

Contributors

HIROYUKI KAGAMIYAMA	Osaka Medical College, Takatsuki
WŁODZIMIERZ S. OSTROWSKI	Jagellonian University, Kraków
IGNACY Z. SIEMION	Wrocław University
ANDRZEJ JERZMANOWSKI	Warsaw University
KENJI SODA	Kansai University, Suita
RENATA DĄBROWSKA	Institute of Experimental Biology, Warsaw
KUNIO YAGI	Institute of Applied Biochemistry, Gifu
JACEK OTLEWSKI	Wrocław University
ANTONI POLANOWSKI	Wrocław University
NOBUO KATO	Kyoto University, Kyoto
ANDRZEJ B. LEGOCKI	Institute of Bioorganic Chemistry, Poznań
FUMIO SAKIYAMA	Osaka University, Suita
CLARE E. SANSOM	London University, Institute of Bioorganic Chemistry
GRZEGORZ BUJACZ	Łódź University
YOSHIYUKI KAMIO	Tohoku University, Sendai
JERZY CIARKOWSKI	Gdańsk University
ANDRZEJ BIERZYŃSKI	Institute of Biochemistry and Biophysics, Warsaw
YASUSHI KAWATA	Tottori University, Tottori
MACIEJ ŻYLICZ	Gdańsk University
NIKODEM GRANKOWSKI	Lublin University
JACEK KUŹNICKI	Institute of Experimental Biology, Warsaw
JAN BARCISZEWSKI	Institute of Bioorganic Chemistry, Poznań

Polish Academy of Sciences
Japan Society for Promotion of Science

Polish-Japanese Seminar

TRENDS
in
PROTEIN RESEARCH

Poznań, 18-19 November, 1997

editors:
Andrzej B. Legocki & Kenji Soda

Scientific Publishers OWN
Poznań 1997

Polish-Japanese Seminar
TRENDS in PROTEIN RESEARCH

editors:

Andrzej B. Legocki & Kenji Soda

Poznań 18-19 November, 1997
Science Center, Polish Academy of Sciences
Wieniawskiego 17/19, 61-713 Poznań, Poland

Publication partially supported by
the State Committee for Scientific Research

Copyright © Institute of Bioorganic Chemistry
Poznań 1997
All rights reserved

ISBN 83 85481-59-1

Published by:

Scientific Publishers OWN
Polish Academy of Sciences
Wieniawskiego 17/19, 61-713 Poznań
tel.: +4861 8528503 fax: +4861 8520671
e-mail: own@man.poznan.pl

Technical Editor: Andrzej Wójtowicz

Cover contains etching *Summer with a girl* (1982) by Leszek Rózga

Printed in Poland (TotalDruk, Poznań)

Contents

PROTEIN FUNCTION

Preface	7
HIROYUKI KAGAMIYAMA, HIDEYUKI HAYASHI The role of the pyridoxal phosphate-lysine Schiff base in the catalytic action of aromatic-L-amino acid decarboxylase	11
WŁODZIMIERZ S. OSTROWSKI Phosphorylo-enzyme intermediate in catalytic reaction of non-specific phosphomonohydrolases	15
IGNACY Z. SIEMION Thymopoinetin-like motifs in proteins	21
MARTA PRYMAKOWSKA-BOSAK, RADOSŁAW TOMASZEWSKI, MARCIN PRZEWŁOKA, JOANNA ŚLUSARCZYK, JOANNA DŁUŻNIEWSKA, NICOLE CHAUBET, STEVEN SPIKER, MIECZYŚLAW KURAŚ, ANDRZEJ JERZMANOWSKI Histone H1 – a well known protein with the unknown function	27
KENJI SODA, TATSUO KURIHARA, NOBUYOSHI ESAKI Structure and reaction mechanism of 2-haloacid dehalogenases	31
RENATA DĄBROWSKA, ROBERT MAKUCH Structural and functional aspects of smooth muscle regulation by actin-associated proteins	37

