were fitted to a simple linear equation and the fitted parameters such as slope and positive correlation coefficient were found to be 0.996 and 0.999, respectively, for both of the proteins. These observations unambiguously demonstrate that the reliability of the CIntX on calculating the k_{rc} of NHs in proteins. The program, then, calculates residue-specific free energies and protection factors of proteins with the only requirement of having NMR-derived k_{ex} of NHs in the proteins. In the third stage, the program recalculates the k_{rc} of NHs by accounting the effect of ionisable groups in proteins and calculates k_{rc} of labile protons that are not NHs in the fourth stage of the program (discussed in the next headings).

3.4 Revaluation of k_{rc} values of backbone labile protons in proteins after accounting the effect of ionisable side chains

In general, polypeptide chains consist of about 25% amino acids having ionisable groups (COOH/NH_3/SH/OH) in their side chains [16, 17]. These ionisable groups will be fully or partially protonated depending on their pK_R under defined experimental conditions. Bai *et al.* have determined the effect of protonated and deprotonated forms of a few of

Table 3 :The percentage of protonated and deprotonated forms of ionisable groups in proteins at pD 3.6 and their dissociation constants (pK_R).

Residues	pK_R	% Protonated forms	% Deprotonated forms
Asparatate	3.67	54	46
Glutamate	4.25	18	82
Histidine	6.54	100	0
Cysteine	8.55	100	0
Tyrosine	9.6	100	0
Lysine	10.4	100	0
N-terminal	8.00*	100	0
C-terminal	3.67**	54	46

* pK_R of N-terminal amino group.
** pK_R of C-terminal carboxylic acid group.

Table 4: Values of k_{rc} and ΔG_{HX} of NHs of a few amino acids in cobrotoxin at pD 3.6, 298 K have been tabulated, herein.

Sl. No.	Residue#	Uncorrected*		Corrected**		$\Delta\Delta G_{HX}$
		k_{rc}	ΔG_{HX}	k_{rc} (min^{-1})	ΔG_{HX}	
1	E2	0.13	2.26	0.10	2.08	0.18
2	C3	1.04	3.62	0.84	3.48	0.14
3	H4	2.18	NA@	2.18	NA	NA
4	N5	2.52	NA	2.52	NA	NA
5	E21	0.31	1.4	0.23	1.22	0.18
6	T22	0.25	NA	0.20	NA	NA
7	D31	0.97	NA	0.35	NA	NA
8	H32	3.01	NA	1.32	NA	NA
9	R33	0.97	NA	0.97	NA	NA
10	E38	0.33	3.57	0.24	3.38	0.19
11	R39	0.36	2.93	0.29	2.8	0.13
12	E51	0.12	2.19	0.09	2.01	0.18
13	I52	0.06	2.23	0.05	2.11	0.12
14	D58	0.93	NA	0.33	NA	NA
15	R59	0.58	3.46	0.26	2.98	0.48
16	C-terminal$	0.26	NA	0.07	NA	NA

#Amino acids are denoted by single letter codes.
*The k_{rc} values of NHs are calculated without accounting the effect of ionisable groups in the protein.
**The k_{rc} values of NHs are calculated by accounting the effect of ionisable groups in the protein.
@NA represents the residues for which the values of k_{ex} are not available.
$C-terminal carboxylic acid group.

Table 5: Estimation of the k_{rc} values of labile protons in the side-chains of cobrotoxin at pD 3.6, 298K.

Sl. No.	Residue#	Side Chain Proton	Rate Constant k_{rc} (sec^{-1})	Time Constant τ (sec)
1	N5	H_E	1.47	0.68
2	Q6	H_E	0.98	1.02
3	Q6	H_Z	0.41	2.42
4	Q7	H_E	0.71	1.40
5	Q7	H_Z	0.30	3.32
6	Q10	H_E	0.57	1.76
7	Q10	H_Z	0.25	4.02
8	N23	H_E	1.07	0.93
9	R28	NδH	0.17	5.97
10	W29	N_1H	0.56	1.79
11	R30	NδH	0.10	10.13
12	R33	NδH	0.86	1.17
13	R36	NδH	0.14	7.01
14	R39	NδH	0.25	4.01
15	N48	H_E	1.53	0.66
16	N53	H_E	0.76	1.32
17	R59	NδH	0.22	4.50
18	N61	H_E	0.99	1.01
19	N62	H_E	2.22	0.45

Amino acids are denoted by single letter codes.

