THESIS FOR DEGREE OF DOCTOR OF PHILOSOPHY

STRUCTURAL INVESTIGATION OF TRANSGLUTAMINASES

Attila Ambrus

University of Debrecen
Medical and Health Science Center
Department of Biochemistry and Molecular Biology
Debrecen, 2001

STRUCTURAL INVESTIGATION OF TRANSGLUTAMINASES

Thesis for degree of Doctor of Philosophy

Attila Ambrus

Supervisor: Prof. Dr. László Fésüs

University of Debrecen
Medical and Health Science Center
Department of Biochemistry and Molecular Biology

Debrecen

2001

INTRODUCTION

Transglutaminases are Ca^{2+} dependent acyltransferases that catalyze the formation of amide bonds between the γ–carboxamide groups of peptide-bound glutamine residues and the primary amino groups of different amines including the ε–amino groups of lysine sidechains in proteins. For understanding the TGase-reaction, the most critical data are related to the Ca^{2+} activation processes of these enzymes. We investigated in this study the FXIII-A_2, TGase 2 and TGase 3 (each of human origin).

Two forms of FXIII exist. The plasma enzyme contains two A and two B subunits together in a heterotetramer, while the cellular one consists of two identical A subunits. Plasma coagulation factor XIII is the last enzyme in the blood coagulation cascade stabilizing the fibrin soft clot making it unsusceptible to fibrinolysis by crosslinking fibrin and/or α_2-antiplasmin (the inhibitor of plasmin) onto the clot. The function of the cellular enzyme is still obscure. By X-ray crystallography, the structure of the cellular zymogen has solved at 2.1Å resolution. ^{43}Ca NMR has elucidated that the B subunit does not bind Ca^{2+}. X-ray studies also showed that each A subunit binds one Ca^{2+} with high affinity causing very small conformational changes, while equilibrium dialysis suggested two tightly bound Ca^{2+} and up to six more with smaller affinities on the FXIII-A_2B_2 tetramer at slightly higher Ca^{2+} concentrations. The overall Ca^{2+} affinity of plasma FXIII was determined to be Kd≅0.1 mM. In physiological conditions, both plasma and cellular FXIII require a proteolytic cleavage (by thrombin) between Arg37 and Gly38 for their activation. However, *in vitro,* in the presence of low Ca^{2+} and high inert salt concentration the cellular FXIII, while at high Ca^{2+} and low salt concentration both forms of FXIII assume active conformations without any proteolytic cleavage. There have been no data on the Ca^{2+} binding properties of FXIIIs under these conditions. Earlier, it is elucidated that several divalent and trivalent metal ions except Mg^{2+} can be bound to plasma FXIII.

The human tissue-type transglutaminase (TGase 2) has been implicated in stimulus secretion coupling, receptor mediated endocytosis, programmed cell death, extracellular matrix organization and cell adhesion, cell growth and proliferation and tumor growth. The protein, which is a monomer molecule, could not be crystallized, yet. However, because of homology to FXIII it is possible to build its homology model. Recently, a 2.5Å resolution X-ray structure (relatively low resolution compared to the available FXIII structure the resolution of which is 2.1Å) of the Red Sea Bream TGase 2 is published. This structure contains 666 amino acids from the total of 695. The sequence identity of this fish TGase 2 to its human analogue is 43.6%.

The number of Ca^{2+} binding to one molecule of human TGase 2 was estimated to be six in high affinity sites by equilibrium dialysis and all binding sites were proposed to have similar Ca^{2+} affinities giving a hyperbolic saturation curve with an apparent affinity constant of 90μM titrated at

low Ca^{2+} concentrations. At higher Ca^{2+} concentrations, a “fast-moving” (on SDS gel) form of TGase 2 was detected, and the existence of low affinity Ca^{2+} binding sites was proposed without giving their number, localization and affinity constants.

GTP lowers the Ca^{2+} affinity and transamidase activity of human TGase 2 binding to the core domain of the molecule at the amino acids 159-173 and causing considerable conformational changes; however it has not been clarified how many and which Ca^{2+} binding sites are affected. Ca^{2+} and GTP can also influence the proteolytic and folding behavior of the enzyme suggesting another level of pharmacological regulation related to the two affectors.

The monomer TGase 3 is essential in keratinization participating in the formation of the insoluble cornified envelope in the dying keratinocytes. There have been no studies so far regarding its Ca^{2+} binding properties. It is known that the activity of the enzyme is retained in a 50kDa part after the limited proteolysis necessary for its activation; it is likely that this fragment binds Ca^{2+}.

