulations were performed in explicit solvent with counterions and snapshots from each simulation were saved every 1ps for analysis.

The full statistical transfer entropy analysis was performed on the ERK2 trajectories using the method of Kamberaj and van der Vaart¹¹ as reported in previous work.¹⁶ The variance-covariance matrix of protein fluctuations was calculated for CA atoms using Carma.³⁰ Preparatory calculations for the dGNM method were made in python using ProDy³¹ to parse the pdb files and set the atom definitions, as well as to calculate the Kirchhoff matrix. Calculation of the transfer entropy for the dGNM method and our new covariance-based method was performed following reference⁵ using python. Our method follows the form of the dGNM calculation entirely, with the inverse of the covariance matrix used instead of the Kirchhoff matrix. We use a time delay $\tau = 5$ following the recommendation of reference⁵ and consistent with cross-correlation delay times in molecular dynamics simulations. We note that $\tau = 5$ typically maximizes the transfer entropy obtained from our method (using larger values of τ gives negligible differences in the results); smaller values of τ tend to underestimate the transfer entropy, though not in any way that would alter the analysis — the peak shapes and patterns are consistent for all values of τ we tested in the range of 1–50, and in no case did we observe changing the value of τ to result in the switch of any residue from driving to responding or vice versa (Figure 2). We note that this is not necessarily the case with the original dGNM method, where in addition to choosing the value for τ , there is a cutoff parameter that needs to be chosen for each system, and different values for the cutoff used to define the Kirchhoff matrix can have a significant impact on the calculated transfer entropy, including switching which residues/regions appear to be driving or responding (Figure 2).

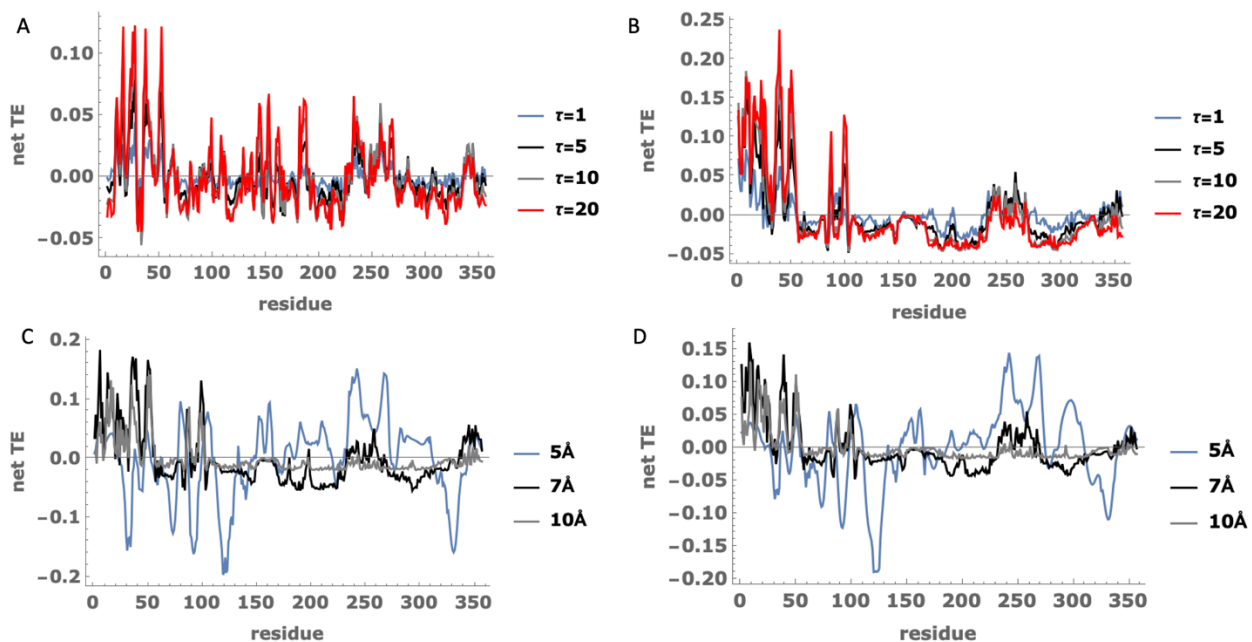


Figure 2: Net transfer entropy out from each residue calculated from the covariance method (A) and dGNN method (B) for the active state of ERK2 as a function of increasing τ value. Net transfer entropy out from each residue calculated from the dGNN method for inactive (C) and active (D) ERK2 as a function of increasing cutoff distance used in creating the Kirchhoff matrix.

