T. Richa et al /J. Pharm. Sci. & Res. Vol.4(6), 2012, 1852 - 1858

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

CIntX: A Software Tool for Calculating the Intrinsic Exchange Rates of Labile Protons in Proteins

T. Richa and T. Sivaraman*
Structural Biology Laboratory,
Department of Bioinformatics, School of Chemical and Biotechnology,
SASTRA University,
Thanjavur – 613401, Tamil Nadu, India.

Abstract
Estimation of extrinsic (k_{ex}) and intrinsic (k_{rc}) exchange rate constants for labile protons of proteins is indispensable to determine the residue-specific free energies. While k_{ex} of labile protons in proteins are determined by experimental methods, the k_{rc} of those protons are calculated based on the parameters derived from the model compound studies, in general. We, herein, describe a computational tool, CIntX, which calculates the k_{rc} of labile protons in proteins at defined experimental conditions such as pH, temperature, solvents and ionic strength using three-dimensional structures of the proteins. The program also accounts structural parameters (amino acid sequence, *cis-trans* proline conformations, disulfide bonds and ionisable groups in the proteins) influencing the k_{rc} of labile protons in a fully automated manner. The robustness of the tool on predicting the k_{rc} of labile protons in proteins has been validated using data reported in the literature. The program is publicly available at http://sblab.sastra.edu/cintx.html

Keywords: Algorithm, Computational tool, Disulfide bonds, Free energy, H/D exchange, Intrinsic Exchange rates, Ionic strength, Proline conformation and Protection factor.

1. INTRODUCTION

The structure-function relationships of proteins have intrigued the researchers for many decades [1]. Each protein adopts a unique three-dimensional structure with defined stability, which is important for its biological activities. The stability of proteins is calculated by measuring the free energy differences between the folded (N) and the unfolded conformations (U) of the proteins at ambient conditions [2,3]. Hydrogen-deuterium (H/D) exchange put forward by Linderstram-Lang is a powerful method to estimate free energy of exchange (ΔG_{HX}) of proteins at residue level [4]. In a typical H/D exchange experiment, when a protein is dissolved in deuterium oxide (D_2O), labile amide protons (NHs) of the protein begin to exchange with solvent deuterium and the exchange phenomena is represented by the following two-state model equation proposed by Hvdit [5,6].

$$\text{Closed (NH)} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{Open (NH)} \xrightarrow{k_{rc}} \text{Exchanged.} \qquad (1)$$

In this model, the k_{op} and k_{cl} are rate constants for the unfolding and the folding reactions of proteins under native conditions, respectively. Exchange takes place only from the unfolded state of proteins with the rate constant of k_{rc}, which can be predicted based upon the parameters derived from studies on the model compounds [7]. In this context, it should be mentioned that SPHERE, a program developed by Roder, calculates k_{rc} values of backbone amide protons (NHs) of proteins from their primary structures [8]. In this paper, we are herein describing a novel computational tool, CIntX, which calculates k_{rc} values of labile protons located in the backbone, side-chains and termini of proteins at defined experimental conditions from their tertiary structures. The program requires only four inputs: PDB file of the protein, Temperature in Kelvin, pH and Ionic strength. The CIntX, in a fully automated fashion, accounts for the effect of structural parameters such as amino acid sequence, *cis-trans* proline conformations, disulfide bonds and ionisable groups in proteins on the estimation of k_{rc} values for labile protons in the proteins. The uniqueness of the tool on predicting the k_{rc} values for the labile protons of proteins and its application on determining ΔG_{HX} of the proteins have also been discussed in detail.