There are several proteins studied by ^{43}Ca, ^{25}Mg and other metal NMR using different theoretical and technical approaches. These methods provide unique and completely ion-specific means of obtaining both qualitative and quantitative information about the metal ion binding proteins.

During the production of a recombinant protein, we often face to the problem of solubility of the product. If the amount of the expressed protein in the cytoplasm is not considerable, it is very probable that it forms so-called *inclusion bodies.* The development of the successful refolding technique provides the recombinant active protein as well as kinetic and structural data relating to the folding properties of the molecule.

OBJECTIVES

Transglutaminases have several physiologically important functions; therefore, detailed structural studies are necessary to clarify the related molecular mechanisms. The results above clearly show that a number of unanswered questions exist about the Ca^{2+} binding properties of TGases. Therefore, we have decided to initiate a series of ^{43}Ca NMR studies combined with surface polarity analysis and multiple sequence alignment targeting three TGases. We have intended to reveal all potential Ca^{2+} binding sites on these TGases. The construction of a precise homology model of human TGase 2 was also desirable for this purpose. Furthermore, the determination of the dissociation constants and binding site symmetry parameters of Ca^{2+} complexes of the enzymes and the ^{25}Mg NMR study of FXIII-A_2 were also considered to be done in the first part of our investigations. Our second focus was related to the elucidation of the Ca^{2+}/GTP effect on the folding properties of the recombinant human TGase 2.

MATERIALS AND METHODS

NMR experiments

^{25}Mg experiments were performed at 22.05MHz (360MHz of proton) with a 5mm inverse broad-band probehead on a Bruker AM-360 spectrometer. Typically 256 transients were used in order to obtain good signal to noise ratio for integration. The ^{43}Ca NMR experiments were performed at 33.7MHz (500MHz of proton) on a Bruker DRX-500 spectrometer with a 10mm direct broad-band probehead. Typically 64 or more transients were applied for good signal to noise ratio. T_1 measurements were carried out by inversion recovery experiments using the standard Bruker microprogram. T_2 measurements were carried out by means of the CPMG sequence with the standard Bruker microprogram. The determination of T_2 with the CPMG sequence has a great advantage compared to the usual line-width method. In this case the inhomogenity of the applied magnetic field is excluded. This inhomogenity sometimes can cause 50% of the effect measured, however what is worst is that the effect changes from sample to sample because of the geometry and the quality of the NMR tubes. Temperature was always stable at 300±1K. After all NMR experiments the protein concentration was determined again.

Evaluation of NMR data

FIDs were analyzed by the Bruker 1D WinNMR program package. The heights and integrals of the peaks were fit to time as single exponentials in Microsoft Excel 5.0. The evaluation of relaxation parameters was also performed by non-linear parameter fitting using a homemade Gauss – Newton – Marquard algorithm. More exponential terms gave no significantly better fits. For hyperbolic fitting of the relaxation rate dependence on the total Ca^{2+} concentration to the theoretical equation, the program Sigmaplot Version 2001 was used with 100 iteration steps. The exact molecular masses of the enzyme molecules were retrieved from the Swissprot database.

TGase activity assay

TGase activity measurements were carried out by an amine-incorporation assay in an ELISA plate in which a biotynylated amine substrate was incorporated into dimethylated-casein coated onto a surface. The color reaction was developed by a streptavidin-conjugated alkaline-phosphatase.

Multiple sequence alignment

The multiple sequence alignment of TGases was created by the GCG program PILEUP and FASTA.

Homology modeling

In order to construct the homology model of TGase 2 using the refined high resolution structure of FXIII as a template, all known primary structures of TGases and related proteins were identified as a first step using the program package GCG. With the GCG program STRINGSEARCH and the search term 'transglutaminase', all TGase sequences from the GENEMBL database were retrieved and with the GCG program FASTA and the sequence of TGase 2 all homologous sequences were identified from the SWISSPROT database. Pairwise alignments were constructed using BESTFIT. Structure alignments were carried out using the UCLA Fold Prediction server. From the comparison of the alignments, highly conserved regions of the sequences were identified and distinguished from the less-conserved regions. Using the program "O", the side chains of the factor XIII structure in the conserved regions were replaced with the ones of the TGase 2 sequence. The less-conserved regions were deleted and rebuilt from a database, which contains sequences of known structures. Energy minimization was carried out using the program XPLOR. The first rounds were calculated without electrostatic potential and explicit hydrogen atoms in order to eliminate the worst close contacts between atoms. In later rounds electrostatic energy as well as polar hydrogen atoms were included. After each round, a superposition of the model onto the starting structure of factor XIII was calculated using the program LSQKAB of the CCP4 program suite. For structure validation, the program PROCHECK was used. All bond lengths and bond angles that deviated too much from their ideal values were checked manually and adjusted where it was possible. Potential hydrogen bonds and salt bridges were identified. The presumably conserved active center geometry was restrained during the final energy minimization rounds.

