Bringing order to disorder: genomic analysis uncovers three distinct forms of protein disorder

Magali Michaut*[1], Jeremy Bellay*[2], Sangjo Han*[1], Michael Costanzo[1], Brenda J. Andrews[1], Charles Boone[1], Gary D. Bader[1], Chad L. Myers[2] and Philip M. Kim[1]

[1] Donnelly CCBR, University of Toronto, 160 College Street, M5S 3E1 Toronto, Ontario, Canada

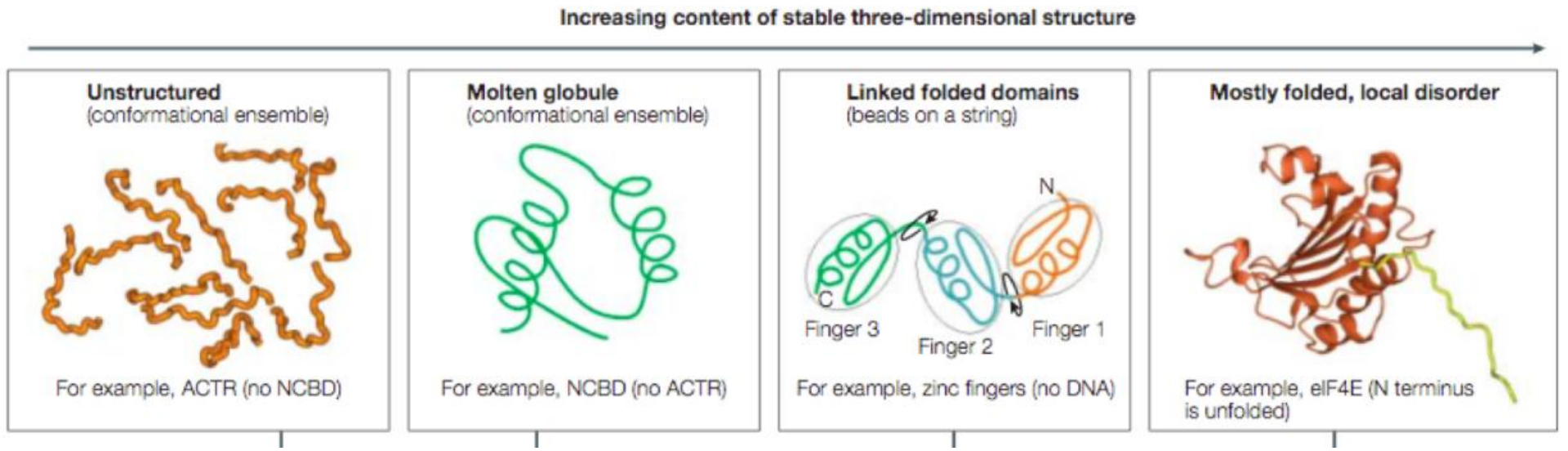
[2] Department of Computer Science and Engineering, University of Minnesota, MN 55455 Minneapolis, Minnesota, USA