these amino acids (Asp, Glu, and His) on the intrinsic exchange rates of labile protons [7]. It is usually assumed (in the calculations of exchange rates of NHs in proteins) that an ionisable groups will be fully protonated when its pK_R is greater than the solution pD and fully deprotonated when its pK_R is lower than the pD of the solution. However, when there is not much difference (< 1.0) between the pK_R of an ionisable group and the solution pD, the ionisable group in a polypeptide chain will have significant amount of its protonated and deprotonated forms under the defined experimental conditions. Table 3 shows the pK_R of ionisable groups of amino acids (Asp, Glu, His, Cys, Tyr, and Lys) and their percentage of protonated/deprotonated forms at pD 3.6, 298 K. The CIntX determines the fraction of protonated (F_{HA}) and deprotonated (F_A) forms of those ionisable groups using equation (3) and equation (4) (refer method section) and then recalculates the k_{rc} of NHs that are influenced by the ionisable groups using the equation (15) described below herein in the third stage of the program.

$$k_{rc} = 10^{(X - pD)} + 10^{(Y + pD - pKD)} + 10^{(Z)}. \qquad (15)$$

where,

$$X = \log k_a + \log (F_{HA}{*}A_{L+} + F_A{*}A_{L-}) + \log (F_{HA}{*}A_{R+} + F_A{*}A_{R-}) + (E_a/4.57){*}(1/293 - 1/T). \qquad (16)$$

$$Y = \log k_b + \log (F_{HA}{*}B_{L+} + F_A{*}B_{L-}) + \log (F_{HA}{*}B_{R+} + F_A{*}B_{R-}) + (E_b/4.57){*}(1/293 - 1/T). \qquad (17)$$

$$Z = \log k_w + \log (F_{HA}{*}B_{L+} + F_A{*}B_{L-}) + \log (F_{HA}{*}B_{R+} + F_A{*}B_{R-}) + (E_w/4.57){*}(1/293 - 1/T). \qquad (18)$$

wherein '+' and '-' signs in the subscripts of the standard reference parameters of a residue indicate the protonated and deprotonated forms of the corresponding amino acid in the polypeptide chain, respectively [7,15]. Table 4 shows the recalculated k_{rc} values for the NHs of cobrotoxin, which are influenced by ionisable groups present in the protein. Residue-specific free energy (ΔG_{HX}) values of the NHs have also been calculated using NMR-derived k_{ex} values of the NHs at identical exchange conditions (pD 3.6, temperature 298K) of the protein [18]. The highest $\Delta\Delta G_{HX}$ was found to be 0.48 kcal/mol for R59 of the protein at pD 3.6, 298 K, as the residue is preceded by D58, an amino acid having ionisable group. This observation reveals that discrepancy between the k_{rc}-corrected and the k_{rc}-uncorrected will be remarkable for NHs depending on their positions in the polypeptide chain and the exchange-conditions such as pH and temperature. In the fourth stage, the CIntX calculates k_{rc} values of labile protons present in the side chains of arginine, asparagine, glutamine and tryptophan using equation (11) in a single step. Table 5 shows k_{rc} values of labile protons from a few side chains of cobrotoxin at pD 3.6 and 298 K. The labile protons of these side-chains in the protein are exchanging in the time span of 0.5 – 10.1 seconds (Table 5). The k_{rc} values of side-chain labile protons will be powerful probes to determine the folding rate(s) of tertiary structural formations of proteins in which the labile protons are protected from the solvent exchange under native conditions.

Error-free estimation of k_{rc} values for labile protons of proteins is indispensable to estimate accurate ΔG_{HX} ($\Delta G_{HX} = -RT \ln (k_{ex}/k_{rc})$) of proteins. The exchange of free energy is important to understand the relationship between three-dimensional (3D) conformations and stabilities of proteins as each protein adopts a specific, well-defined 3D structure, which is important for its biological activities [19]. Moreover, ΔG_{HX} analysis provides clues on understanding the mechanism of unfolding of proteins (two-state/multi-state processes) and on analysing the 3D structural architectures (domains organization) of proteins [20, 21]. Thus, it is obvious that estimation of an accurate ΔG (free energy change) for proteins at ambient conditions is indispensable to unambiguously address the thermodynamic and kinetic events of proteins. In these backgrounds, we strongly believe that CIntX, fully automated and user-friendly computational tool facilitated with web server (*http://sblab.sastra.edu/cintx.html*), is very useful for estimating k_{rc} values of all types of labile protons (backbone, side chains, N-terminal and C-terminal positions) of proteins.