Transfer entropy analysis was performed on the directional matrix $DNTE_{i \rightarrow j} = NTE_{i \rightarrow j} - NTE_{j \rightarrow i}$, where positive values indicate that the residue on the vertical axis drives the residue on the horizontal axis. The most driving residues were identified in each state as the peak with the maximum $NTE_{i \rightarrow \text{out}}$ calculated as the sum over the i^{th} row of the directional matrix. Pathways of information flow were followed by sorting the directional matrix row-wise to determine the residue j to which residue i sends the most information, allowing for a directional analysis of information flow. All pathways eventually formed a repeating transfer entropy loop. Pathways were often remarkably consistent within individual states but varied between different methods and states of the kinase.

Data for the betweenness of the nodes in ERK2 were obtained from the MCPATH server³² using PDB codes 1ERK²⁹ and 2ERK²¹ to identify functional residues in the inactive and active states, respectively, with a path length of 100,000. The mutational and configurational frustration of the residues in ERK2 was calculated for the same PDB codes using the frustratometer server³³ with default settings. In cases where residues were missing from the PDB files, values of zero were used at those positions to keep numbering consistent for analysis.

Results and Discussion

Our method is consistent with previous theory and experiment and provides new mechanistic insights

We used two approaches to analyze and compare the transfer entropy results from the three methods tested in this study. We identified the top twenty most driving residues in the protein in each state by each method (Table 1), and checked for consistency among the methods in identifying the same driving regions. We note that in this analysis, the full simulation method seems on the face of it to give remarkably different results from the other two methods; however, when we realized that many peaks are in the same immediate region, and expanded our analysis to look at the top 4-5 most driving “regions” of the kinase, rather than a fixed residue count, we found that in fact the methods are in good agreement, as visual inspection of Figure 3 shows. We note that the 7 most driving residues in the inactive state for our method are also shown as strongly driving in the full-histogram and dGNM method. Almost half of the residues shown in the active state are also seen in at least one of the other two methods.

Inactive State	Active State	G83A	Q103A	L73P	I84A
MET 36 *dh	TYR 28 **	CYS 38 **	GLN 64 *	LEU 341 *	ARG 259 **
VAL 37 *+dh	PHE 17 d	MET 36 *	ARG 68	GLU 343	ARG 65
GLN 15 dh	LYS 53 **	VAL 37 **	MET 36 *	PHE 346	LEU 144
LYS 53 *+d	CYS 38 *+d	ALA 40 +	CYS 63 *	THR 179	ASN 260
VAL 16 *+dh	LEU 26 +	TYR 41 +	ARG 65	SER 200	VAL 171
CYS 38 *+dh	THR 24 +	GLU 31 *	LYS 53 **	LYS 342	GLN 64 *
TYR 28 *+h	ARG 259 **	LYS 53 **	GLU 31 *	ASN 199	VAL 143
GLY 14 d	GLN 234 **	ILE 51 +	THR 61	ALA 5	HSD 145
GLN 64 *	VAL 16 *+d	ILE 54	LEU 67	LYS 201	ARG 170
GLN 234 **	SER 27 **	LEU 341 *	TYR 41 +	ALA 4	CYS 63 *
ARG 259 **	GLU 10 dh	LEU 26 +	VAL 37 **	ASP 177	THR 61
LEU 154 *+	LYS 52 +	LEU 113	GLN 60	LYS 340 *	ASN 236
LEU 341 *	ILE 51 +	THR 24 +	ARG 259 **	LEU 344	LEU 235
GLU 31 *h	MET 11 *+	TYR 28 **	TYR 62	LEU 161	TYR 62
ALA 6 *d	TYR 41 +dh	LEU 153	LYS 52 +	ILE 345	TYR 356
CYS 63 *	ALA 40 +d	ASP 42 +	GLN 234 **	GLY 180	GLU 69
ILE 29 h	ASP 42 +d	ASN 25	THR 24 +	LEU 154 *+	THR 66
LYS 340 *	VAL 37 *+d	GLN 64 *	ALA 258	ALA 3	HSD 139
SER 27 *+	SER 39 d	ASP 233	ALA 40 +	PHE 181	GLN 234 *+
MET 11 *+h	LEU 154 *+	HSD 237	LEU 26 +	ALA 6 *	ARG 357

