IMPORTANCE OF C-TERMINAL KNOT STRUCTURE IN CARBONIC ANHYDRASE II (PART I): ROLE OF C-TERMINAL KNOT FORMING GLN253 ON ENZYMATIC PROPERTY OF HUMAN CARBONIC II

Mohammad Taufiq Alam[1]

Department of Applied Chemistry and Chemical Technology
University of Rajshahi, Rajahahi-6205, Bangladesh
e-mail: alam@ru.ac.bd

Abstract

In both, bovine and human carbonic anhydrase II, a conserved glutamine residue occupies the position in the middle of the knot, which is formed by intercrossing of C-terminal end with N-terminal region. Previous studies have indicated that C-terminus is not the part of an active site, but truncation of 7 amino acid residue in this region can have marked effects on stability of the enzyme (data not published). To gain further insight into the role of specific amino acid residue in C-terminal region, site directed mutagenesis was used to introduce point mutation. Substitution of glutamine with cysteine was chosen because the cysteine residue is less hydrophilic as compared with glutamine and thus, may disrupt the hydrophilic environment in this region. Result indicates that Gln253 located within the C-terminus knot topology plays a significant role in normal function of the enzyme. Thus, C-terminal region might mediate cooperativity between the central active site of the enzyme through proper formation of knot.

Key words: Human carbonic anhydrase II, knot topology, point mutation

Introduction

An extensive computational search of protein data bank together with our present knowledge of protein folding shows that there are several proteins having a trefoil knot topology in their molecular structure and some carbonic anhydrases (CAs) belongs to this group (Taylor 2000, Takusagawa 1996 and Liang and Mislow 1994). CAs are ubiquitous metalloenzymes, virtually present in all living organisms, including animals, plants, algae, bacteria, and archaebacteria. These enzymes catalyze the reversible hydration of CO_2 to form HCO_3^-, which is involved in many biosynthetic reactions, among which are gluconeogenesis, lipogenesis, the synthesis of certain amino acids, and pyrimidine nucleotide biosynthesis. Moreover, these enzymes are involved in pH homeostasis, bone resorption, ion transport, electrolyte secretion in a variety of tissues, calcification, and tumorigenesis (Chegwidden 2000). Humans express at least 14 CA isozymes (HCA I-XIV) with varying catalytic activity and cellular distribution (Chegwidden 2000 and Duda 2004) among which HCA I, HCA II, HCA IV and HCA V posses distinct trefoil knot topology in their molecular structure. HCA II is the most extensively studied of these isozymes because it has widespread tissue distribution with large amounts located in red blood cells, where it is involved in respiration, and because it is among the most efficient of the HCAs with a catalytic turnover rate of 106 s^{-1} (Khalifah 1971).

---

[1] To whom all correspondence should be addressed.

Fig. 1 a shows a "ribbon" representation of the 3-D structure of HCA II obtained from X-ray crystallography (Eriksson *et al.* 1988 and Kannan *et al.* 1975). The molecule consists of a single polypeptide chain with 259 amino acid residues giving a molecular weight of ~29,000. The C-terminal is intercrossed with Gln253 to form a knot structure (Fig. 1, black arrow).
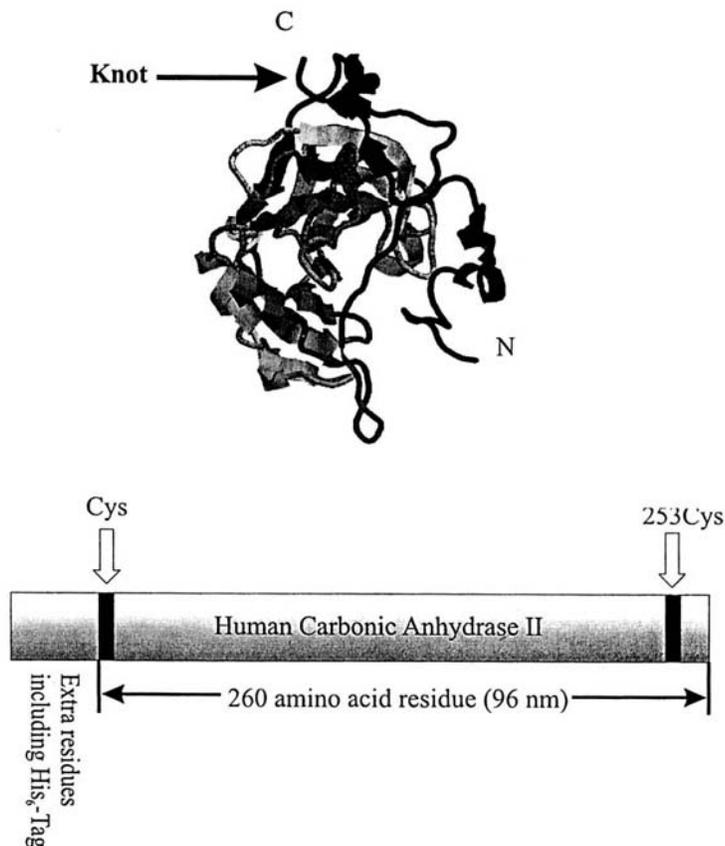