2. METHODS

The CIntX algorithm has been implemented using PERL scripting language [9]. The program accepts both amino acid sequences (represented by single letter codes) and PDB (Protein Data Bank) co-ordinates of proteins as query file. In order to calculate the k_{rc} values, the program takes temperature in Kelvin, pH in pD (pD = pH+0.4), ionic strength in molarities and activation energies in cal/mol [7,10]. The values of k_{rc} and ΔG_{HX} of labile protons are expressed in per minute and kcal/mol, respectively. While a PDB file is an input, the program automatically defines the *cis/trans* conformations of Xaa-Pro (Xaa can be any one of the twenty standard alpha amino acids and Pro denotes proline) peptide bonds and patterns of disulfide bonds in the protein using the following relationship:

$$D = \{(U_X\text{-}P_X)^2 + (U_Y\text{-}P_Y)^2 + (U_Z\text{-}P_Z)^2\}^{1\backslash 2} \qquad (2)$$

where D is the distance between the C_α of Xaa and C_α or C_δ of proline residue in angstrom; U and P stand for Xaa and Pro residues of Xaa-Pro peptide bond, respectively; X, Y and Z are the atomic co-ordinates of an atom considered. The Xaa-Pro peptide bond is considered to have *trans*-conformation, when the distance between C_α of Xaa and C_α of Pro (C_α—C_α) is greater than the distance between C_α of Xaa and C_δ of Pro (C_α—C_δ). Similarly, when C_α—C_δ > C_α—C_α, the Xaa-Pro peptide bond is considered to adopt *cis*-conformation [11]. In

order to determine the disulfide bond patterns in proteins, the distance between the sulphur atoms of thiol groups from two cysteins is calculated using the equation (2) and the two cysteines are considered to be disulfide bonded, if the calculated distance is within 2.3 Å.

For each successful run, the program provides two sets of the k_{rc} values for labile protons in proteins (one for accounting the effect of ionisable side-chains and another one for not accounting the effect). The CIntX determines the fraction of protonated (F_{HA}) and deprotonated (F_A) forms of ionisable groups in proteins using equations (3) & (4), respectively and the equations were derived from Henderson–Hasselbalch equation (mathematical derivations are not shown).

$$F_{HA} = 1/(1+10^{x-y}). \quad (3)$$

$$F_A = 1 - F_{HA}. \quad (4)$$

where x and y denote the pD of protein solution and the pK_R of ionisable groups, respectively. The program then recalculates the k_{rc} of NHs that are influenced by the ionisable groups (detailed description is given under the results and discussion heading). In addition, the program displays structural parameters such as Xaa-Pro conformations and disulfide bond patterns and derived parameters such as protection factors and residue-specific free energy values in kcal/mol provided NMR-derived k_{ex} values of labile protons are given as input parameters. A detailed list of instructions to use the tool is made available at *http://sblab.sastra.edu/cintx.html*

3. Results and Discussion

3.1 Mathematical derivations of equations for calculating k_{rc} of labile protons in proteins

The values of k_{rc} for labile protons of proteins can be estimated using the equation (5) as described by Bai et al. [7,10].

$$k_{rc} = k_a * R_a * 10^{(-pD)} + k_b * R_b * 10^{(pD-pKD)} + k_w * R_w. \quad (5)$$

where k_a, k_b and k_w are rate constants of acid, base and water catalyzed exchange reactions, respectively; R_a, R_b and R_w are the effect of residues that are on the left and right sides of amide protons at acidic, basic and neutral conditions, respectively; pK_D is the molar ionization constant of D_2O; pD is the pH-meter reading corrected to deuterium effect. Thus, the equation (5) accounts the effect of primary sequence and solution pH on the k_{rc} of labile protons under consideration. The effect of temperature on the intrinsic exchange rates of the labile protons is calculated using the following equation.

$$k(T) = k(293)\exp(-E[1/T - 1/293]/R). \quad (6)$$

where $k(T)$ and $k(293)$ are rate constants at desired temperature and 293 K, respectively; E is the activation energy and its value for acidic, basic and neutral exchange reactions are 14 kcal/mol/K, 17 kcal/mol/K and 19 kcal/mol/K, respectively; T is the absolute temperature in Kelvin and R is the gas constant. The effect of temperature on the acid catalyzed exchange rate constant is as shown in the equation (7).

$$\log k(acid)_T = \log k(acid)_{293K} + (E_a/4.57) * (1/293 - 1/T) \quad (7)$$

By combining the equation (7) and the component representing acid-catalyzed reaction in equation (5), the following equation (8) could be derived.

$$\log k(acid)_T = \log k_a + \log R_a - pD + (E_a/4.57) * (1/293 - 1/T). \quad (8)$$