** These authors equally contributed to this work*

There is a continuum of structure and disorder in the proteome

disorder

structure



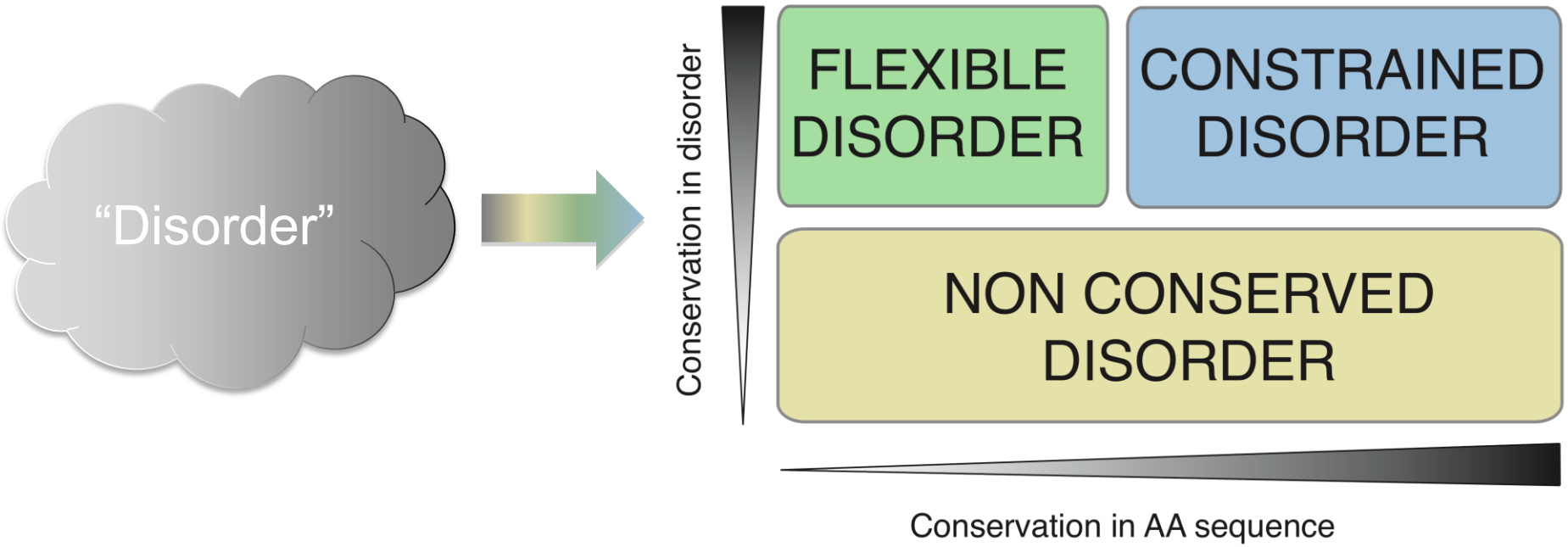
Dyson & Wright, 2005, Nature Reviews

Protein disorder is a diverse and complex phenomenon

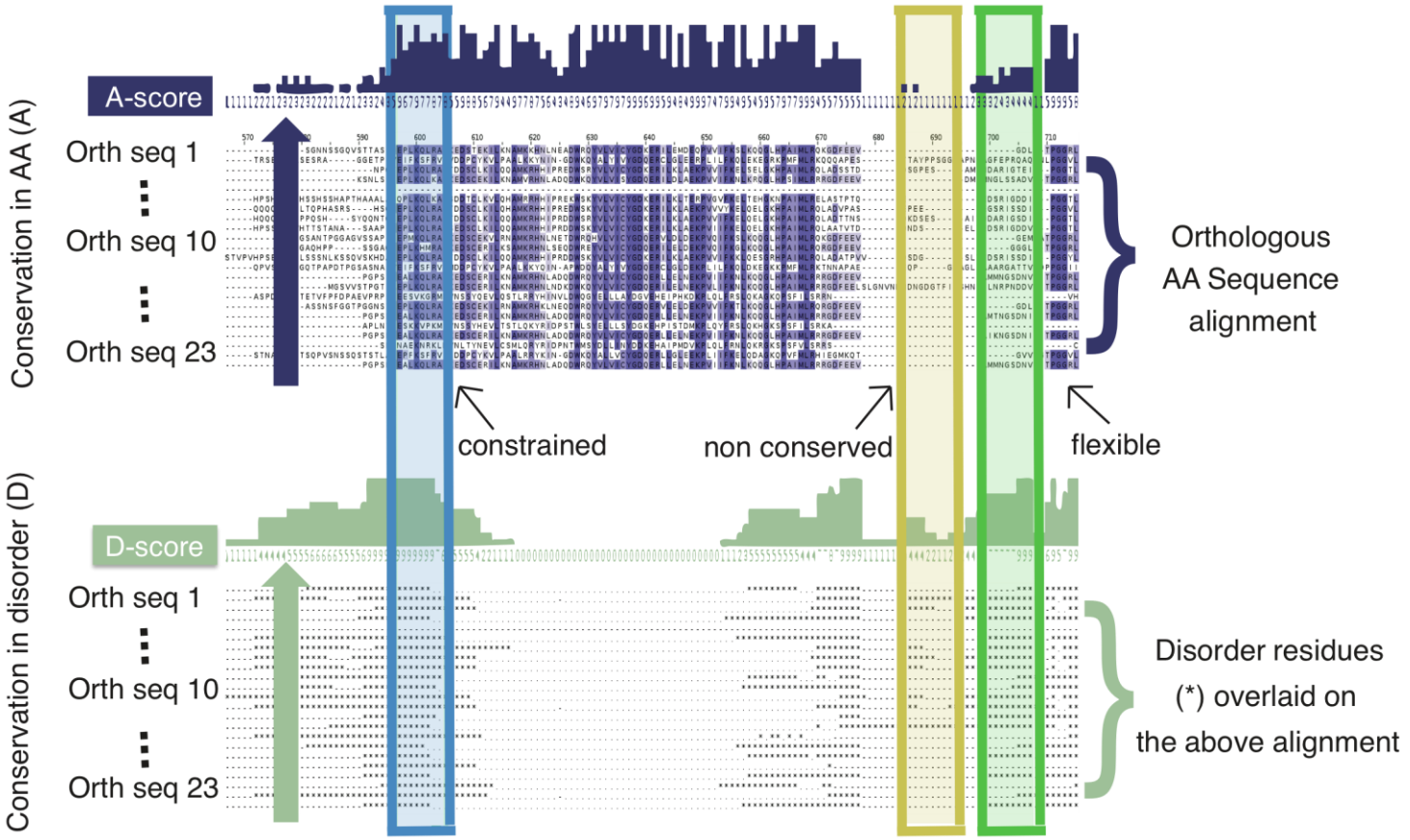
- More prevalent in complex organisms
  - 33% of the residues in Human, a few % in *E. coli*
- Various functional roles
  - Signaling, cellular regulation, nuclear localization, chaperone activity, RNA and DNA binding, antibody creation, multiple splicing
- Implicated in a variety of diseases
  - Cancer, neurodegenerative and cardiovascular diseases