Fig. 1. Three-dimensional structure of HCAII indicates the presence of a knot structure close to it C-terminus (Eriksson 1988). Lower panel is the schematic representation of the construct encoding engineered HCAII. It was constructed by adding an extra Cys residue at N-terminus and replacing Gln253 with Cys.

The importance of the unique C-terminal folding pattern of HCA II having a trefoil knot was implicated in a series of refolding studies done by Carlsson and his colleagues (Fransson 1992, Carlsson 1974 and Freskgard 1991). They showed that HCA II could be refolded from the denatured state with several distinct kinetic steps, the slowest among them having a half-time of 9 min. Using various labeling techniques, they showed that refolding of the C-terminal region constituted one of the slow reactions and was almost synchronous with the recovery of the enzymatic activity. They also showed that proteolytic truncation of three consecutive amino acid residues from the C-terminus of HCA I under denaturing conditions reduced the extent of recovery of refolded protein to only 25% of the starting material and the refolded protein showed decreased stability (Carlsson 1974). These observations on HCAs indicate that the C-terminal sequence

involved in the knot topology contains vital information for the correct folding of the enzyme leading to the formation of an active site. In order to better understand the role of knot on enzymatic activity, Gln253 was replaced with Cys. Gln253 has a propensity to form helix structures; in contrast Cys has a propensity to form beta-sheet structure (Xuhua and Xie 2002). Thus, one can expected, such a destabilization in the region of intercrossing may influence the proper formation of knot, as a result the evolved protein will show no or reduce catalytic activity. Using this variant (Gln253Cys), I have showed that mutant protein only produce 50% of the original activity in compared with wild type HCA II.

Materials and Methods

Construction of the expression plasmid

A mutant (HCAIICys205Ser) gene of human carbonic anhydrase II was given by Dr U Carlsson, Linköping University, Sweden. Two cysteine residues were introduced in the protein at both of its N- and 253 positions (N/253), where Gln253 was replaced with Cys. The construction of the gene which encode N/253 was carried out by PCR-based site-directed mutagenesis (Sambrook and Russell 2001) using the following primer set {5'-GGGCATATGCACCACCACCACCACGGTGGT- AGCGGTGGTAGCTGTGCCCATCACT-GGGGGTAC-3' and 5'-GGGCTCGAGTTAA-CATTTGAAGGAAGCTTTGATTTG-CCTGTTCTTCA-3'} and inserted into expression vectors (pET) (Novagen, Madison, WI) between *Nde* I and *Xho* I sites (Invitrogen, USA), and the sequences were confirmed.

