Presence and Utility of Intrinsically Disordered Regions in Kinases

Jaymin J Kathiriya[1,6], Ravi Ramesh Pathak[1,6], Eric Clayman[1], Bin Xue[7], Vladimir N. Uversky[3,4,5,6], and Vrushank Davé[1,2]

[1]Morsani College of Medicine, Department of Pathology and Cell Biology, [2]Department of Molecular Oncology, H. Lee Moffitt Cancer Center and Research Institute, [3]Department of Molecular Medicine, [4]USF Health Byrd Alzheimer's Research Institute, University of South Florida, Tampa, FL, 33612, [5]Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia, [6]Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia, and [7]Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, Tampa, Florida, 33620.

SUPPLEMENTARY INFORMATION CAPTIONS

Supplementary Figure S1: Predicted IDRs not seen in crystal structures

Crystal structures of 10 kinases were obtained from RCSB Protein Data Bank (http://rcsb.org/pdb/home/home.do). These structures were visualized using PyMol. The predicted IDRs of these 10 kinases overlapped with residues which were missing in the crystal structures. The blue bar represents structured regions of the kinases, while red lines represent predicted IDRs. Figures not to scale.

Supplementary Table S1: Prediction of Intrinsically Disordered Regions in the Human Kinome identifies 83% of the kinases as IDPs. DISPROT was used to predict IDRs in each of the human kinases. Original list of the human kinome was retrieved from Manning et al (2002). Only the proteins with kinase domain found on UniprotKB were considered for the analysis. Each row of the table is a unique kinase. Table key is provided at the bottom of the table. The table identifies 417 out of 504 (83%) kinases with IDRs.

Supplementary Table S2: Disease Enrichment of the 10 kinase Groups reveal Cancer has the most enriched disease. Disease enrichment profile of the 10 kinase groups using IPA core analysis reveals significance of each group (Group name – column 2 and p-value – column 3) and the molecules participating (Molecules – column 4) to enrich a particular disease or disorder (Category – column 1). This raw data was used to identify TK and AGC group as two of the most enriched kinase groups in Cancer.

Supplementary Table S3: Disordered kinases participate in 90% of KKIs. (i) The table lists total number of inter- and intra-group KKIs and (ii) prevalence of IDRs in the kinases participating in these interactions. (iii) Two sets of kinases were identified. Set A: participates in Kinase-Kinase interactions and Set B: does not participate in Kinase-Kinase interactions. DO refers to a kinase with an IDR and O refers to a kinase without any IDR.

Supplementary Table S4: Identification of 40 common Kinases. Top interactors from each of the 10 kinase groups were identified. (i) The first column is the list of uniprotID of kinases participating in the most number of intergroup interactions. Second column contains the list of uniprotID of kinases participating in the most number of intragroup interactions. Both inter and intra group interactions were combined for each group and a primary topological analysis was applied to identify topologically most relevant kinases for each of the 10 groups. UniprotIDs of these kinases are listed in the third column. Kinases common to all of the three lists in each group were identified (background green). If there were no common kinases, topologically most significant kinase was selected for that group. (ii) A list of 40 kinases common to all three lists were identified. Their names, uniprotIDs, their groups are listed.

Supplementary Table S5: Identification of the Global Interactome of the Kinome Modulators. The 5 kinome modulators were probed to identify known global interactions with kinases and other proteins. UniprotIDs of the kinases are listed. We identify a total of 1200 interactions of the 5 kinases with 963 unique proteins. Interestingly, we found that 25% (127 of 504) of the analyzed kinome interacts with these 5 kinome modulators.

Supplementary Table S6: Disease Enrichment of the primary interactome of the 5 kinome modulators reveals cancer as the most enriched disease. Disease enrichment profile of the 963 proteins interacting with the 5 KMs using IPA core analysis reveals significance of each disease (Category – column 1, p-value – column 2) and the molecules participating to enrich a particular disease or disorder (Molecules – column 3).

Supplementary Table S7: Molecular and Cellular Functions of the primary interactome of the 5 kinome modulators reveal Cell Death and Differentiation as the most significant function. Molecular and Cellular Functions of the 963 proteins interacting with the 5 KMs using IPA core analysis reveals significance of each function (Category – column 1, p-value – column 2) and the molecules participating to enrich a particular function (Molecules – column 3). This analysis identified that 5 KMs are able to influence cellular death and differentiation by interacting with 963 proteins.

Supplementary Table S8: 85% of the IDRs have 80% or more concordance with experimentally validated structures from PDB database. 100 IDRs were randomly picked from 43 kinases. These IDRs were compared with known structures of their protein using MobiDB, a database of protein disorder and mobility annotations. Our analysis revealed that 66% of IDRs had 100% concordance with the experimentally validated structures while 85% of IDRs had at least 80% concordance. Legend of the table is given at the bottom of the table.

Supplementary Table S9: Molecular Recognition Features (MoRFs) prediction reveals positive degree of correlation between number of MoRFs and percentage disorderliness of a protein. MoRFPred was used to predict MoRFs in 501 kinases. Column 1: UniProtID refers to the protein identifier. “Start” refers to starting AA residue number of the MoRF and “End” reers to ending AA residue number of the MoRF.