Table 1: Top 20 most driving residues in each state of ERK2 for our new method. Residues in important regions of the kinase are highlighted in colors representing the N-loop (purple), ATP binding site (green), helix C (orange), activation lip (blue), and D site (red). Residues that are known to be of particular importance for protein function are shown in bold. For purposes of comparison between states, residues which appear in the inactive state are marked in all states with an asterisk, and residues which appear in the active state are marked with a plus sign. Residues which appear in the top 20 most driving of the dGNM and full histogram methods are marked with a *d* and an *h*, respectively.

Both Table 1 and Figure 3 show strong driving behavior from the N-loop in both active and inactive states. Residues from helix C are seen as strong drivers in the inactive state, mutants G83A, Q103A, and I84A. From this region, residues Q64 (important role in domain restructuring), R65 (important for auto-activation), R68 (prevents refolding), and salt-bridge residue Q69 are seen as drivers of information. Salt-bridge residue K52 is a strong driver in the activate state and in mutant Q103A. Residues L153 and L154 are strong drivers of information from the ATP binding site. Mutation of residues L73 and I84 also seems to increase the driving behavior of the activation lip. Of particular interest, residue R170 shows driving behavior for mutant I84A. Literature suggests R170 is in contact with D334 along with R68 to prevent from refolding of loop L16 in the inactive state.¹⁶ Mutant L73P is the only state that shows driving behavior from the D site.

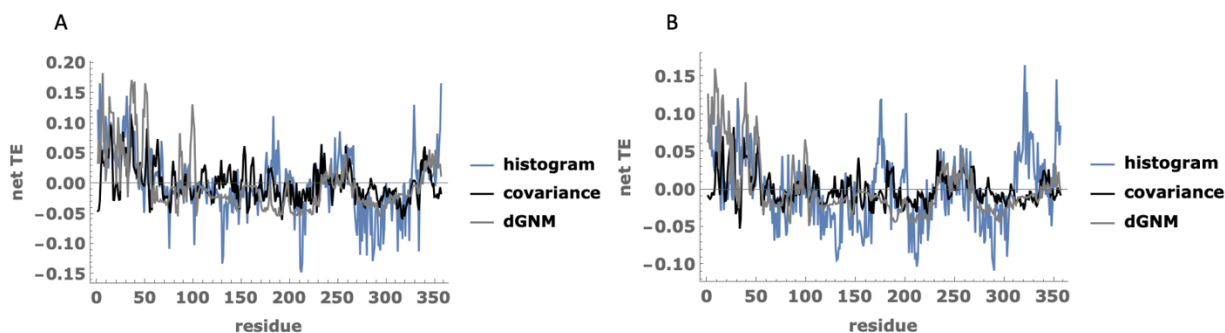


Figure 3: Net transfer entropy out from each residue calculated from the full histogram (blue), covariance (black), and dGNM (gray) methods for inactive (A) and active (B) states of ERK2.

We also compared the methods on the basis of the information flow pathways they produced from these top 20 most driving residues. From each pathway, we looked to see whether key regions of the kinase were coupled in an information-flow relationship. We note that in both our method and the dGNM method, the DFG site appears in information flow pathways of the inactive state as responding to helix C (both methods), the ATP binding site (dGNM method), and the D site (our method).

From our pathways analysis of active state ERK2, we see that the ATP site drives the motions of the D site (and the F site in the dGNM method). In the mutant pathways, we also observe driving behavior from the ATP site to the activation lip (Q103A) and loop L16 (G83A). We note, consistent with our previous observations of the crucial role of G83 for the domain closure and activation of the kinase,¹⁶ that residue 83 is found in many information flow pathways for the active state of ERK2 and all mutants except L73P.

Of particular interest is the L73P mutant, where pathways of information flow show that the F site drives ATP site and D site, and that loop L16 and the activation lip drive ATP and D site; these pathways are markedly different from those seen in all other states and suggest that the L73P mutant may activate by a different mechanism than the other mutants. This is consistent with previously-reported data that the L73P mutant shows greater fluctuations at the beginning of the activation lip than other mutants,¹⁶ but our analysis provides deeper insights into the possible different pathways of autoactivation.