Expression and purification of mutant protein

The constructed recombinant plasmid pET was transformed into competent *Escherichia coli* BL21. One colony of *E. coli* BL21 harboring the plasmid pET was grown overnight at 37°C in 5 ml SOB medium supplemented with kanamycin (50 μg/ml). The whole culture was inoculated into 500 ml fresh SOB medium and the culture was grown to an OD_{600} of 0.5-0.6 at 37°C. Mutant protein expression was initiated by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to 0.1 mM. After 4 h induction, cells were collected by centrifugation at 12,000 rpm for 10 min at 4°C, resuspended in lyses buffer (50 mM phosphate buffer saline containing 300 mM NaCl, and 10 mM imidazole pH 8), and lysed on ice by sonication at 400 W for 100 cycles (5 s working, 5 s free). The supernatant of the cell lysate resulting from centrifugation at 12,000 rpm at 4°C for 20 min was applied to a Ni-NTA chelating column (Qiagen, Hilden, Germany). After extensive washing with binding buffer, then with the same buffer containing 20 mM imidazole, the mutant protein was eluted with six column volumes of elution buffer (50 mM phosphate buffer saline containing 300 mM NaCl, and 250 mM imidazole pH 8). The peak fractions containing the mutant protein were pooled and dialyzed overnight at 4°C against Tris-SO$_4$, pH 7.5.

Size Exclusion Chromatography

The dialyzed mutant protein was subsequently applied to size exclusion chromatography (SEC) through a 100 μL sample loop, using the column Superdex 75 HR 10/30 (Amersham-Pharmacia Biotech, Sweden) installed on the HPLC-system (Jasco PU 1580, Japan). Elution with 50mM Tris-SO4 buffer, pH 7.5 at a flow rate of 0.5 mL/min was carried out.

Protein determination

The concentration of purified protein was determined by measuring the absorbance at 280 nm using a UV/Vis spectrophotometer (Shimadzu, Japan). An absorption coefficient at 280 nm of ε_{280}=8730 M^{-1}cm^{-1} was used to determine concentrations (Gill and Hippel 1989).

Esterase activity measurement

Enzyme activity was measured on a Shimadzu UV-1601 spectrophotometer using *p*-nitrophenyl acetate (*p*NPA) as substrate. This assay was performed as described by Whitney (1973) with some modifications. In this method, esterases catalyse the breakdown of *p*-nitrophenyl acetate, an artificial substrate, into *p*-nitrophenol and acetate. The conditions of enzyme activity are 25 °C and pH 7.5. 50 μl of enzyme solution were mixed with 930 μl 50 mM Tris-SO$_4$ buffer (pH 7.5), and maintained at 25 °C for 5 min in a 1 ml spectrophotometer cuvette. The reaction was started by addition of the 20 μl of substrate solution (50 mM *p*NPA dissolved in acetone and stored at –20 °C). Increase in absorbance at 348 nm was read against a blank without enzyme for 5 min. One enzyme unit (IU) was the amount of enzyme liberating 1 μmol of *p*-nitrophenol (*p*NP) per minute under the assay conditions. All measurements were performed three times and the mean values were calculated.

N-terminal amino acid sequence and amino acid composition analysis

Edman degradation was performed with a gas-phase sequencer (Shimadzu, Japan). Identification and quantification of phenythiohydantoin derivatives was done by reversephase HPLC with an internal standard for quantification. For amino acid analysis, protein samples were hydrolysed in constant-boiling HCl at 110°C for 24 h in nitrogen-flushed evacuated tubes in the presence of 1 mg/ml phenol. Hydrolysates were analysed RP-HPLC column.

Fluorescence measurements

Intrinsic fluorescence measurements of proteins were done using a Model Shimadzu RF-5300PC spectropolarimeter. The samples were continuously mixed by a magnetic stirrer and the temperature of the samples was maintained at 25°C using a thermostatically controlled cell holder. Excitation and emission slits with a nominal band pass of 5 nm and 1 cm path length quartz cuvettes which are coated on two surfaces were used for all measurements. Background intensities of samples were subtracted from each sample spectrum to cancel out any contribution due to the solvent Raman peak and other scattering artifacts. The excitation light was vertically polarized and a magic-angle polarizer preceded the emission monochromator. All emission spectra were collected at an excitation wavelength of 280 mm.

