

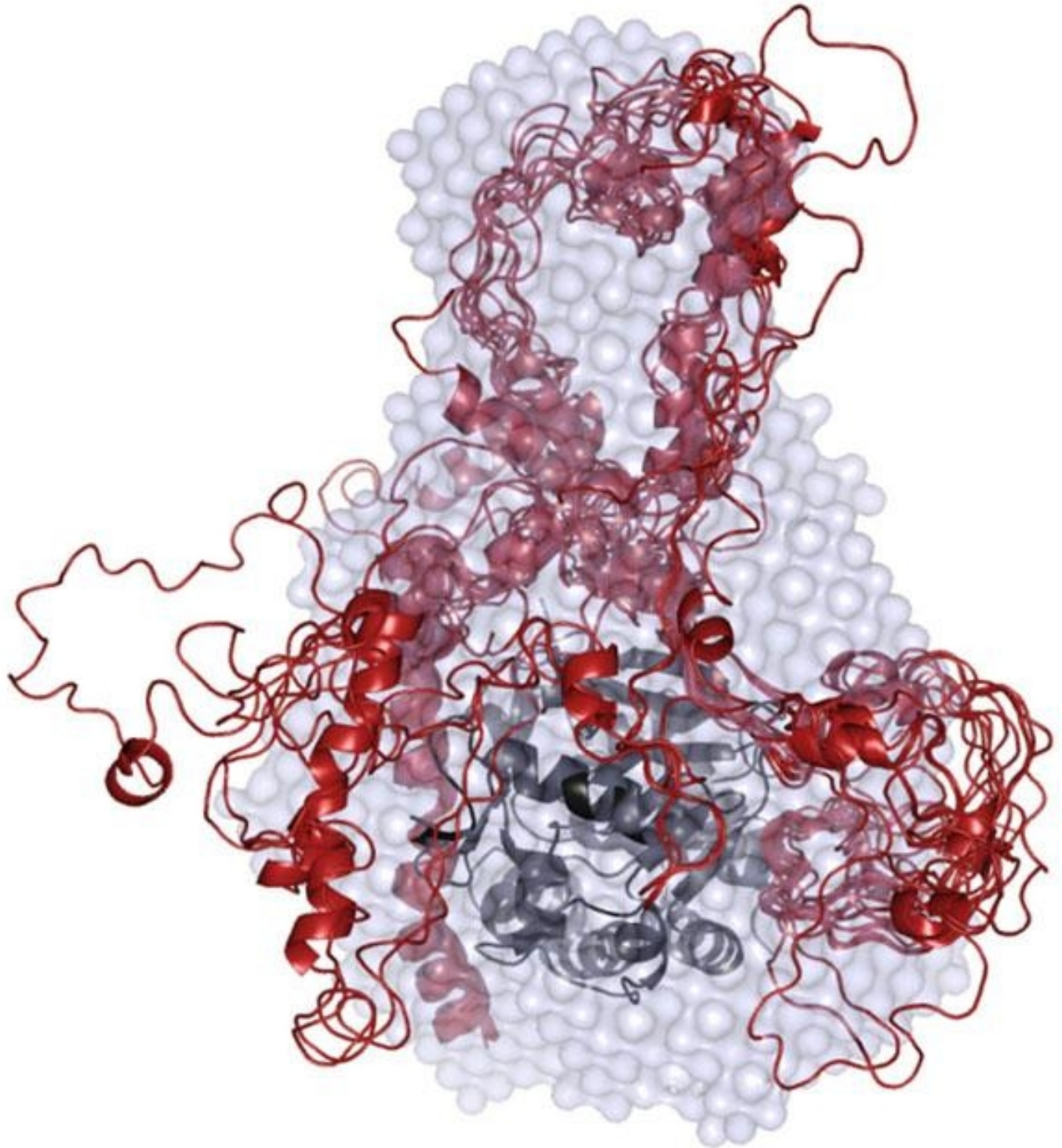
---

## **Intrinsically disordered proteins' predictors and databases: An overview**

**Author :** Tariq Abdullah

**Categories :** [Algorithms](#), [Softwares](#), [Tools](#)

**Date :** October 17, 2017



Intrinsically unstructured proteins (IUPs) are the natively unfolded proteins which must be unfolded or disordered in order to perform their functions. They are commonly referred as intrinsically disordered proteins (IDPs) and play significant roles in regulating and signaling biological networks [1]. IDPs are also involved in the assembly of signaling complexes and in the dynamic self-assembly of membrane-less nuclear and cytoplasmic organelles [1]. The disordered regions in a protein can be highly conserved among the species in respect of both the composition and the sequence [2].

IDPs have been found to be interacting frequently with the protein interaction networks [3,4]. The

computational and bioinformatics analysis helps to identify and characterize disordered protein regions. Around fifteen years ago, there was only one predictor available for identifying the disordered protein regions, known as PONDR [5]. Today, around 50 predictors can be used to predict the disordered regions in a protein [6] such as FoldIndex [7], GlobPlot [8], FoldUnfold [9-11], and so on. All these predictors are based on different algorithms. Structural disorders account to different states thereby rendering the prediction approach of single predictors ineffective. Therefore, some other combined algorithms were developed to predict IDPs more efficiently such as PONDR-FIT [12]. Some of the databases for IDPs also exist, for example, DisProt [13] and D2P2 [4].