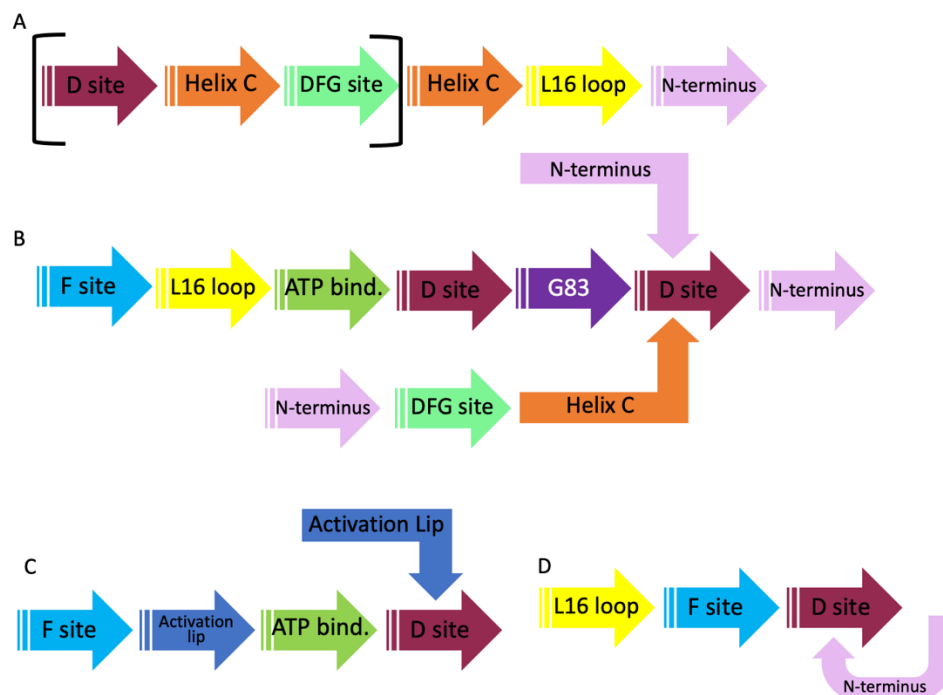


Figure 4: Pathways of information flow in the inactive state (A), active state (B), and L73P mutant (C and D). Brackets indicate branches that occur in some but not all pathways for the state.

Full atomic simulations give insight into the importance of anharmonic conformational motions

The full statistical method shows considerably stronger driving behavior for the activation lip (175-176) and loop L16 in the active state than in either of the other methods (Figure 3). These regions show significant conformational flexibility and even some degree of folding (in L16) or reorganization (in the activation lip) during the course of simulations of even moderate length. It is not expected that such motions would be well-described by a quasi-harmonic analysis,²⁸ and this is reflected in the transfer entropy descriptions of these regions. What is less clear, however, is whether the simulations are sufficiently long for the full statistical treatment of the transfer entropy to be converged; our pathways analysis suggests not. We believe that the relative heights of the peaks (and thus the pathways generated by following the most driving residues) may not be converged in this method at this timescale, and that as a result the finer details of information flow in these systems cannot be resolved. However, we highlight again the importance of our observations that the first-level analysis (of the regions that are most strongly driving or receiving) is consistent across all three transfer entropy methods studied in this work.

Regions of high structural frustration show inconsistencies in transfer entropy calculations

Apart from the differences in the transfer entropy of the highly flexible regions mentioned above, the greatest discrepancy among the methods occurs around residue 100, where the dGNM method appears to dramatically overestimate the transfer entropy. This region is

highlighted by other network analysis methods as being highly constrained/frustrated (Figure 5). We note that the dGNM method tends to overestimate the TE of these regions compared to simulation-based methods. We note that the covariance method shows driving behavior for the phosphorylated residues on the activation lip that are missed by the dGNM method (residues 180-190).