Results and Discussion

Mutant Protein Expression

To examine the specific roles of Gln253 within the knot structure, we expressed an engineered protein of HCA II in which Gln253 was replaced with Cys. For this purpose, a mutant gene (Gln253Cys) was prepared through PCR-based site-directed mutagenesis. The strategy of synthesis is described in the Materials and Methods, and the final PCR product digested with two restriction enzymes (*Nde* I and *Xho* I) was cloned in the expression vector pET and the sequence of gene was confirmed by automated DNA sequencing.

Recombinant plasmid pET-Gln253Cys containing accurate gene was transform into *E. coli* BL21(DE3).The construct for Gln253Cys expression, containing a 6xHis-tag for affinity purification and an enterokinase (EK) cleavage site for removing extra amino acid sequence. After transformation into the expression host *E. coli* BL21 (DE3), the mutant protein was cultured in SOB media. When the culture reached at the mid-logarithm time, *E. coli* cells harboring plasmid was treated with 1mmol/L IPTG.

Extraction and purification of mutant protein

Many recombinant proteins expressed at high levels in *E.coli* form insoluble cytoplasmic inclusion bodies (Schoner *et al.* 1985), which require extraction with denaturing solvents followed by 'renaturation' steps before the protein in question assumes its active form. To determine whether the recombinant Gln253Cys was soluble, we subjected cell extracts to the differential centrifugation scheme described in Materials and Methods. Gln253Cys protein was found exclusively in the high-speed supernatant fraction (Fig. 2, lane 1). This is in contrast to the fractionation of aggregated recombinant proteins which are usually found in the low-speed pellet (Schoner *et al.* 1985)**.** This observation indicates that Gln253Cys protein was contained within the cytoplasmic space in a soluble form. SDS-PAGE of expressed protein at various stages in the purification scheme is shown in Fig.2. Moreover, the existence of the 6xHis-tag on the carrier protein provided an effective means of purification for the mutant protein by Ni^{2+}-chelating chromatography. SDS-PAGE analysis revealed the effectiveness of this purification step (Fig. 2 lane 6-7). The protein migrated as a single band with an estimated relative molecular mass *(M,)* of ~29,114 which is close to that predicted by the covalent structure (Auron 1984 and March *et al.* 1985). A small amount of *E. coli* proteins were also detected by Coomassie blue staining. Gel filtration HPLC was carried out where the protein eluted as a symmetrical peak at 21.5 min with Tris-SO4 buffer (Fig. 3). The purity of the protein was also verified by SDS PAGE (Fig. 2, lane 8).
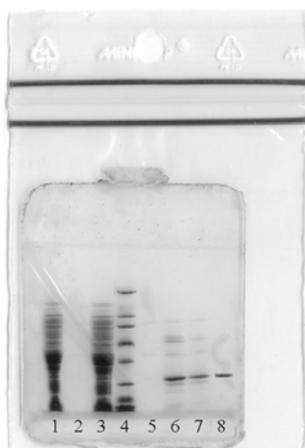

Fig. 2. Purification of Gln253Cys. At each step in the purification an aliquot of the sample was subjected to SDS-PAGE and the proteins were visualized by staining with Coomassie blue. Lanes 4 is standard proteins (from the top: 200,000; 116,250; 87,400; 66,200; 45,000; 31,000; 21,500;14,400 and 650. Lanes 1-3 and 5–8 are various steps in the purification: 1, crude extract; 2, washing; 3, flow; 6-7 elution on Ni-NTA column and 8 purified enzyme after gel filtration.
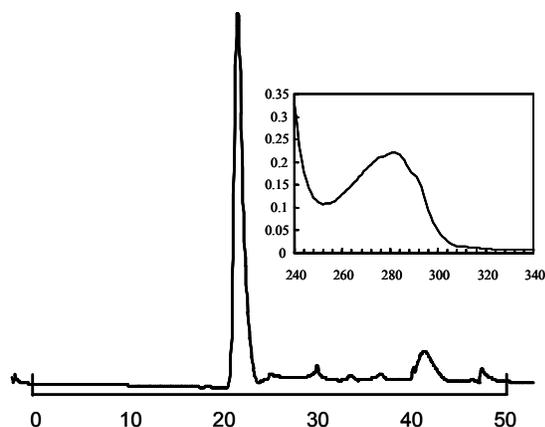
Fig. 3 Gel filtration of Gln253Cys on Superdex 75 HR 10/30 HPLC column. Conditions were as described in Materials and Methods. About 50 μl of Gln253Cys in 100 mM Tris-SO4, pH 7.5 containing 2 mM sodium azide, was analysed. The inset shows ultraviolet absorption spectra of the mutant protein. Spectra was recorded at room temperature in 50 mM Tri-SO4 buffer, pH 7.5.