BIOTECHNOLOGICAL IMPLICATION OF PROTEIN RESEARCH

KUNIO YAGI Expression of phospholipid hydroperoxide glutathione peroxidase	45
JACEK OTLEWSKI, WŁODZIMIERZ APOSTOLUK, OLGA BUCZEK, LILIANA CHOLAWSKA, AGNIESZKA GRZESIAK, KATARZYNA KOŚCIELSKA, IZABELA KROKOSZYŃSKA, DANIEL KROWARSCH, DAMIAN STACHOWIAK, MICHAŁ DADLEZ Structural and energetic aspects of protein inhibitor – serine proteinase recognition	49
TADEUSZ WILUSZ, ANTONI POLANOWSKI Designing of protein proteinase inhibitors in the aspect of both protection and nutritional improvement of plants	59
NOBUO KATO, YASUYOSHI SAKAI Heterologous gene expression system developed in the methylotrophic yeast, <i>Candida boidinii</i> , for useful enzyme production	65
MICHAŁ M. SIKORSKI, ANDRZEJ B. LEGOCKI Pathogenesis-related protein genes structure and expression	71

PROTEIN CONFORMATION AND MODELLING

FUMIO SAKIYAMA, SHIGEMI NORIOKA, SHIGEKI OHTA, SEONG-IL LIM, KENTARO SHIRAKI, NARUTO KATSURAGI, TAKEHARU MASAKI, YUKITERU KATSUBE, YASUYUKI KITAGAWA, KAZUMICHI SASHI Structural basis for restricted substrate specificity and high catalytic potency of a lysylendopeptidase, <i>Achromobacter</i> protease I	79
CLARE E. SANSOM, MARIUSZ JASKÓLSKI Modelling enzyme action	85
GRZEGORZ BUJACZ, JERRY ALEXANDRATOS, ALEXANDER WLODAWER, GEORGE MERKEL, MARK ANDRAKE, RICHARD KATZ, ANNA MARIE SKALKA Binding of different divalent cations to the active site of ASV integrase	89
HIROFUMI NARIYA, AKIHITO NISHIYAMA, YOSHIYUKI KAMIO Identification of the minimum segment in which Threonine ²⁴⁶ residue is phosphorylated by protein kinase A for the LukS-specific function of staphylococcal leukocidin	95
JERZY CIARKOWSKI, CEZARY CZAPLEWSKI, RAJMUND KAŻMIERKIEWICZ, EWA POLITOWSKA Molecular modelling of G protein-coupled receptor (GPCR) – bioligand interactions	101
ANDRZEJ BIERZYŃSKI α -Helix formation as an early step of protein folding: a critical examination of the helix-coil transition theories	107

REGULATORY FUNCTION OF PROTEINS

YASUSHI KAWATA Structure and function of molecular chaperone GroE: the chaperonin-facilitated protein folding and the role of nucleotide	113
ALICJA WAWRZYNÓW, BOGDAN BANECKI, KRZYSZTOF LIBEREK, ADAM BŁASZCZAK, JAROSŁAW MARSZAŁEK, IGOR KONIECZNY, MACIEJ ŻYLICZ Hsp70 molecular chaperone machinery: structure and function	119
MAREK PILECKI, PIOTR DUKOWSKI, NIKODEM GRANKOWSKI Acidic Ribosomal Proteins and their phosphorylation in yeast cells	125
WIESŁAWA LEŚNIAK, GERTRUDA POKRYWIŃSKA, JACEK KUŹNICKI Use of the matinspector program and gel mobility shift assays for analysis of regulatory sequences in the S100A6 (Calcyclin) gene promoter	131
MAŁGORZATA GIEL-PIETRASZUK, MACIEJ SZYMAŃSKI, GENOWEFA ŚLÓSAREK, THOMAS SPECHT, VOLKER A. ERDMANN, PIOTR MUCHA, PIOTR REKOWSKI, GOTFRYD KUPRYSZEWSKI, JAN BARCISZEWSKI The structure and function of systemin	141
Index of authors	149

Preface

The history of co-operation between the Polish Academy of Sciences and the Japanese Society for Promotion of Science is relatively short, but fruitful. Over the last three years, these contacts resulted in three bilateral scientific conferences. In 1995, a Symposium on Medicinal Applications of Molecular Biology was held in Łódź. In the following year, Plant Biotechnology conference was held in Osaka, and in 1997, a symposium on Trends in Protein Research was held in Poznań.

Thus, the subject of the Polish-Japanese symposium included the priority areas of biotechnology showing the fields of potential co-operation between the two countries.