We structure the different functions of disorder

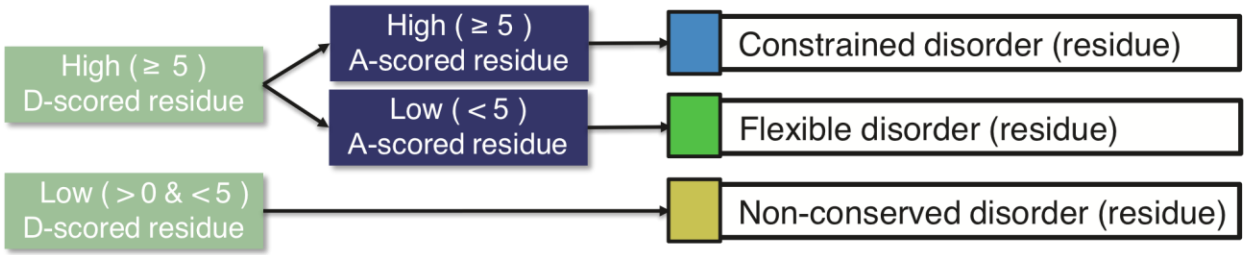
- Systematic analysis
- Use genetic interactions and comparative genomics
- Partition disorder into 3 classes



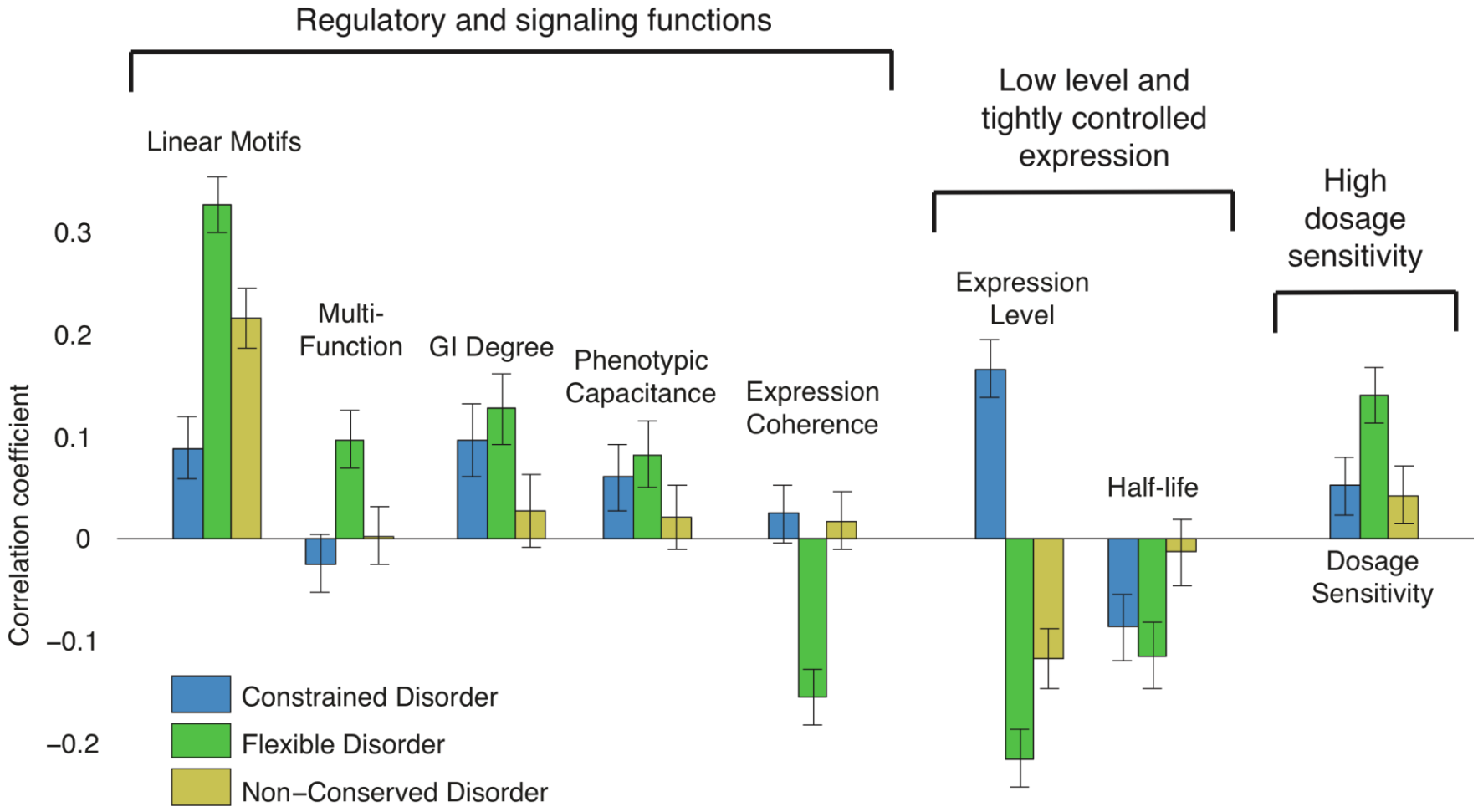
The 3 classes are defined based on sequence and disorder conservation scores among the yeast clade