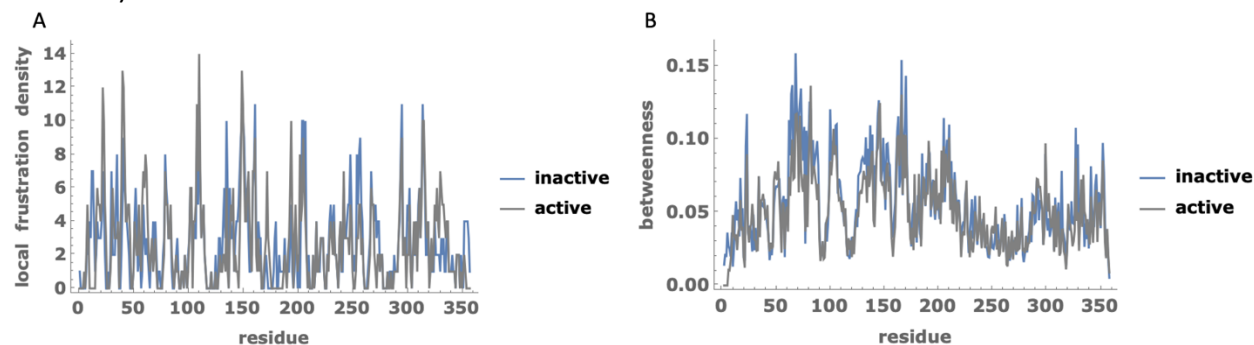


Figure 5: (A) Number of highly frustrated (mutational) contacts within a 5Å sphere of each residue in ERK2. (B) Betweenness calculated for each residue in ERK2 from the MCPath server.

All methods show the importance of the N-terminal region for information flow in ERK2

We note that in all three methods, residues in the N-terminal region show strong driving and responding behavior and couple with the substrate binding sites in many pathways. While there has been some evidence of involvement of in this region (particularly the Gly-rich loop) reported in the literature,²⁰ it has been attributed largely to instability in the active site and the activation loop. We believe that there may be significant allosteric interactions taking place that involve the extreme N-terminal domain of the kinase that interact with the substrate binding sites and might be of interest for further studies in inhibitor binding and dynamics.

Conclusions

We demonstrate that using the covariance matrix of protein fluctuations leads to identification of the same important driving and responding residues, and most importantly, the same allosteric interactions as the full statistical method. We note that there are fewer adjustable parameters in this method than in the dGNM method, and that this method is sensitive to dynamical changes in mutant proteins, whereas dGNM is not. Our method provides rapid, reliable insight into the allosteric behavior of ERK2 at reasonable computational cost.

Importantly, by comparing all three methods to one another in a single, well-studied system, we provide insights about where each method is likely to become unreliable for transfer entropy calculations in biological systems. The dGNM method may overestimate the transfer entropy of highly frustrated regions of the protein network. The covariance method may underestimate the transfer entropy of highly flexible regions of proteins where the quasi-harmonic approximation fails to hold. In such cases, the full statistical method is likely the best option, but very long simulation timescales are likely to be needed to achieve good convergence and the ability to obtain detailed interpretation of information transfer pathways.

Finally, by combining a variety of approaches for the study of allosteric behavior in ERK2, we are able to provide insights about the structure and dynamics of the kinase that may help to guide future experiments and drug design. We note especially the consistent importance of the N-terminal region in our analysis and its regular coupling with the substrate binding sites. We expect this to be of interest in future studies of inhibitor binding and dynamics.

Acknowledgements

The authors declare no competing financial interest. The authors gratefully acknowledge Aysima Hacısuleyman and Burak Erman for providing their Matlab code for use in this study and for several insightful conversations. We thank Baylee Grinsteiner, Mary Claire Talbot, Autumn Bertch, and Cassandra Boehm for helpful discussions and assistance with calculations. DB thanks James Peliska and David Ronderos for consultation and guidance.

Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103442.

References:

- (1) Schreiber, T. Measuring Information Transfer. *Phys. Rev. Lett.* **2000**, *85* (2), 461–464. <https://doi.org/10.1103/PhysRevLett.85.461>.
- (2) Prokopenko, M.; Lizier, J. T.; Price, D. C. On Thermodynamic Interpretation of Transfer Entropy. *Entropy* **2013**, *15* (2), 524–543. <https://doi.org/10.3390/e15020524>.
- (3) Ay, N.; Polani, D. Information Flows in Causal Networks. *Adv. Complex Syst.* **2008**, *11* (01), 17–41. <https://doi.org/10.1142/S0219525908001465>.
- (4) Lizier, J. T.; Prokopenko, M. Differentiating Information Transfer and Causal Effect. *Eur. Phys. J. B* **2010**, *73* (4), 605–615. <https://doi.org/10.1140/epjb/e2010-00034-5>.
- (5) Hacısuleyman, A.; Erman, B. Causality, Transfer Entropy, and Allosteric Communication Landscapes in Proteins with Harmonic Interactions. *Proteins Struct. Funct. Bioinforma.* **2017**, *85* (6), 1056–1064. <https://doi.org/10.1002/prot.25272>.
- (6) Hacısuleyman, A.; Erman, B. Entropy Transfer between Residue Pairs and Allostery in Proteins: Quantifying Allosteric Communication in Ubiquitin. *PLOS Comput. Biol.* **2017**, *13* (1), e1005319. <https://doi.org/10.1371/journal.pcbi.1005319>.
- (7) Camacho, M.; Romeu, A.; Ruiz-Marin, M. Symbolic Transfer Entropy Test for Causality in Longitudinal Data. *Econ. Model.* **2020**. <https://doi.org/10.1016/j.econmod.2020.02.007>.
- (8) Borge-Holthoefer, J.; Perra, N.; Gonçalves, B.; González-Bailón, S.; Arenas, A.; Moreno, Y.; Vespignani, A. The Dynamics of Information-Driven Coordination Phenomena: A Transfer Entropy Analysis. *Sci. Adv.* **2016**, *2* (4), e1501158. <https://doi.org/10.1126/sciadv.1501158>.
- (9) Korbel, J.; Jiang, X.; Zheng, B. Transfer Entropy between Communities in Complex Financial Networks. *Entropy* **2019**, *21* (11), 1124. <https://doi.org/10.3390/e21111124>.
- (10) Kamberaj, H. *Tools for Statistical Analysis of Molecular Dynamics Simulation Data*; International Balkan University: Skopje, Republic of Macedonia, 2017.
- (11) Kamberaj, H.; van der Vaart, A. Extracting the Causality of Correlated Motions from Molecular Dynamics Simulations. *Biophys. J.* **2009**, *97* (6), 1747–1755. <https://doi.org/10.1016/j.bpj.2009.07.019>.
- (12) Wollstadt, P.; Lizier, J.; Vicente, R.; Finn, C.; Martinez-Zarzuela, M.; Mediano, P.; Novelli, L.; Wibral, M. IDTxL: The Information Dynamics Toolkit XL: A Python Package for the Efficient Analysis of Multivariate Information Dynamics in Networks. *J. Open Source Softw.* **2019**, *4* (34), 1081. <https://doi.org/10.21105/joss.01081>.
- (13) Hacısuleyman, A.; Erman, B. Entropy Transfer and Allosteric Communication in Proteins. *J. Drug Des. Res.* **2017**, *4* (5), 1054.
- (14) Kamberaj, H. A Theoretical Model for the Collective Motion of Proteins by Means of Principal Component Analysis. *Open Phys.* **2010**, *9* (1), 96–109. <https://doi.org/10.2478/s11534-010-0048-2>.
- (15) Kamberaj, H. Sampling Convergence Of Collective Motions In Proteins. *J. Appl. Phys. Sci. Int.* **2017**, 101–112.
- (16) Barr, D.; Oashi, T.; Burkhard, K.; Lucius, S.; Samadani, R.; Zhang, J.; Shapiro, P.; MacKerell, A. D.; van der Vaart, A. Importance of Domain Closure for the Autoactivation of ERK2. *Biochemistry* **2011**, *50* (37), 8038–8048. <https://doi.org/10.1021/bi200503a>.
- (17) Emrick, M. A.; Lee, T.; Starkey, P. J.; Mumby, M. C.; Resing, K. A.; Ahn, N. G. The Gatekeeper Residue Controls Autoactivation of ERK2 via a Pathway of Intramolecular Connectivity. *Proc. Natl. Acad. Sci.* **2006**, *103* (48), 18101–18106. <https://doi.org/10.1073/pnas.0608849103>.
- (18) Piserchio, A.; Warthaka, M.; Devkota, A. K.; Kaoud, T. S.; Lee, S.; Abramczyk, O.; Ren, P.; Dalby, K. N.; Ghose, R. Solution NMR Insights into Docking Interactions Involving Inactive ERK2. *Biochemistry* **2011**, *50* (18), 3660–3672. <https://doi.org/10.1021/bi2000559>.

