Thirunavukkarasu Sivaraman /J. Pharm. Sci. & Res. Vol. 9(10), 2017, 1979-1981

ISSN:0975-1459
Journal of Pharmaceutical Sciences and Research
www.jpsr.pharmainfo.in

Computational Tools for Analyzing Hydrogen-Deuterium Exchange of Proteins

Thirunavukkarasu Sivaraman
Department of Biotechnology, Karpagam University, Coimbatore – 641021, Tamil Nadu, India.
E-mail: thirunavukarasusivaraman.kahedu.edu.in

Abstract

Understanding structural excursions of proteins under native conditions at residue level resolutions is crucial to map energy landscapes and folding mechanisms of proteins. Moreover, the structural stabilities of proteins provide insights about the forces governing conformations of the proteins. The hydrogen – deuterium (H/D) exchange methods are robust strategies for probing structural interactions, stability, folding and dynamics of proteins at residue level resolution. In this article, principles and applications of various experimental and computational tools that are being useful for analyzing H/D exchange of proteins have been brought into fore in a concise and forthright manner. Moreover, scopes of the methods in the future scenario of protein science have also been brought into fore.

Keywords: Cryptic intermediates, H/D exchange, Free energy of exchange and Folding.

1. INTRODUCTION

Hydrogen-deuterium (H/D) exchange is a process in which solvent deuterium exchange with labile protons of proteins dissolved in deuterium oxide (D_2O) in an irreversible manner. Mechanisms by which H/D exchange occurs in proteins have been proposed nearly 60 years ago and the theory could be well-validated using highly sophisticated recent biophysical techniques [1-5]. The exchange processes in proteins may happen either at pure EX1 (pseudo unimolecular reaction) or EX2 (bimolecular reaction) or mixed EX1/EX2 conditions and the H/D exchange methods are unique, exquisite and robust strategies for probing structural interactions, stability, dynamics and folding of proteins at residue level resolutions [6,7].

Understanding the relationships among structures, stabilities, dynamics and folding of proteins is crucial for designing *de novo* peptides of therapeutic and industrial purposes. In general, conformational stabilities of proteins can be monitored under denaturing conditions (through external chemical or physical forces) by using traditional biophysical techniques such as fluorescence/ circular dichroism / infrared spectroscopic techniques [8]. However, since these methods are insensitive to detect cryptic intermediates that are infinitesimally populating in the equilibrium unfolding pathways of proteins, the free energy of unfolding estimated by these methods are ambiguous in most proteins studied to date [9]. Fortunately, unfolding free energy exchange of proteins can be studied at residue level under native conditions (in absence of denaturants) by using NMR-assisted H/D exchange methods [10,11]. In ambient conditions, native states of proteins are always in equilibrium with the number of microstates defined by the Boltzmann relationship and the exchange reactions of labile protons of proteins are happening through the unfolded states and hence the reactions are independent of population of native states. Thus, unlike traditional methods, signal of the microstates are not swamped by the predominant native state in the H/D exchange methods and in turn, free energy exchange of residues representing global unfolding – folding reactions are considered as free energy of the proteins. Moreover, residue-specific free energy values estimated from the methods are useful to delineate relative rigidities and flexibilities of various structural segments of proteins.

2. NMR/MS - H/D EXCHANGE

H/D exchange methods in conjunction with Nuclear Magnetic Resonance (NMR) and as well with Mass spectrometric (MS) techniques are very useful to estimate residue-specific refolding rate constants of proteins in the time span of sub-seconds and also to characterize the folding pathways (sequential vs. parallel) of proteins [12-14]. Another potential advantage of the H/D exchange methods is detection of short-lived partially unfolded states known as cryptic intermediates (CIs) accumulating in the unfolding kinetics of proteins under native conditions. The methods are robust not only on identifying the CIs structurally but also on mapping out the energetic landscapes of proteins under their native conditions [15,16]. To date, H/D exchange methods are only available experimental tactics to study stability, folding and dynamics of proteins under conditions favoring native folded conformations. Streptomyces subtilisin inhibitor was the first protein to be studied using NMR-H/D exchange method under EX2 conditions [17]. Since then, thermodynamics and kinetics of 83 proteins have been studied using NMR-assisted H/D exchange methods to date (1985-2016). Quite a large amount of H/D exchange data on 83 proteins belonging to all types of classes brings an excellent platform where one can use the wealthy data to figure-out various types of structural and dynamic information that were mostly eluded in the macroscopic experiments [18].