Surface polarity analysis

These calculations were made by the program Grasp using the recently determined 3D structure of FXIII-A zymogen and our homology model of TGase 2. Both molecules were probed by using the same intensity level of color (red=19.73, blue=14.37).

Preparation of inclusion bodies

After the optimization of the procedure, the plasmid expression was carried out in LB Broth Base containing 100 μg/ml ampicillin at 37 ^{o}C. Expression was initiated at OD=0.5 by adding 1mM IPTG for 105min. The culture was centrifuged (4^{o}C, 500g, 15min) and homogenized vigorously at 4^{o}C in 3ml of lysis buffer (50mM Tris, 1mM EDTA, 100mM KCl, 133μM PMSF, pH8.0) per wet gram of bacteria. Sonication was carried out with a Branson Sonifier 450 on ice (duty cycle: 40, output: 6, 3.5min) and the homogenate was centrifuged (12,000g, 4^{o}C, 15min). The pellet was resuspended at room temperature in 5ml 2M urea per wet gram of pellet for 5min. This was followed by a centrifugation (25^{o}C, 12,000g, 15min) and the pellet was kept as the inclusion body preparation.

Electronmicroscopic pictures

The electronmicrographs of the inclusion bodies were made according to Marston et al. Briefly, the samples were fixed by 2% glutaraldehyde and dyed by saturated uranyl acetate at 2^{o}C. The grids were coated with lysine. The samples were used without dehydration and dried on air. The transmission electronmicroscope was a JEOL 100B (1973, Japan).

RESULTS

1. We have constructed a new homology model of human tissue transglutaminase on the basis of the highest resolution (2.1Å) X-ray structure of cellular FXIII zymogen. Our model is the most accurate structure of the human TGase 2 now available.

2. We have performed surface polarity analysis on the highest resolution structure of cellular FXIII zymogen and on our homology model of TGase 2. Results were combined with the multiple sequence alignment of TGases and we could identify probably all the potential Ca^{2+} binding sites on these molecules; their numbers are more than reported earlier. From the previously published three putative sites of TGase 2, two possess negativity on the surface, but the third one does not. Interestingly, in the recently determined high affinity binding pocket of FXIII, only one amino acid shows considerable negative potential on the surface. The potential Ca^{2+} binding sites of all TGases can be predicted using the results of this analysis.

3. In accordance with the high number of negatively charged clusters in the surface polarity analysis, ^{43}Ca NMR provided higher (but still millimolar) average dissociation constants titrated on a wide Ca^{2+} concentration scale than previous studies did with equilibrium dialysis

in shorter ranges. These results suggest the existence of low affinity Ca^{2+} binding sites on both FXIII-A and TGase 2 in addition to the well-known high affinity ones.

4. Increasing the salt concentration or activating with thrombin, FXIII-A_2 partially lost its original Ca^{2+} affinity in contrast to previous results, which showed no effects.

5. The NMR data suggest different mechanisms for the proteolytic and salt activation processes of FXIII.

6. The NMR provided structural evidence for the GTP-induced conformational changes of TGase 2 molecule diminishing all of its Ca^{2+} binding sites.

7. The NMR data on the Ca^{2+} binding properties of the first analyzed TGase 3 are presented here; TGase 3 binds Ca^{2+} the most tightly, which weakens after the proteolytic activation.

8. The investigated TGases have very symmetric Ca^{2+} binding sites and no EF-hand motifs.

9. The refolding of the recombinant TGase 2 molecule to its catalytically active form from inclusion bodies essentially needs the presence of a helper material with higher molecular mass, but only in the initiation phase. In natural conditions, it is probably a chaperon; while *in vitro* it could be e.g. PEG.

10. Ca^{2+} and GTP are ascribed as effector molecules in the early phase of the structural reconstitution of TGase 2. The most efficient condition if they are together in the refolding buffer.

11. Two optimal concentrations of PEG (probably two different folding pathways) and a relatively long time scale (with two activity maxima) were identified for the evolution of the final structure of TGase 2 during its refolding.