- (19) Bardwell, A. J.; Bardwell, L. Two Hydrophobic Residues Can Determine the Specificity of Mitogen-Activated Protein Kinase Docking Interactions. *J. Biol. Chem.* **2015**, *290* (44), 26661–26674. <https://doi.org/10.1074/jbc.M115.691436>.
- (20) Pegram, L. M.; Liddle, J. C.; Xiao, Y.; Hoh, M.; Rudolph, J.; Iverson, D. B.; Vigers, G. P.; Smith, D.; Zhang, H.; Wang, W.; Moffat, J. G.; Ahn, N. G. Activation Loop Dynamics Are Controlled by Conformation-Selective Inhibitors of ERK2. *Proc. Natl. Acad. Sci.* **2019**, *116* (31), 15463–15468. <https://doi.org/10.1073/pnas.1906824116>.
- (21) Canagarajah, B. J.; Khokhlatchev, A.; Cobb, M. H.; Goldsmith, E. J. Activation Mechanism of the MAP Kinase ERK2 by Dual Phosphorylation. *Cell* **1997**, *90* (5), 859–869. [https://doi.org/10.1016/S0092-8674\(00\)80351-7](https://doi.org/10.1016/S0092-8674(00)80351-7).
- (22) Akella, R.; Moon, T. M.; Goldsmith, E. J. Unique MAP Kinase Binding Sites. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2008**, *1784* (1), 48–55. <https://doi.org/10.1016/j.bbapap.2007.09.016>.
- (23) Alakent, B.; Ince, Z. N. G. Elucidating Allosteric Communication in Proteins via Computational Methods. In *Frontiers in Computational Chemistry*; Ul-Haq, Z., Madura, J. D., Eds.; Bentham Science Publishers: Sharjah, UAE, 2017; Vol. 3, pp 260–309.
- (24) Ichiye, T.; Karplus, M. Collective Motions in Proteins: A Covariance Analysis of Atomic Fluctuations in Molecular Dynamics and Normal Mode Simulations. *Proteins Struct. Funct. Bioinforma.* **1991**, *11* (3), 205–217. <https://doi.org/10.1002/prot.340110305>.
- (25) Haliloglu, T.; Bahar, I.; Erman, B. Gaussian Dynamics of Folded Proteins. *Phys. Rev. Lett.* **1997**, *79* (16), 3090–3093. <https://doi.org/10.1103/PhysRevLett.79.3090>.
- (26) Karplus, M.; Kushick, J. N. Method for Estimating the Configurational Entropy of Macromolecules. *Macromolecules* **1981**, *14* (2), 325–332. <https://doi.org/10.1021/ma50003a019>.
- (27) Levy, R. M.; Srinivasan, A. R.; Olson, W. K.; McCammon, J. A. Quasi-Harmonic Method for Studying Very Low Frequency Modes in Proteins. *Biopolymers* **1984**, *23* (6), 1099–1112. <https://doi.org/10.1002/bip.360230610>.
- (28) Chang, C.-E.; Chen, W.; Gilson, M. K. Evaluating the Accuracy of the Quasiharmonic Approximation. *J. Chem. Theory Comput.* **2005**, *1* (5), 1017–1028. <https://doi.org/10.1021/ct0500904>.
- (29) Zhang, F.; Strand, A.; Robbins, D.; Cobb, M. H.; Goldsmith, E. J. Atomic Structure of the MAP Kinase ERK2 at 2.3 Å Resolution. *Nature* **1994**, *367* (6465), 704. <https://doi.org/10.1038/367704a0>.
- (30) Glykos, N. M. Carma: A Molecular Dynamics Analysis Program. *J. Comput. Chem.* **2006**, *27* (14), 1765–1768. <https://doi.org/10.1002/jcc.20482>.
- (31) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **2011**, *27* (11), 1575–1577. <https://doi.org/10.1093/bioinformatics/btr168>.
- (32) Kaya, C.; Armutlulu, A.; Ekesan, S.; Haliloglu, T. MCPATH: Monte Carlo Path Generation Approach to Predict Likely Allosteric Pathways and Functional Residues. *Nucleic Acids Res.* **2013**, *41* (W1), W249–W255. <https://doi.org/10.1093/nar/gkt284>.
- (33) Jenik, M.; Parra, R. G.; Radusky, L. G.; Turjanski, A.; Wolynes, P. G.; Ferreira, D. U. Protein Frustratometer: A Tool to Localize Energetic Frustration in Protein Molecules. *Nucleic Acids Res.* **2012**, *40* (W1), W348–W351. <https://doi.org/10.1093/nar/gks447>.