Notwithstanding the advantages of the H/D NMR and H/D MS methods in protein chemistry, these methods are laborious, expensive, technically challenging, time consuming and also require sound experimental knowledge on the H/D exchange of proteins. Moreover, the methods may not be suitable to the proteins which cannot withstand their folded structures in the solution condition throughout the course of experiments and also to the proteins which are

highly prone to get into aggregation or degradation. In these contexts, computational tools will be excellent alternative to the H/D exchange methods provided the tools are robust and reliable on probing stability and folding pathways of proteins on the basis of their structural architectures [18-21]. There are several computational tools for predicting various parameters for the H/D exchange of proteins and also for analyzing the exchange data derived from NMR spectroscopy and MS spectrometry techniques.

3. COMPUTATIONAL TOOLS FOR ANALYZING H/D EXCHANGE DATA

The basic principle of the H/D MS methods is measuring overall degree of deuteration in proteins that are subjected to the H/D labeling experiments. Deuterium labelling, mass data acquisition and data processing are the three main steps involved in the H/D exchange of proteins by mass spectrometry experiments. Computational tools available in the literature for processing and analyzing H/D exchange data of proteins by mass spectrometry are as follows: AutoHD, DEX, EXMS, HDX Analyzer, HDXfinder, HDsite, HDX Workbench, HeXicon, HX-Express, Hydra, TOF2H and MS Tools [22-33]. Though prime functions of all these software packages are to determine deuterium distributions in the digested peptide fragments, the tools are differing from each other in the algorithm employed for analyzing the mass spectrometric data. Apart from estimating deuterium uptake level of polypeptide fragments, some of these tools also provide graphical outputs and statistical inferences on the exchange processes.

A few software tools and algorithms have been developed for analyzing the H/D exchange data of proteins obtained from NMR experiments. They are SPHERE, CIntX, H-Protection, CamP, COREX/BEST, META, OneG and OneG-Vali [34-41]. Of these tools, SPHERE and CIntX are very useful to calculate intrinsic exchange rates of labile protons in proteins. While the former tool requires many prerequisite inputs in addition to amino acid sequences of proteins, the latter tool predicts residue-specific exchange rates of backbone and as well side-chain labile protons of unfolded proteins by only using three-dimensional structures of the corresponding proteins as input data. The program is publicly available at http://sblab.sastra.edu/cintx.html Residue-specific free energies of proteins can be predicted by means of CamP and COREX/BEST computational tools. However, prediction accuracy of the methods was shown to be around 50% only. H-protection server aims at predicting only protection status of each residue in a protein from its primary sequence. The webservers META, OneG and OneG-Vali are very useful on analyzing unfolding kinetics of proteins under native conditions. It is worthy of mentioning that the OneG-Vali is a unique computational tool (only available program to date) of this kind for qualitatively and quantitatively predicting cryptic intermediates (CIs) that may presumably accumulate in the unfolding kinetics of proteins under native conditions. The tool requires 4 prerequisite parameters (atomic coordinates, ΔG_{HX}, ΔG_U, and C_m of proteins) and completes a successful run within a few minutes. Structural coordinates of CIs, population of each foldon and all calculated parameters are directly downloadable in appropriate formats from the webserver. The tool can also be used to validate CIs characterized by experimental methods. The OneG-Vali can be freely accessed and instantly used at http://sblab.sastra.edu/oneg-vali.html

In addition to the tools mentioned above, several computational strategies have also been reported in the literature to calculate overall rate of folding/unfolding of proteins [42-44]. In these contexts, developing a computational tool for predicting residue-specific free-energies and folding/unfolding rate constants at defined conditions on the basis of three-dimensional structures of proteins would also be a quite interesting and a highly challenging task in near future in the area of protein folding pathways by *in silco*. The success on the task, in turn, will lead to computationally explore the energetic levels of residues that are unfolding and refolding by various mechanisms (global, sub-global and local structural fluctuations) under native conditions of proteins. Thus, there are great scopes to develop unprecedented experimental strategies and as well computational tools for addressing various structural excursions of protein molecules under conditions favoring folded conformations.