Spectroscopic Properties

The absorption spectrum of Cys-substituted HCA II at pH 7.5 in Tris-SO$_4$ buffer is shown in Fig. 2 (inset). There is one distinct maximum at 280 nm. The absorption spectrum is quite similar to that of wild-type HCA II. Fluorescence has provided to be a particularly valuable technique for studying proteins structure. HCA II contains 7 tryptophan and 8 tyrosine residues in their molecular structure. These residues are located in different part of the 3D structure. Excitation at 280 nm results in both tryptophan and tyrosine fluorescence, so it is easy to use fluorescence spectroscopy at this wavelength to monitor the global structure of the protein. The emission spectra of the Gln253Cys protein together with wild-type HCA II were shown in Fig 4. Excitation at 280 nm showed a single emission peak with a λ_{max} at 340.0 together with a shoulder at 305 nm respectively. Considering the amounts of tyrosine and tryptophan, one could expect two emission peaks. But even one tyrosine and tryptophan residue can give a spectrum with a peak corresponding to tryptophan emission in the interval 331–342 nm at excitation wavelength 280 nm (Khan *et al.* 1980 and Gruen *et al.* 1987). This is due to the fact that if two molecules are within förster distance, there should be a transfer of energy from the excited tyrosine to tryptophan, which corresponds to a single emission peak (Khan 1980 and Gryczynski 1988). Crystallographic structure of HCA II showed that 5 tyrosine molecules are located within the length of förster distance. Thus the present result is consistence with the 3D structure of the protein. The similarities of emission spectrums indicate that all the tryptophan and tyrosine residues of wild-type and genetically engineered proteins are located in similar apolar environments.

The Authenticity of the recombinant protein

In order to confirm its authenticity, recombinant protein was subjected to N-terminal amino acid sequencing, amino acid composition and MALDI-TOF mass spectroscopy analysis. According to the sequencing result, the first six amino acid of N-terminal of recombinant protein was SER, HIS, HIS, TRP, GLY and TYR which was as same as the native HCA II sequence. On the other hand, the result of amino acid composition

analysis indicated that the amino acid composition of the recombinant protein was also in accordance with the anticipated composition (data not shown). Moreover, the molecular weight of recombinant protein analyzed by MALDI-TOF mass spectroscopy was 29114 Dalton, which was similar to the theoretical molecular weight of HCA II (Fig. 5)
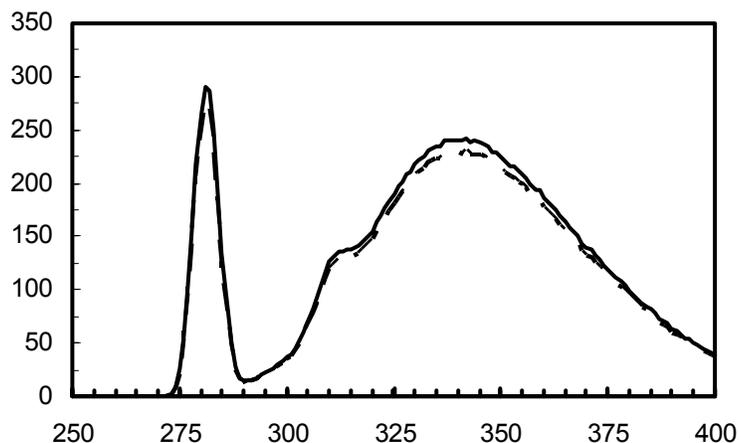

Fig. 4. Intrinsic fluorescence spectra of native (dot line) and engineered protein (solid line) in TrisSO4 buffer at pH 7.5. The excitation wavelength was 280 nm and the protein concentration was slightly different for each spectrum. When normalized with respect to concentration, the fluorescence intensities of the two were virtually identical.