Considering, $X = \log k_a + \log R_a + (E_a/4.57) * (1/293 - 1/T)$, the equation (8) can be rewritten as shown either in the equation (9) or in the equation (10).

$$\log k(acid)_T = X - pD. \quad (9)$$

$$k(acid)_T = 10^{(X-pD)}. \quad (10)$$

Similarly, temperature-corrected base catalyzed and water catalyzed equations could also be derived. Substituting all the three temperature-corrected parameters in the equation (5), the following equation (11) is obtained. Hence, the equation (11) calculates k_{rc} values of labile protons in proteins in a single step accounting for the effect of neighbouring amino acids, pH and temperature.

$$k_{rc} = 10^{(X - pD)} + 10^{(Y + pD - pKD)} + 10^{(Z)}. \quad (11)$$

where,

$$X = \log k_a + \log A_L + \log A_R + (E_a/4.57)*(1/293 - 1/T). \quad (12)$$

$$Y = \log k_b + \log B_L + \log B_R + (E_b/4.57)*(1/293 - 1/T). \quad (13)$$

$$Z = \log k_w + \log B_L + \log B_R + (E_w/4.57)*(1/293 - 1/T). \quad (14)$$

In the above equations, A_L and A_R denote the effect of residues that are on the left and right sides of amide protons at acidic conditions, respectively. Similarly, B_L and B_R denote the effect of residues that are on the left and right sides of amide protons at basic conditions, respectively [7].

Table 1: Comparing the actual and the predicted (by CIntX program) conformations of Xaa-Pro peptide bonds in Cytochrome C, Cobrotoxin and RNase A.

Protein (PDB ID)#	No. of Prolines	Position of Prolines	Distance measured by manual method*		Actual Conformation of Xaa-Pro peptide bond	Distance calculated by CIntX		Predicted conformation of Xaa-Pro peptide bond
			C_α-C_α	C_α-C_δ		C_α-C_α	C_α-C_δ	
1HRC	4	30	3.78	2.80	*Trans*	3.781	2.796	*Trans*
		44	3.81	2.96	*Trans*	3.814	2.957	*Trans*
		71	3.81	2.85	*Trans*	3.809	2.854	*Trans*
		76	3.79	2.64	*Trans*	3.790	2.638	*Trans*
1COD	2	12	3.69	2.64	*Trans*	3.694	2.644	*Trans*
		44	3.66	2.02	*Trans*	3.658	2.017	*Trans*
5RSA	4	42	3.86	2.90	*Trans*	3.858	2.898	*Trans*
		93	3.04	3.88	*Cis*	3.039	3.883	*Cis*
		114	2.91	3.82	*Cis*	2.912	3.817	*Cis*
		117	3.83	2.88	*Trans*	3.818	2.878	*Trans*

#1HRC: Cytochrome C, 1COD: Cobrotoxin and 5RSA: Rnase A.

*PyMol, molecular visualization tool [22], was used for determining the conformations of Xaa-Pro peptide bonds in the proteins manually. Distances are shown in angstrom.

Table 2: Comparing the actual and the predicted (by CIntX program) cysteine(s) and cystine(s) of Cytochrome C, Cobrotoxin and Rnase A.

Protein (PDB ID)	No. of Cys[#]	Position of Cys	Distance measured by manual method[*]		Actual Conformation of Cys residue	Distance calculated by CIntX		Predicted conformation of Cys residue
			S-S Pair	Distance		S-S Pair	Distance	
1HRC	2	14 17	none	8.71	Cysteine	none	8.712	Cysteine
1COD	8	3 24	C3-C24	1.40	Cystine	C3-C24	1.404	Cystine
		17 41	C17-C41	1.44	Cystine	C17-C41	1.442	Cystine
		43 54	C43-C54	1.33	Cystine	C43-C54	1.331	Cystine
		55 60	C55-C60	1.84	Cystine	C55-C60	1.836	Cystine
5RSA	8	26 84	C26-C84	2.04	Cystine	C26-C84	2.037	Cystine
		40 95	C40-C95	1.97	Cystine	C40-C95	1.968	Cystine
		58 110	C58-C110	1.95	Cystine	C58-C110	1.954	Cystine
		65 72	C65-C72	1.94	Cystine	C65-C72	1.941	Cystine

[#]Cys denotes cysteine residue.
[*]The manual distance measurements for determining the cysteine and cystine residues in the proteins were carried-out using PyMol molecular visualization tool [22]. Distances are shown in angstrom.