In the last few years, the interactions between the IDPs and the other proteins have been an interesting research topic as their detailed analysis opens opportunities for therapeutic targeting. Some of the studies based on IDP interactions has led to successful pharmaceutical targeting [15]. DIBS (DIordered Binding Sites) is a recently developed database which stores interactions between the IDPs and the ordered proteins [16].

The annotations of order and disorder are grouped into three categories:

1. Proteins are marked as disordered from the direct proofs collected from the databases such as DisProt [13], and these proteins are referred as 'Confirmed'.
2. Besides the direct experimental validation, the proteins are also marked as disordered in DIBS if its close homolog found to be lacking the intrinsic structure.
3. the third group comprises the proteins regions in the disordered state that bind via a known, short functional motif (either from ELM [17], UniProt [18], Pfam [19] or the literature).

The predictors' accuracy for predicting the disordered regions in proteins is assessed as a part of the critical assessment of structure prediction (CASP) experiment, and the best accuracy among the predictors has been found to be 85% [21].

*This article is just a short introduction to IDPs predictors and the interaction databases. We will try to cover the algorithms of the predictors in detail in the upcoming articles. Meanwhile, keep telling us some topics related to bioinformatics which you are interested in, or tools/software to get their tutorials. You can write us at [info@bioinformaticsreview.com](mailto:info@bioinformaticsreview.com).*

## References

1. Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signaling and regulation. *Nature reviews. Molecular cell biology*, 16(1), 18.
2. Dyson, H. J., & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nature reviews. Molecular cell biology*, 6(3), 197.
3. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., & Uversky, V. N. (2005). Flexible nets. *The FEBS journal*, 272(20), 5129-5148.
4. Kim, P. M., Sboner, A., Xia, Y., & Gerstein, M. (2008). The role of disorder in interaction networks: a structural analysis. *Molecular systems biology*, 4(1), 179.
5. Garner, E., Cannon, P., Romero, P., Obradovic, Z., & Dunker, A. K. (1998). Predicting disordered

- regions from amino acid sequence. *Genome Informatics*, 9, 201-213.
6. He, B., Wang, K., Liu, Y., Xue, B., Uversky, V. N., & Dunker, A. K. (2009). Predicting intrinsic disorder in proteins: an overview. *Cell research*, 19(8), 929.
  7. Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., ... & Sussman, J. L. (2005). FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*, 21(16), 3435-3438.
  8. Linding, R., Russell, R. B., Neduva, V., & Gibson, T. J. (2003). GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic acids research*, 31(13), 3701-3708.
  9. Garbuzynskiy SO, Lobanov MY, Galzitskaya OV (2004). To be fold-ed or to be unfolded.(13), 2871-2877
  10. Galzitskaya, O. V., Garbuzynskiy, S. O., & Lobanov, M. Y. (2006). FoldUnfold: web server for the prediction of disordered regions in protein chain. *Bioinformatics*, 22(23), 2948-2949.
  11. Galzitskaya, O. V., Garbuzynskiy, S. O., & Lobanov, M. Y. (2007). Expected packing density allows prediction of both amyloidogenic and disordered regions in protein chains. *Journal of Physics: Condensed Matter*, 19(28), 285225.
  12. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K., & Uversky, V. N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1804(4), 996-1010.
  13. Sickmeier, M., Hamilton, J. A., LeGall, T., Vacic, V., Cortese, M. S., Tantos, A., ... & Obradovic, Z. (2006). DisProt: the database of disordered proteins. *Nucleic acids research*, 35(suppl\_1), D786-D793.
  14. Oates, M. E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M. J., Xue, B., ... & Dunker, A. K. (2012). D2P2: database of disordered protein predictions. *Nucleic acids research*, 41(D1), D508-D516.
  15. Corbi-Verge, C., & Kim, P. M. (2016). Motif mediated protein-protein interactions as drug targets. *Cell Communication and Signaling*, 14(1), 8.
  16. Schad, E., Fichó, E., Pancsa, R., Simon, I., Dosztányi, Z., & Mészáros, B. (2017). DIBS: a repository of disordered binding sites mediating interactions with ordered proteins. *Bioinformatics*, btx640.
  17. Dinkel, H., Van Roey, K., Michael, S., Kumar, M., Uyar, B., Altenberg, B., ... & Dahl, S. L. (2015). ELM 2016—data update and new functionality of the eukaryotic linear motif resource. *Nucleic acids research*, 44(D1), D294-D300.
  18. UniProt Consortium. (2014). UniProt: a hub for protein information. *Nucleic acids research*, gku989.
  19. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., ... & Studholme, D. J. (2004). The Pfam protein families database. *Nucleic acids research*, 32(suppl\_1), D138-D141.c
  20. Monastyrskyy, B., Fidelis, K., Moulton, J., Tramontano, A., & Kryshtafovych, A. (2011). Evaluation of disorder predictions in CASP9. *Proteins: Structure, Function, and Bioinformatics*, 79(S10), 107-118.

**Sharing is caring. Spread the love!**

- [Print](#)
- [Email](#)
- [LinkedIn](#)
- [Twitter](#)
- [Facebook](#)
- [Google](#)

•