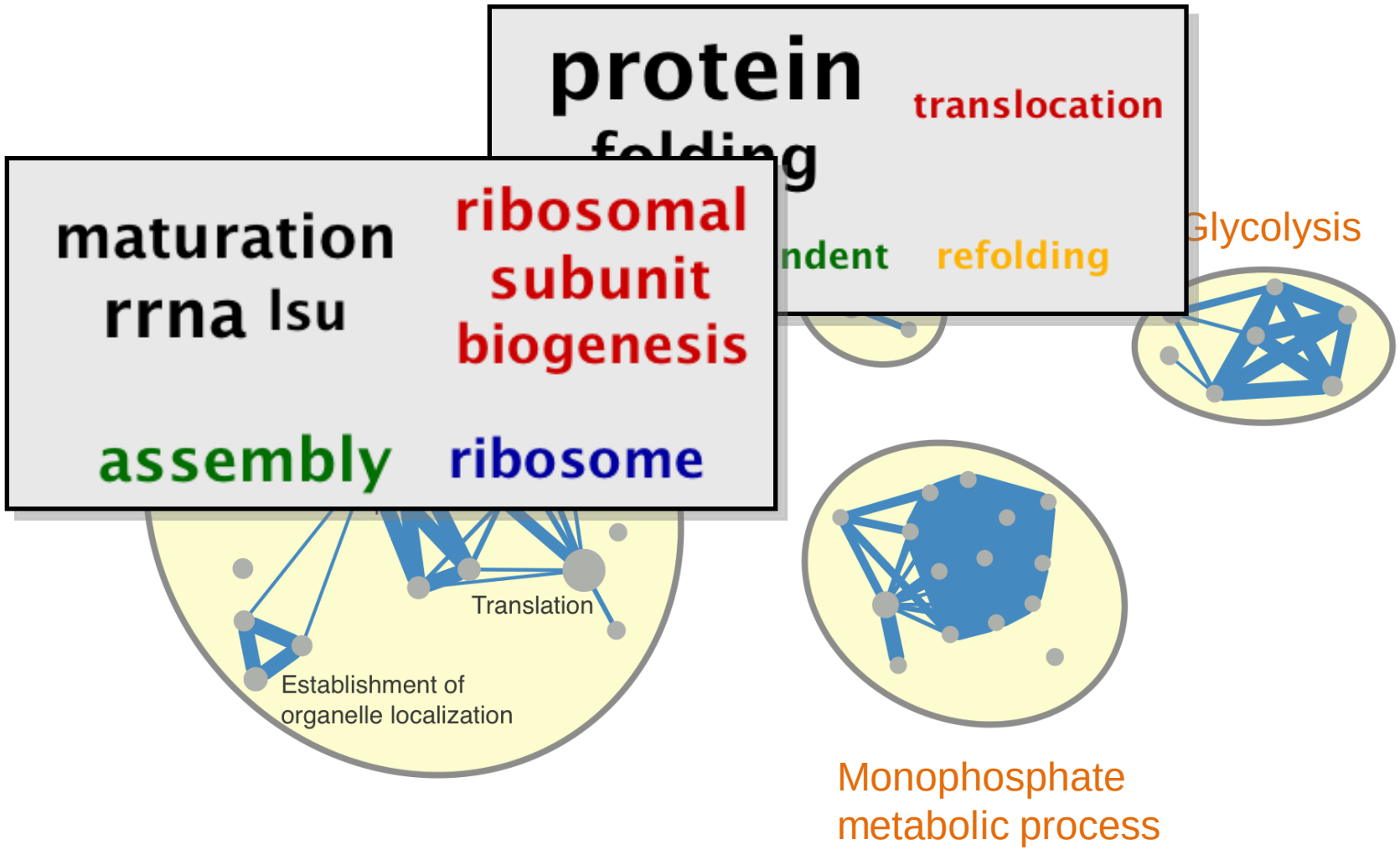
Define three distinct types of disorder residues across species



‘Classic’ disorder is more specifically related to flexible disorder



Proteins enriched in constrained disorder are involved in …



Bringing order to disorder: genomic analysis uncovers three distinct forms of protein disorder

Magali Michaut[1*], Jeremy Bellay[2*], Sangjo Han[1*], Michael Costanzo[1], Brenda J Andrews[1], Charles Boone[1], Gary D Bader[1], Chad L Myers[2] and Philip M Kim[1]
[1] Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada
[2] Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA
* These authors contributed equally to this work

Introduction

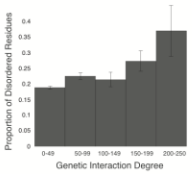
Many proteins include extended regions that do not fold into a native fixed conformation. These are referred to as being unstructured or disordered. Recent advances in computational prediction of disordered regions in amino acid sequences have greatly expanded our awareness of the widespread occurrence of disordered regions. Intrinsically disordered regions are widespread, especially in proteomes of higher eukaryotes, and have been associated with a plethora of different cellular functions. Here, we attempt to better understand the different roles of disorder using a novel analysis that leverages both comparative genomics and genetic interactions.