Esterase Activity

After conforming the structural similarity, enzymatic activity of engineered protein was measured. The result indicates that the activity was about 50% of that of a commercial wild-type HCA II. An addition of Zn^{2+} did not increase the activity significantly. This indicates that either half of the Gln253Cys protein molecules were not in active state, or the particular mutation employed in this experiment reduced the activity of all the molecules to abot one half that of the wild-type enzyme.
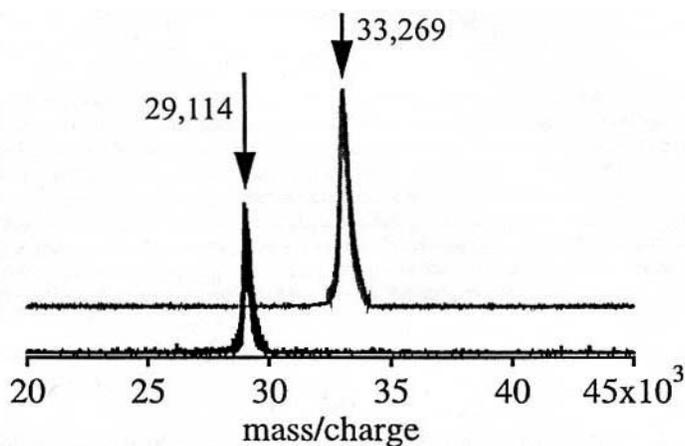

Fig. 5. MALDI-TOF mass spectra of HCA II together with engineered protein. One μ l of sample and 1 μ of a saturated α-cyano-4hydroxy-trans-cinnamic acid (CHCA) solution in 50% acetonitrile, 0.1% of trifluoroacetic acid (TFA) were spotted on a stainless steel MALDI target. Mass spectra were acquired over the range 15,000-30,000 Da.

In summary, we have replaced one hydrophilic residue located in the knot structure of HCA II with residues of approximately equivalent size but with less hydrophilic characteristics (Gln253Cys). Amino acids interact with each other, especially with neighboring amino acids as well as surrounding water molecules, to generate protein structures. Different amino acids have different preferences either for or against water. Glutamine residue in wild-type HCA II is hydrogen-bonded to conserved water in a network of ordered hydrogen-bonded waters observed in the X-ray crystal structure. Though the expressed and purified protein posses similar structure from spectroscopic point of view, however, we have determined that these substituted amino acids have effects on both esterase catalysis activity as well as quaternary structure of the enzyme (direct evidence will appear in the part-II of the series of publication). HCA II is a metalloenzyme that contains a zinc ion (Zn^{2+}) in its active site which is positioned deep in the central cleft of the enzyme and liganded to the ε-Ns of His-93 and His-95, the δ-N of His-119, and water molecules (Eriksson 1988). This active site is located far form the C-terminal Gln253. Therefore, it is not clear how the above-mentioned amino acid residue that is not in direct contact with the active site amino acid residues induced a direct effect on the active site. However, observations that make it likely that at least some of these hydrogen bonds are disrupted as a result the evolved enzyme exhibit reduced enzymatic activity.

Acknowledgements

References

Auron P E, Webb A C, Rosenwasser L J, Mucci S F, Rich A, Wolff S M and Dinarello C A (1984) Nucleotide sequence of human monocyte interleukin-1 precursor cDNA. *Proc Natl Acad Sci* **81**: 7907–7911

Carlsson U, Henderson L E, Nyman P O and Samuelsson T (1974) Studies on the influence of carboxyl-terminal amino acid residues on the activity and stability of human erythrocyte carbonic anhydrase B. *FEBS Lett*. **48**: 167-71.