The last decade has seen rapid development of protein research. New methodologies allowed to combine research on proteins at the DNA level, including regulation of translation process. Progress in research on protein conformation and relating it to biological functions offer a chance to understand one of the key issues of molecular biology, i.e. the correlation between the structure of macromolecules and their function.

On 18-19 November 1997, the Polish-Japanese symposium organised by the Institute of Bioorganic Chemistry of the Polish Academy of Sciences was held in Poznań. This book contains the high-lights of the 22 lectures presented at the symposium.

Presenting the readers with this publication we strongly believe that this work should well contribute to the Polish-Japanese co-operation. The future will verify whether our expectations in that respect were justified.

Andrzej B. Legocki
Kenji Soda

Poznań, November 1997

Japan, November 1971

1971

PROTEIN FUNCTION

THE ROLE OF THE PYRIDOXAL PHOSPHATE-LYSINE SCHIFF BASE IN THE CATALYTIC ACTION OF AROMATIC -L-AMINO ACID DECARBOXYLASE

Hiroyuki Kagamiyama and Hideyuki Hayashi

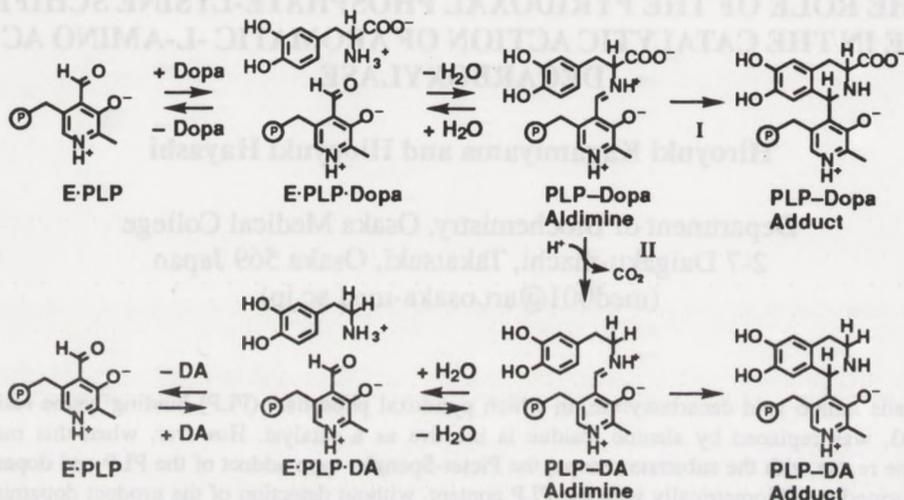
Department of Biochemistry, Osaka Medical College
2-7 Daigaku-machi, Takatsuki, Osaka 569 Japan
(med001@art.osaka-med.ac.jp)

Aromatic amino acid decarboxylase, in which pyridoxal phosphate (PLP) binding lysine residue, Lys303, was replaced by alanine residue is inactive as a catalyst. However, when this mutant enzyme reacts with the substrate L-dopa, the Pictet-Spengler type adduct of the PLP and dopamine was formed stoichiometrically with the PLP content, without detection of the product dopamine in the reaction mixture and the adduct of the PLP and dopa. By using dopa methyl ester, we found that the corresponding adduct was formed when the decarboxylation was blocked by esterifying the carboxylate group of dopa, suggesting that the dopa-PLP adduct might be formed at the rate much smaller than that of the decarboxylation. Analyses of the anomalous side reaction of Lys303Ala mutant AADC indicates that Lys303 is not essential for the decarboxylation step, but is important for increasing the rate of the product release and probably the external aldimine formation.

In higher animals, decarboxylation of amino acid provides biologically active amines. Aromatic amino acid decarboxylase (AADC) catalyzes the irreversible decarboxylation reaction of L-dopa and 5-hydroxytryptophan, producing dopamine (DA) and serotonin, respectively.