12. An optimized refolding procedure of the non-fused recombinant TGase 2 is reported.

SUMMARY

- a new and precise homology model of human TGase 2 was constructed,
- Ca^{2+} affinities, Ca^{2+} binding site symmetry parameters and probably the localization of all the potential Ca^{2+} binding sites were determined for FXIII-A_2, TGase 2 and TGase 3,
- elucidated that the thrombin and salt activation of FXIII-A_2 are on different structural basis,
- found that TGases have very symmetric Ca^{2+} binding sites which are not EF-hands,
- a successful refolding procedure is developed providing structural and kinetic data for TGase 2,
- found two optimal concentrations of PEG for the initiation of the refolding of TGase 2,
- found that Ca^{2+} and GTP together affect the most positively the PEG-enhanced refolding of the human TGase 2.

LIST OF SCIENTIFIC PRESENTATIONS

Papers which these theses are based on:

Attila Ambrus and László Fésüs, Polyethylene glycol enhanced refolding of the recombinant human tissue transglutaminase. Prep. Biochem. &Biotechnol. 31(1) 59-70 (2001) IF: 0.447

Attila Ambrus, István Bányai, Manfred Weiss, Rolf Hilgenfeld, László Muszbek, Zsolt Keresztessy and László Fésüs, Calcium binding of transglutaminases: a ^{43}Ca NMR study combined with surface polarity analysis. J Biomol Struct Dyn (in press) IF: 1.643

ΣIF: 2.090

Posters:

Attila Ambrus, L. Fésüs (1999) Structural investigations of human tissue transglutaminase. 4th Work-meeting of the Molecular Biology Section of the Hungarian Biochemical Society, Eger, Hungary, Abstract No. DP-1, p.66.

Attila Ambrus, István Bányai, Manfred Weiss, Rolf Hilgenfeld, László Muszbek, Zsolt Keresztessy, László Fésüs (2000) Calcium binding of transglutaminases: a ^{43}Ca NMR study combined with surface polarity analysis. Sixth International Transglutaminase Conference, Lyon, France, Abstract No. 1.

Attila Ambrus, István Bányai, Manfred Weiss, Rolf Hilgenfeld, László Muszbek, Zsolt Keresztessy, László Fésüs (2000) Calcium binding of transglutaminases: a ^{43}Ca NMR study combined with surface polarity analysis. "The present students – the future scientists", The Day of Science, Budapest, Hungary, Abstract No.: 12.

Other Papers:

Ambrus Attila, Batta Gyula, Kövér E. Katalin, Examination of the structure of Calretinin: Backbone assignment and secondary structure elements of domain I-II. Biokémia (in press) IF: -.

Małgorzata Palczewska, Attila Ambrus, Patrick Groves, Gyula Batta, Katalin E. Kövér, Werner Klaus, Agata Kaleta and Jacek Kuźnicki, Calcium-dependent properties of neuronal calretinin modules I-II

(residues 1-100) differ from those of homologous calbindin D_{28k} modules I-II (residues 1-93). (manuscript in preparation)

Imre Tóth, Attila Ambrus and István Bányai, On the mechanism of cyanide substitution of $Tl(EDTA)CN^{2-}$ complex. (manuscript in preparation).

Attila Ambrus and István Bányai, Study of deuterium isotope effect in proton exchange reaction of HCN. (manuscript in preparation)

Gyula Batta, Attila Ambrus, Patrick Groves and Jacek Kuznicki, Backbone assignment and secondary structure elements of the calcium-dependent neuronal calretinin modules I-II (residues 1-100), (manuscript in preparation)

Other Posters:

Batta Gyula, Ambrus Attila, E. Kövér Katalin, Patrick Groves, Malgorzata Palczewska, Jacek Kuznicki (2000) NMR assignment and dynamic examination of Calretinin protein modules I-II. Conference of Chemists 2000, Debrecen, Hungary.

Batta, G., Ambrus, A., Kövér, K.E., Groves, P., Palczewska, M. and Kuznicki, J. (2000) The first (CR I-II) domain of Calretinin, a neuronal calcium binding protein. XIX. International Conference on Magnetic Resonance in Biological Systems, Florence, Italy, Abstract: p. 122 (top).

Batta Gyula, Ambrus Attila, E. Kövér Katalin, Patrick Groves, Malgorzata Palczewska, Jacek Kuznicki (2000) NMR assignment and dynamic examination of Calretinin protein modules I-II. "The present students – the future scientists", The Day of Science, Budapest, Hungary, Abstract No.: 13.