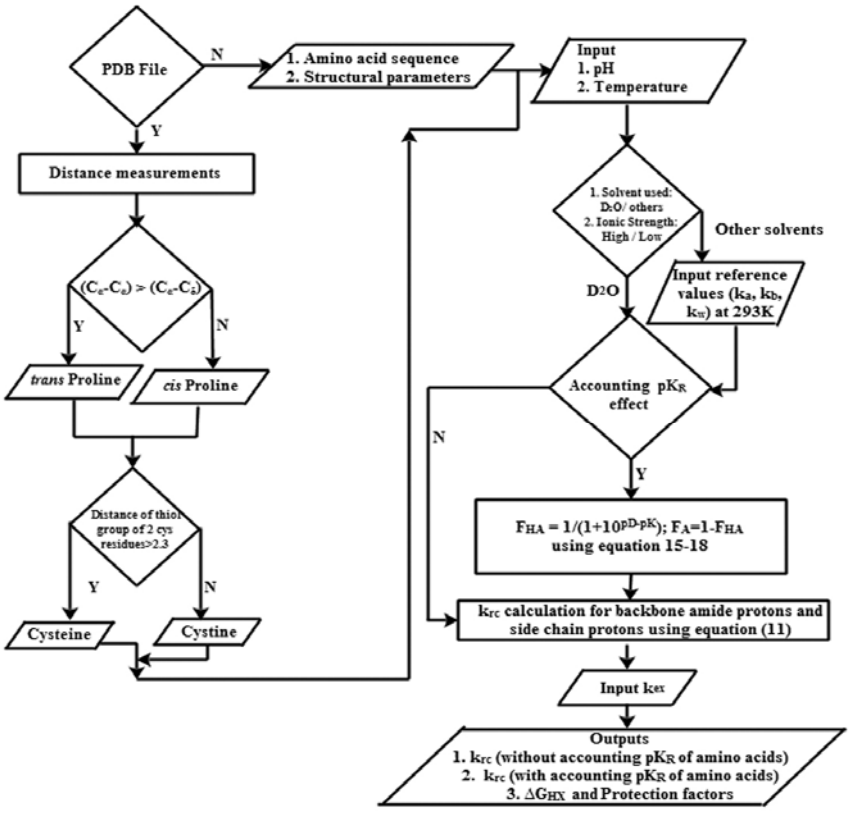

Figure 1. Flowchart outlines key-steps of the CIntX program.

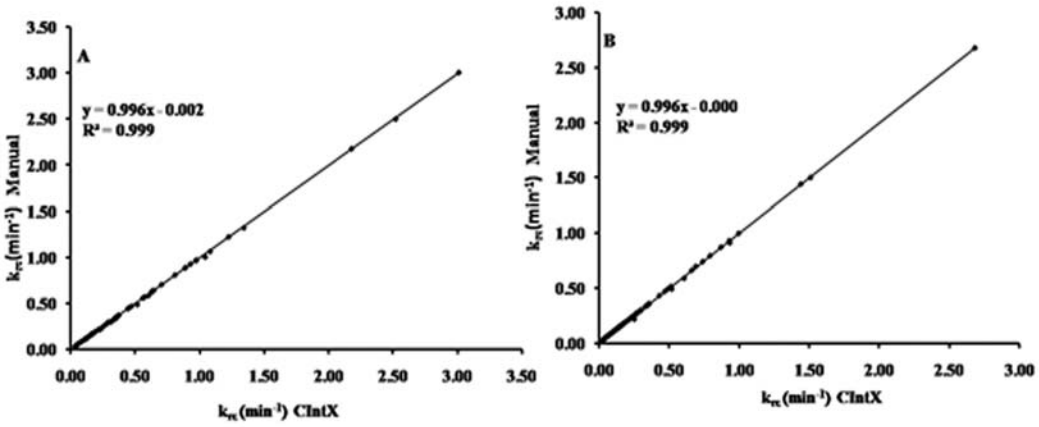

Figure 2. Correlation between the k_{rc} values estimated by manual calculations and the CIntX program for the NHs in proteins (A) Coborotoxin (1COD) and (B) Cytochrome C (1HRC) at pH 3.2, 298 K.