AADC is an enzyme that utilizes pyridoxal phosphate (PLP) as a coenzyme. Like other PLP-dependent enzymes, the carbonyl group of the PLP is bound to enzymes through aldimine formation with the ϵ -amino group of a specific lysine residue at the active site (internal aldimine). When the amino acid substrate enters the active site, it displaces the lysine amino group by a transaldimination process to form a substrate-PLP Schiff base (external aldimine). Concomitant with the formation of the external aldimine with the substrate, the free lysine amino group is released. There is evidence that in most of the PLP-dependent enzyme reaction, α -proton of substrate amino acid must be released as the prerequisite step to the subsequent processes, in which the released ϵ -amino group of the lysine residue participates as the base catalyst for the α -proton removal (1, 2). However, in the decarboxylase reaction, the α -proton is not required to be removed. Therefore, the lysine residue that forms an internal aldimine in amino acid decarboxylase is considered to serve different roles.

This paper describes the functional roles of the lysine residue that binds PLP in the active site of amino acid decarboxylase.



We expressed the gene of rat liver AADC in *E.coli* to study AADC at the protein level (3).

AADC is a homodimer and the subunit is composed of 480 amino acids. Lysine 303 is the residue that forms the internal aldimine. To define its functional role, Lys303 was mutated to alanine. The mutant AADC was catalytically inactive.

Spectral Properties of AADC

The wild type enzyme has a large absorption bands at 335 nm and 425 nm (3). The 335 nm band is ascribed to the enolimine form and 425 nm band to the ketoenamine form of the internal aldimine between the PLP and Lys303. The mutant enzyme shows the absorption bands at 330 nm and 397 nm (4). The 397-nm absorption can be ascribed to the PLP molecule in the active site of the mutant enzyme without forming the internal aldimine, because free PLP has an absorption maximum at 390 nm. However, the structure of the 330-nm absorption species is less clear. Fluorescence spectroscopic analyses of the mutant AADC indicated that the one possible structure was hydrated form of the PLP. These results strongly suggested that the PLP did not form an internal aldimine in the active site of the mutant AADC.

Reaction Product of the Mutant AADC

The content of PLP in the mutant enzyme was estimated to be 0.8 mol/subunit by the well-known phenylhydrazine method.

The mutant AADC was incubated with L-dopa. Although the enzymatic activity to produce dopamine was not detected, the absorption at 330 nm gradually increased with the concomitant decrease in the absorption at 390 nm. To identify this 330 nm species, the reaction mixture was analyzed on HPLC column (Cosmosil 5C18, 4.6 × 250 mm, 50 mM K-phosphate buffer, pH 2.2, containing 50 mM perchlorate) after the

spectral change was completed. The effluent was monitored at 330 nm. A single peak was observed at the retention time of 9.5 min. The absorption spectrum of the compound in this peak, showing the absorption maxima at 325 nm and 280 nm, was in good agreement with that of the sum of the spectra of pyridoxamine phosphate and dopamine. This suggested that the compound has a catechol ring and a hydroxypyridine ring and has sp^3 structure at C4' position of the coenzyme.

Dopa and dopamine are easily condensed with aldehydes (Pictet-Spengler reaction, 5). When the reaction mixture, in which dopa and dopamine were incubated with PLP, was subjected to the HPLC, the peaks with absorption at 330 nm were observed at the retention time of 9.5 min for the dopamine-PLP adduct and 11.5 min for the dopa-PLP adduct. Both peaks showed a spectral pattern identical to that of the enzyme product. Interestingly, on the chromatogram of the reaction mixture of the mutant enzyme and L-dopa, only a peak corresponding to the dopamine-PLP adduct was detected. Incubation of 10 μM of the mutant AADC with 200 μM L-dopa produced 7.5 μM of the dopamine-PLP adduct, showing that almost all of the PLP in the active site of the mutant AADC was converted to the dopamine adduct. When the same amount of the wild type enzyme was incubated with 1 mM dopa, the dopamine-PLP adduct, as well as the dopa-PLP adduct, was below the detection limit, indicating that such a side reaction does not occur in the catalytic reaction of the wild type enzyme.

Dopa methyl ester (DopaOMe) is an analogue of dopa. It can form an external aldimine with the PLP, but is unable to undergo decarboxylation, because of the presence of a methyl group attached to the carboxylate group. When the mutant AADC was reacted with DopaOMe, the spectrum of the enzyme changed in the same manner as that observed for L-dopa or dopamine, and the Pictet-Spengler adduct was obtained. This indicated that, although the dopa-PLP adduct was not detected in the reaction of the mutant AADC with L-dopa, this pathway could be detected when the decarboxylation reaction was blocked by methylating the carboxylate group of L-dopa. Therefore, the dopa-PLP adduct might be formed from dopa-PLP external aldimine with much smaller rate compared with that of decarboxylation of the external aldimine.