One sentence summary

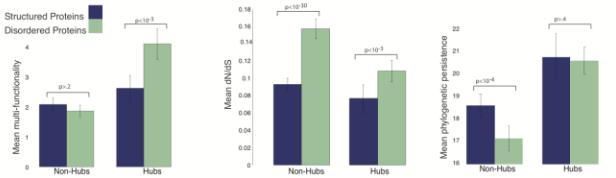
We use a novel comparative genomics analysis to uncover that protein disorder can be split into three biologically and biophysically distinct phenomena.

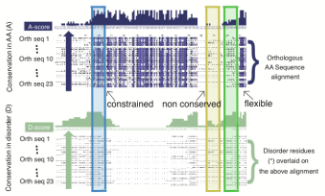
Genetic interactions distinguish different roles of disorder

The percentage of disorder is correlated with the genetic interaction degree. This suggests that disorder is related to the multifunctionality of proteins.



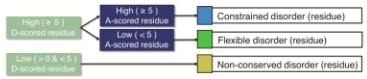
Disorder seems to play a functional role in hubs and not outside. Interesringly we ound that disordered proteins evolve fast but disordered hubs are conserved. Thus we decided to undertake a comparative analysis of disordered regions across the yeast clade.

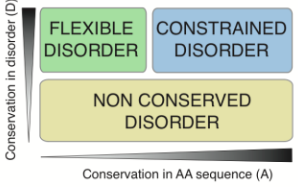




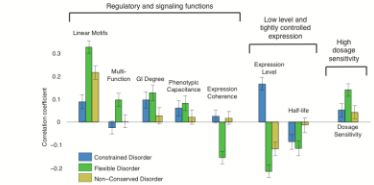
We define three distinct forms of disorder

We define a sequence and a disorder conservation score using comparative analysis across the yeast clade which we use to define three classes of disorder residues in the yeast clade.

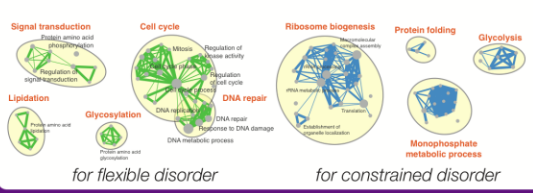




The 'classic' disorder is closest to flexible disorder



Proteins enriched in conserved disorder are involved in ...



Concluding remarks

In this work, we show that protein disorder can be partitioned into three biophysically and biologically distinct phenomena. Flexible disorder is closest to canonical protein disorder and is associated with signaling pathways and multi-functionality. Conversely, constrained disorder has markedly different functional attributes and is involved in RNA binding and protein chaperones. Finally, non-conserved disorder appears largely non-functional. These distinctions provide both an informative division of disorder and imply common underlying mechanisms that support these functions.

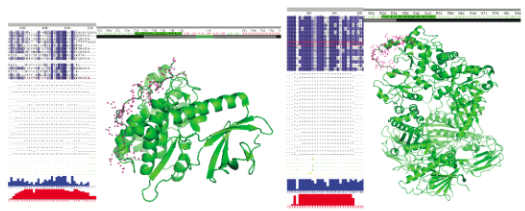
Acknowledgments
Thanks for the ISCB Travel Fellowship to attend ECCB.

Flexible disorder in Sky1

This C-ter disordered loop interacts with the activation loop of the kinase.

Constrained disorder in HSP90

This region is implicated in the chaperone activity of the protein.



Poster #3

Come talk to me!

Contacts
☺ at ECCB
@ magali.michaut@utoronto.ca

Acknowledgments

- Myers lab (Minneapolis)
  - Chad Myers
  - Jeremy Bellay
- Kim lab (Toronto)
  - Philip Kim
  - Sangjo Han
  - Simon Kim
- Bader lab (Toronto)
  - Gary Bader
- Boone lab (Toronto)
  - Charlie Boone
  - Michael Costanzo
- Andrews lab (Toronto)
  - Brenda Andrews

Donnelly Centre
for Cellular + Biomolecular Research