3.2 Determining prerequisite structural parameters for estimating k_{rc} of labile protons in proteins

Proline residues are linked to its preceding amino acid in polypeptides through an imide bond instead of amide linkages by which other amino acids are connected. Amide bonds prefer negligible percentage of (about 0.03%) *cis*-conformations in the unfolded states, whereas imide bond (Xaa-Pro) prefers remarkable percentage of *cis*-conformations in the unfolded states and the percentage varies (6-38%) depending on the chemical properties of the residue (Xaa) preceding proline [12,13]. The Xaa-Pro bond isomerises between its *cis* and *trans* conformations in the unfolded states of proteins under native conditions. In order to accurately predict the H/D exchange behaviour of labile protons influenced by the Xaa-Pro peptide bonds, the program determines the conformations of Xaa-Pro peptide bonds in proteins using the equation (2). Table 1 validates the robustness of the program on predicting *cis*-*trans* proline conformations in proteins such as Cytochrome C (1HRC), Cobrotoxin (1COD) and Rnase A (5RSA). A quick inspection to the table suggests that the actual conformations and the CIntX predicted conformations of proline residues in the proteins are in perfect agreement. Similarly, the k_{rc} of NHs of cysteine and cystine residues in proteins remarkably differ from each other [7]. Moreover, the k_{rc} of residues that are succeeded by cysteine/cystine residues are also differentially influenced by them. In order to account the effect of these residues, CIntX determines the conformations of each cysteine residue in proteins. When the distance between the sulphur atoms of thiol groups from two cysteins is within 2.3 Å, the two cysteines are considered to be disulfide bonded [14]. The program determines the cysteine/cystine conformation by sampling all possible cysteine pairs in proteins. The distance between any two sulphur atoms is calculated using the equation (2) described in the method section. Table 2 shows robustness of the program on predicting the cysteine/cystine residues in Cytochrome C, Cobrotoxin and Rnase A.

3.3 Estimation of k_{rc} values of backbone labile protons in proteins from their three-dimensional (3D) structures

The outline of the program is schematically represented in Fig.1.The program calculates the values of k_{rc} for labile protons in proteins with four inputs: PDB file of the protein, pH, Temperature, and Ionic strength. In outline, the program has four stages for each complete cycle. First, the program determines number of proline and cysteine residues and their conformations using the input PDB file, itself (discussed in the above section). Second, k_{rc} of labile protons are calculated using equation (11), which accounts the effect of pH, temperature, and primary structure of proteins on the k_{rc}. The effect of neighbouring residues on the NH is calculated in the program as shown herein.

/* Calculation of k_{rc} for the NH of s_3 of S */
$S = (s_1, s_2, s_3.....s_n)$;
for ($i=1$; $i<n$; i++)
$R_a = As_2 * As_3$;
$R_b = Bs_2 * Bs_3$;
$k_{rc} = k_a * R_a * 10^{(-pD)} + k_b * R_b * 10^{(pD-pKD)} + k_w * R_b$
end for;

In the above representations, '*S*' denotes total number of residues in the given input polypeptide chain. Each amino acid in the polypeptide chain is denoted as '*s*' and its subscript defines the position of the amino acid in the chain. Residue at i-1th position and ith position are considered to be on right and left parts to the NH of ith residue, respectively. *A* and *B* are standard reference parameters of the amino acids at acidic and basic/neutral conditions, respectively [7,15]. Moreover, depending on the ionic strength (high > 0.5 M or low < 0.5 M) and solvents (H_2O or D_2O) defined, the program assigns the values for k_a, k_b and k_w that are stored as default parameters in the program and calculates k_{rc} for NHs of proteins in the second stage. The k_{rc} values of NHs in Cobrotoxin (1COD) and Cytochrome C (1HRC) have been calculated using the program and the calculated data have been compared with the k_{rc} values of the NHs as determined by manual method (Fig. 2). The data in Figure 2A and 2B

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.