Kinetic Consideration

The dopamine-PLP adduct accumulated up to 7.5 μM , as described above, when 10 μM of the mutant AADC was reacted with 200 μM of dopa. However, accumulation of the dopa-PLP adduct was lower than the detection limit of 0.05 μM . Therefore, the rate of step II (Figure) is at least 150 times as fast as that of step I (Figure). The maximal value of the rate constant for the spectral change on the reaction of the mutant AADC with DopaOMe was 0.068 s^{-1} . This must be a lower limit of the rate constant of step I for DopaOMe. This value can be considered to be the rate constant for dopa-PLP adduct formation. Therefore, the rate constant for step II (decarboxylation step) can be estimated to be more than 10.2 s^{-1} . The rate constant for decarboxylation step for the wild type enzyme had been obtained to be 11.2 s^{-1} (3). Therefore,

we can conclude that Lys303 is not essential for the decarboxylation in AADC reaction.

Role of Lys303

Since the PLP-dopamine adduct is formed from PLP-dopamine aldimine without detection of dopamine in the reaction mixture, hydrolysis of the PLP-dopamine aldimine does not occur in K303A AADC. Therefore, we can conclude that in the wild type AADC, the ϵ -amino group of Lys303 displaces the amino group of dopamine by nucleophilic attack on the imine bond of the PLP-dopamine complex. This transimination could promote the rate of cleavage of the imine bond of the PLP-dopamine aldimine and also prevent the unfavorable equilibrium shift toward the formation of the PLP-dopamine, thus accelerating the product release.

The small rate constant for the aldimine formation between DopaOMe and PLP (0.068 s^{-1}) indicates that Lys303 may also participate in accelerating the aldimine formation between PLP and α -amino group of the substrate or product, because aldimine is more easily attacked by nucleophiles than free aldehyde of PLP. Actually, in aspartate aminotransferase, representative of the PLP enzymes, the internal aldimine bond had been shown to permit 10^4 - 10^5 fold faster external aldimine formation. Furthermore, in aminotransferase and most of the PLP enzymes, the lysine residue which forms the internal aldimine has an additional role in activating α -proton, but in decarboxylase this lysine residue is not essential for the activation of α -carboxylate.

In this study, the rate constants of all the steps were not determined, because the dopa aldimine and the dopamine aldimine were not quantitated spectrophotometrically. This is most probably because these species are accumulated in very little amount during the reaction of the mutant enzyme. To discuss the roles of Lys303 more quantitatively, the complete determination of all the kinetic parameters of the wild type and the mutant enzymes must be obtained. Furthermore, the mechanism of reprotonation at the α -carbon atom after the removal of α -carboxylate group (step II) must be elucidated.

References

1. Metzler, D.E. (1977) *Biochemistry The Chemical Reactions of Living Cells*, Academic Press, New York
2. Toney, M.D., and Kirsch, J.F. (1993) *Biochemistry* 32, 1471-1479
3. Hayashi, H., Mizuguchi, H., and Kagamiyama, H. (1993) *Biochemistry* 32, 812-818
4. Nishino, J., Hayashi, H., Ishii, S., and Kagamiyama, H. (1997) *J. Biochem.* 121, 604-611
5. Abbott, E.H., and Martell, A.E. (1970) *J. Am. Chem. Soc.* 92, 1754-1759

PHOSPHORYLO-ENZYME INTERMEDIATE IN CATALYTIC REACTION OF NON-SPECIFIC PHOSPHOMONOESTERASES

WŁODZIMIERZ S. OSTROWSKI

*Institute of Medical Biochemistry, Collegium Medicum of Jagiellonian University,
Kopernika 7, 31-034 Kraków, Poland*

Hydrolysis of phosphate monoesters by acid and alkaline phosphatases is accomplished by a diverse group of enzymes. As the model enzyme to study the reaction catalyzed in acidic pH, human prostatic acid phosphatase (hPAP) was used. Crystals of hPAP to 12.7 and to 3.0 Å resolution was used for the structure determination and modeling of the active center of the enzyme. Catalytic properties of the enzyme and comparison with other phosphomonoesterases will be discussed.