4. Conclusion

We have herein developed a software tool (CIntX) to calculate the intrinsic exchange rate constants (k_{rc}) of labile protons from backbone, side-chains and termini positions of proteins at defined experimental conditions. CIntX is a fully automated and user-friendly tool in its functions and requires four inputs to achieve the task: PDB file of protein, temperature in Kelvin, pH, and ionic strength. The program accounts the effect of Xaa-Pro peptide bonds, disulfide bonds and ionisable side chains on the k_{rc} estimations using the 3D structural co-ordinates of the given protein. The program also calculates residue-specific free energies and protection factors of proteins with the only requirement of having NMR-derived k_{ex} of NHs in the proteins. Estimation of precise ΔG_{HX} (free energy of exchange) for proteins at ambient conditions is indispensable to unambiguously address the thermodynamic and kinetic events of the proteins. In these backgrounds, CIntX is a unique tool of this kind for calculating k_{rc} of labile protons in proteins. The web server of the program is publicly available at *http://sblab.sastra.edu/cintx.html*

References

[1] Dobson, C. M., Karplus, M., *Curr. Opin. Struct. Biol.* 1999, *9*, 99 – 101.
[2] Pace, C. N., Scholtz, C. M., in: Creigton, T. E. (Ed), *Protein Structure - A Practical Approach,* Oxford University Press Inc., New York 1997, pp. 535 – 566.
[3] Brockwell, D. J., Smith, D. A., Radford, S. E., *Curr. Opin. Struct. Biol.* 2000, *10*, 16 – 25.
[4] Englander, S. W., Mayne, L., Bai, Y., Sosnick, T. R. *Protein Sci.* 1997, 6, 1101 – 1109.
[5] Hvidt, A., Nielsen, S. O., *Adv. Protein Chem.* 1966, *21*, 287 – 366.
[6] Englander, S. W., Sosnick, T. R., Englander, J. J., Mayne, L., Bai, Y., *Curr. Opin. Struct. Biol.* 1996, *6*, 18 – 23.
[7] Bai, Y., Milne, J. S., Mayne, L., Englander, S. W., *Proteins.* 1993, *17*, 75 – 86.
[8] Zhang, Y., *Protein and peptide structure and interactions studied by hydrogen exchange and NM,* Ph.D. Thesis, University of Pennsylvania 1995.

[9] Wall, L., Christiansen, T., Orwant, J., *Programming Perl*, O'Reilly Media Inc., Sebastopol 2000.
[10] Connelly, G. P., Bai, Y., Jeng, M., Englander, S. W., *Proteins*. 1993, *17*, 87 – 92.
[11] Wuthrich, K., *NMR of proteins and nucleic acids,* John Wiley & Sons., New York 1986.
[12] Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., Fischer, G., *J. Mol. Biol.* 1998, *279*, 449 – 460.
[13] Huyghues –Despointes, B. M. P., Scholtz, J. M., Pace, C. N., *Nat. Struct. Biol.* 1999, *6*, 910 – 912.
[14] Bhattacharyya, R., Pal, D., Chakrabarti, P., *Protein. Eng. Des. Sel.* 2004, *17*, 795 – 808.
[15] Molday, R. S., Englander, S. W., Kallen, R. G., *Biochemistry* 1972, 11, 150 – 158.
[16] Grimsley, G. R., Scholtz, J. M., Pace, C. N., *Protein Sci.* 2009, *18*, 247 – 251.
[17] Thurlkill, R. L., Grimsley, G. R., Scholtz, J. M., Pace, C. N., *Protein Sci*. 2006, *15*, 1214 – 1218.
[18] Sivaraman, T., *Understanding the folding and the unfolding of three – finger toxins from Taiwan cobra (Naja naja atra).* Ph.D. Thesis, National Tsing Hua University, Taiwan, 1999.
[19] Huyghues-Despointes, B. M. P., Pace, C. N., Englander, S. W., Scholtz, J. M., in: Murphy, K. M. (Ed.), *Protein Structure, stability and folding,* Humana Press Inc., New Jersey 2001, pp. 69 – 92.
[20] Arrington, C. B., Teesch, L. M., Robertson, A. D., *J. Mol. Biol.* 1999, *285*, 1265 – 1275.
[21] Krishna, M. M. G., Hoang, L., Lin, Y., Englander, S. W., *Methods.* 2004, *34*, 51– 64.
[22] DeLano, W. L., *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA 2002.