REFERENCES

1. Hvidt A, Nielsen SO (1966) Adv Protein Chem. 21:287–366.
2. Bai Y, Milne JS, Mayne L, Englander SW (1993) Proteins. 17:75–86.
3. Huyghues-Despointes BMP, Scholtz JM, Pace CN (1999) Nat Struct Biol. 6:910–912.
4. Sivaraman T, Arrington CB, Robertson AD (2001) Nat Struct Biol. 8:331–333.
5. Woodward CK (1994) Curr Opin Struct Biol. 4:112-116.
6. Englander SW, Mayne L, Krishna MM (2007) Q Rev Biophys. 40:287-326.
7. Baldwin RL (2011) Proteins. 79:2021-2026.
8. Pace CN, Shirley BA, Mcnutt M, Gajiwala K. (1996) FASEB Journal. 10:75-83.
9. Richa T, Sivaraman T (2012) PLoS ONE. 7(3): e32465.
10. Bai Y (2006) Chem Rev. 106:1757-1768.
11. Neira JL, Itzhaki LS, Otzen DE, Davis B, Fersht AR (1997) J Mol Biol. 270:99-110.
12. Roder H, Elove GA, Englander SW (1988) Nature. 335:700-704.
13. Udgaonkar JB, Baldwin RL (1988) Nature. 335:694-699.
14. Sivaraman T, Kumar TKS, Chang DK, Lin WY, Yu C (1998) J Biol Chem. 273:10181-10189.
15. Bai Y, Sasnick TR, Mayne L, Englander SW (1995) Science. 269:192–197.
16. Chamberlain AK, Handel TM, Marqusee S (1996) Nat Struct Biol. 3:782–787.
17. Akasaka K, Inoue T, Hatano H, Woodward CK (1985) Biochemistry. 24:2973-2979.
18. Richa T, Sivaraman T (2013) International Journal of Research in Pharmaceutical Sciences. 4(4):550-562.
19. Englander SW, Mayne L, Krishna MMG (2007) Q Rev Biophys. 40: 287-326.
20. Compiani M, Capriotti E (2013) Biochemistry. 52: 8601-8624.
21. Hilser VJ, E Freire (1996) J Mol Biol 262: 756-772.
22. Guttman M, Weis DD, Engen JR, Lee KK (2013) J Am Soc Mass Spectrom 24: 1906-1912.
23. Hotchko M, Anand GS, Komives EA, Ten Eyck LF (2006) Protein Sci 15: 583-601.
24. Kan ZY, Mayne L, Sevugan Chetty P, Englander SW (2011) J Am Soc Mass Spectrom 22: 1906-1915.
25. Kan ZY, Walters BT, Mayne L, Englander SW (2013) Proc Natl Acad Sci USA 10: 16438-16443.
26. Kavan D, Man P (2011) Int J Mass Spectrum 302: 53-58.

27. Liu S, Liu L, Uzuner U, Zhou X, Gu M, Shi W, Zhang Y, Dai SY, Yuan JS (2011) BMC Bioinformatics 12: S43.
28. Lou X, Kirchner M, Renard BY, Köthe U, Boppel S, Graf C, Lee C-, Steen JAJ, Steen H, Mayer MP, Hamprecht FA (2010) Bioinformatics 26: 1535-1541.
29. Miller, D.E., Prasannan, C.B., Villar, M.T., Fenton, A.W., Artigues, A (2012) J Am Soc Mass Spectrum 23: 425-429.
30. Nikamanon P, Pun E, Chou W, Koter MD, Gershon PD (2008) BMC Bioinformatics 9: 387.
31. Palmblad M, Buijs J, Hakansson P (2001) J Am Soc Mass Spectrum 12: 1153-1162.
32. Pascal BD, Willis S, Lauer JL, Landgraf RR, West GM, Marciano D, Novick S, Goswami D, Chalmers MJ, Griffin PR. HDX Workbench (2012) J Am Soc Mass Spectrum 23: 1512-1521.
33. Slysz GW, Baker CAH, Bozsa BM, Dang A, Percy AJ, Bennett M, Schriemer DC (2009) BMC Bioinformatics 10: 162.
34. Bai Y, Milne JS, Mayne L, Englander SW (1993) Proteins 1993; 17: 75-86.
35. Richa T, Sivaraman T (2011) J Pharm Sci and Res 3: 1486-1490.
36. Richa T, Sivaraman T (2012) J Pharm Sci and Res 4: 1852-1858.
37. Richa T, Sivaraman T (2012) PLoS ONE 7: e32465.
38. Richa T, Sivaraman T (2014) RSC Adv 4: 36325-36335.
39. Lobanov MY, Suvorina MY, Dovidchenko NV, Sokolovskiy IV, Surin AK, Galzitskaya OV (2013) Bioinformatics 29: 1375-1381.
40. Hilser VJ, E Freire (1996) J Mol Biol 262: 756-772.
41. Tartaglia GG, Cavalli A, Vendruscolo M (2007) Structure 15: 139-143.
42. Plaxco KW , Simons KT, Baker D. (1998) J Mol Biol. 277: 985-994.
43. Gromiha MM, Selvaraj S. (2001) J Mol Biol. 301: 27–32.
44. Gromiha MM. (2010) Protein Bioinformatics: From sequence to function. Academic Press: Elsevier India Pvt Ltd. New Delhi.