Introduction

Hydrolytic reactions of phosphate esters in biological systems is an important process combined with energy production, metabolic regulations, transport and with wide variety of cellular signal transduction pathways which are dependent on the activity of phosphomonohydrolases, the enzymes we generally call acid and alkaline phosphatases. This enzymes can split C-O-P and N-P bonds of simple esters, phosphoproteins and terminal phosphate of nucleotides. Phosphatases are recently classified into five groups [1]: a) alkaline phosphatases, b) purple acid phosphatases, c) low MW acid phosphatases, d) high MW acid phosphatases, and e) protein phosphatases which are specific for phospho-aminoacids of proteins. High MW acid phosphatases show hydrolytic and transphosphorylating activity [2]. As the model enzyme for our study we selected acid phosphatase from human prostate gland (hPAP) which was the first tumor marker identified in blood serum about 60 years ago [3], and can be easily obtained in pure state from seminal plasma by affinity chromatography [4].

Properties and structure of hPAP

The enzyme catalyzes hydrolysis of phosphomonoesters using a broad spectrum of substrates including phosphoproteins with phosphorylated serine, threonine, tyrosine and histidine residues [5]. The biosynthesis of hPAP is coded by separate gene located in chromosome 3 in the region 3q21-23 [6] and the process is androgenic hormones dependent [7]. As the secretable enzyme, the gene encodes a 386 residue protein and 32-amino acid signal peptide at N-terminal and which is proteolytically split off during transmembrane transport [8]. The enzyme is a glycoprotein of MW 100 kDa comprised of two identical subunits. There are three asparagine-linked carbohydrate chains attached to the phosphatase polypeptide composed mostly of mannose, fucose and sialic acid. Their sequence has been established by ¹H NMR spectrometry [9].

We have obtained several crystal forms of hPAP using the vapor diffusion technique in hanging drop set-up composed of ammonium sulfate solution and polyethylene glycol

4K over pH range 4.5-7.5 [10]. Changing the pH and ratio of phases finally gave us diamond shaped crystals grown reproducibly from several protein preparations. The crystals diffracted X-rays to 2.7 Å and are belonging to the orthorhombic space groups P212121 measured 0.5 × 0.5 × 0.3 mm showing unit cell dimensions: $a = 126.3 \text{ \AA}$, $b = 207.9 \text{ \AA}$ and $c = 73.0 \text{ \AA}$ (Fig. 1). There are two dimers per asymmetric portion of the unit cell [11].

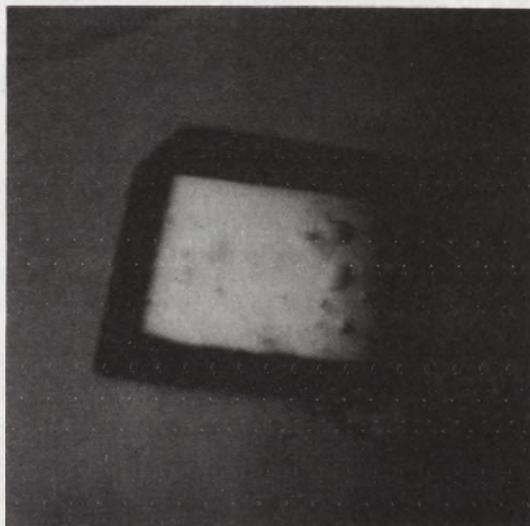


Fig. 1. Photography of monoclinic hPAP crystals grown under biphasic conditions.

Data for the above crystals were measured to 2.7 Å resolution using a R-AXIS II imaging plate system. Subsequent refinement was performed using several programs to refine overall structure of the enzyme. The subunit of hPAP is built up of two domains containing 12 α -helical and 7 β -sheet segments (Fig. 2). Dimer is formed through interactions between α -helices, β -strands and loops. The active site is located at the end of the β -sheet domain in the donut shaped cleft formed by the two domains of the dimeric molecule. The cleft is open and accessible from either side of the dimer.