Chegwidden W R, Dodgson S J and Spencer I M (2000)The Roles of Carbonic Anhydrase in Metabolism, Cell Growth and Cancer in Animals. In *The Carbonic Anhydrases: New Horizons*; Chegwidden W R, Carter N D, Edwards Y H Eds., Birkhäuser Verlag: Basel, Switzerland pp 343-363.

Duda D M and McKenna R (2004) *Handbook of Metalloproteins* (A Messerschmidt A, Ed.) pp 249-263, John Wiley & Sons, New York.

Eriksson A E, Jones T A and Liljas A (1988) Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins* 4: 274-82.

Fransson C, Freskgård P O, Herbertsson H, Johansson A, Jonasson P, Martensson L G, Svensson M, Jonsson B H and Carlsson U (1992) Cis-trans isomerization is rate-determining in the reactivation of denatured human carbonic anhydrase II as evidenced by proline isomerase. *FEBS Lett* **296:** 90-94.

Freskgard P O, Carlsson U, Martensson L G and Jonsson B H (1991) Folding around the C-terminus of human carbonic anhydrase II. Kinetic characterization by use of a chemically reactive SH-group introduced by protein engineering. *FEBS Lett* **289:** 117-122.

Gill S C and von Hippel P H (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**: 319−326.

Gruen L C, Guthrie R E and Blagrove R J (1987) Structure of a major pea seed albumin: Implication of a free sulfhydryl group. *J. Sci. Food Agric.* **41:** 167–178.

Gryczynski I, Eftink M and Lakowicz J R (1988) Conformation heterogeneity in proteins as an origin of heterogeneous fluorescence decays, illustrated by native and denatured ribonuclease T1. *Biochim. Biophys.Acta* **954 :** 244–252.

Kannan K K, Notstrand B, Fridborg K, Lovgren S, Ohlsson A and Petef M (1975) rystal structure of human erythrocyte carbonic anhydrase B. Three-dimensional structure at a nominal 2.2-A resolution. *Proc. Natl. Acad. Sci. USA* **72**: 51-55.

Khalifah R G (1971) The carbon dioxide hydration activity of carbonic anhydrase, *J. Biol. Chem.* **246** :2561-2573.

Khan S M, Darnell D W and Birnbaum E R (1980) Tyrosine fluorescence as a measure of denaturation in thermolysin *Biochim.Biophys. Acta* **624 :** 1–12.

Liang C and Mislow K (1994) Knots in proteins. *J. Am. Chem. Soc.* **116** : 11189-11190.

March C J, Mosley B, Larsen A, Cerretti D B, Braedt G, Price V, Gillis S, Henney C S, Kronheim S R, Grabstein K, Conlon P J, Hopp T P and Cosman D (1985) Cloning, sequence and expression of two distinct human interleukin-1 complementary DNAs. *Nature (Lond.)* **315 :** 641 -647.

Sambrook J and Russell D W (2001) *Molecular Cloning: A Laboratory Manual*. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schoner R G, Ellis L F and Schoner B E (1985) Isolation and purification of protein granules from *Escherichia coli* cells overproducing bovine growth hormone. *Biotechnology* **3**: 151−154.

Taylor W R (2000) A deeply knotted protein structure and how it might fold. *Nature* **406**: 916 - 919.

Takusagawa F and Kamitori K (1996) A real knot in protein. *J. Am. Chem. Soc.* **118 :** 8945–8946.

Whitney P L (1973) Monocarboxamidomethyl Carbonic Anhydrase Purified by Affinity Chromatography. *J. Biol. Chem.* **218**(8): 2785-2789.

Xuhua Xia and Zheng Xie (2002) Protein Structure, Neighbor Effect, and a New Index of Amino Acid Dissimilarities. *Mol. Biol. Evol.* **19**(1):58–67