Catalytic activity

The prostatic enzyme catalyzes two general reactions: hydrolysis of O- and N-linked phosphomonoesters and transphosphorylation with optimum catalytic activity at pH4.5-5.0. The enzyme is hydrolyzing a wide range of alkyl, aryl and acylorthophosphate monoesters. Most of non-specific phosphomonoesterases are acting through the formation of a phosphoryl-enzyme intermediates with elimination of an alcoholic group and in the second step is the hydrolysis of the phosphoryl-enzyme. From the presteady-state burst analysis using p-nitrophenyl phosphate as the substrate, two active centers per dimer of enzyme were observed [12]. Therefore, it was important to study: is association of both identical promoters of hPAP necessary for catalytic function?

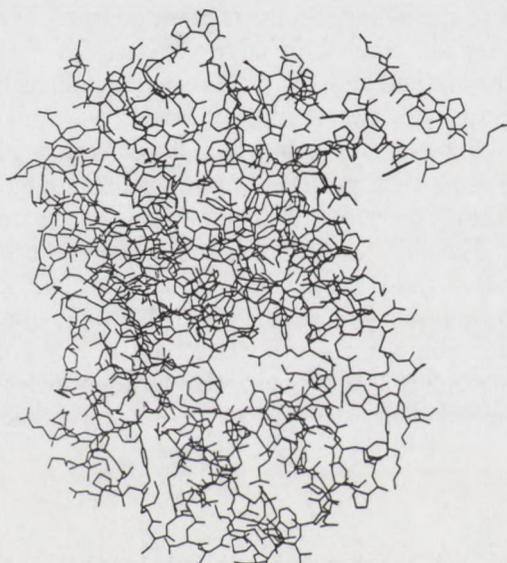
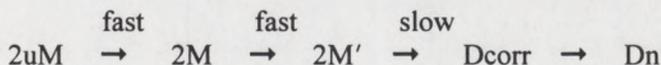


Fig. 2. Three dimensional structure of one subunit of hPAP.

One of the experimental approach to answer this question is kinetic analysis of recovery of enzyme activity after preceding denaturation. We demonstrated, that the enzyme can be reversibly denatured in 6 M urea at pH 2.5 [13]. In this condition the enzyme dissociates into subunits and loses its catalytic activity. But the process is reversible and renaturation leads to reactivation of enzymatic properties and reassociation of the dimer (Fig. 3). Following spectral changes (CD, fluorescence) and fast size-exclusion chromatography during renaturation we were able to propose the folding pathway and assembly of reversible denatured subunits of the enzyme:



where: uM – unfolded monomer, M and M' – partially folded monomers, D_{corr} – dimeric inactive intermediate which form native dimer D_n. The fast steps of renaturation, M and M', are resulting of rapid reconstitution of α-helical fragments, in the next steps interactions between β-sheets, helices and loops form subunit interfaces and fix simultaneously the dimeric structure (D_{corr}, D_n) including spatial arrangements of the active site.

For obtaining more data on active center of hPAP, results from X-ray studies and the homology modeling using described structure of recombinant rat prostatic phosphatase [14] was used. The active site of hPAP consists of the conserved residues, Arg11, His12, Arg15, His257 and Asp258 located in large cavity (Fig. 4). Extensive studies of chemically modified enzyme and series of site directed mutations allowed Van Etten and coworkers [15] to propose a description of the catalytic reaction. Arginine residues are protonated and charged at the pH range in which the enzyme is active. His12, which is the nucleophile in the catalysis is protected from protonation by the electrostatic field produced by the arginines. To verify this hypothesis we calculated the molecular electrostatic potential (Fig. 5) which shows a deep minimum in the active site contributing to the binding of anionic ligands [16]. After formation of enzyme-substrate complex, break-down of the phospho-enzyme occurs by nucleophilic attack by water upon the phospho-enzyme intermediate, then orthophosphate is released or phosphorylation occurs of phosphate acceptor. The strong nucleophilic property of hPAP explains its broad specificity to the molecular structure of the R group of the substrate.

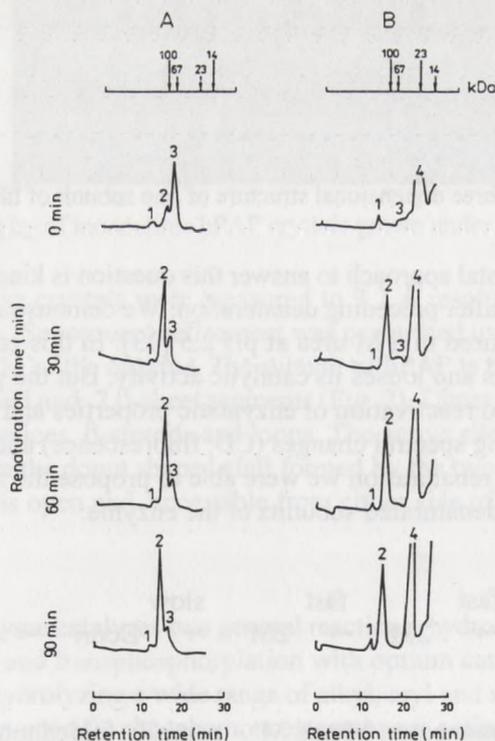


Fig. 3. Fast size-exclusion chromatography of renaturing hPAP. Renaturation by tenfold dilution in 0.1 M phosphate buffer, pH 7.2 containing 0.3 M NaCl: A – untreated with trypsin, B – treated with trypsin (phosphatase is proteolytically split if it is denaturated, 13). Peak 1 – aggregates (12.6 min.), peak 2 – dimer (15.3 min.), peak 3 – monomer (17.0 min.), peak 4 – trypsin (21.4 min.).



Fig. 4. Electron density of the active site of hPAP. The cage marked in the center is the electron density for chloride ions which interact with His12 and Asp258.



Fig. 5. Molecular electrostatic potential calculated for hPAP molecule using the DelPhi software.

The most members of the groups of phosphomonoesterases, hydrolyze phosphate esters, similarly to prostatic acid phosphatase, via the associative mechanisms [1]. The purple acid phosphatases from mammals are distinguished by their purple-pink colors which is due to the presence of a binuclear iron center [17]. Hydrolysis of the phosphate esters proceeds through the two-step mechanism utilizing a base-stable, acid-labile phosphoenzyme intermediate. The reduced form of the enzyme (pink) contains a mixed-valent $\text{Fe}^{3+}\text{-Fe}^{2+}$ center and is enzymatically active; oxidized species containing $\text{Fe}^{3+}\text{-Fe}^{3+}$ center (purple) exhibit no activity. Alkaline phosphatases are usually dimers with MW about 100 kDa and hydrolyze variety of phosphate esters at pH optimum about 8.0. They are distinguished also by the fact that utilize Zn or other bivalent metal ions in catalytic reactions [18]. Alkaline phosphatases form base-labile, acid-stable phosphoenzyme intermediates by the phosphorylation of serine residue as the nucleophile in the active site.

References

- [1] Vincent J. B., Crowder M. W., Averill B. A., *TIBS*, **17**, 105-110 (1992).
- [2] Ostrowski W. S., *Human Prostatic Acid Phosphatase: Physico-chemical and Catalytic Properties*, in: Male Accessory Sex Glands (E. Spring-Mills, E. S. E. Hafez ed.), Elsevier North-Holland Biomed. Press, 197-213 (1980).
- [3] Kutscher W., Wolberg H., *Z. Physiol. Chem.*, **236**, 237 (1935).
- [4] Van Etten R. L., Saini M. S., *Clin. Chem.*, **24**, 1525-1530 (1978).
- [5] Lin M. F., Clinton G. M., *Adv. Prot. Phosphatases*, **4**, 199-228 (1987).
- [6] Li S. S.-L., Sharief F. S., *Genomics*, **17**, 765-766 (1993).
- [7] Lin M. F., Garcia-Arenas R., Chao Y.-Ch., Lai M. M. C., Patel P. C., Xia X.-Z., *Arch. Biochem. Biophys.*, **300**, 384-390 (1993).
- [8] Van Etten R. L., Davidson R., Stevis P. E., MacArthur H., Moore D. L., *J. Biol. Chem.*, **266**, 2313-2319 (1991).
- [9] Riskey J. M., Van Etten R. L., *Arch. Biochem. Biophys.*, **258**, 404-412 (1987).
- [10] Kuciel R., Jakob C. G., Lebioda L., Ostrowski W. S., *J. Crystal Growth*, **122**, 199-203 (1992).
- [11] Jakob C. G., Kuciel R., Lewinski K., Ostrowski W. S., Lebioda L., *Abstr. Amer. Crystallogr. Assn.*, **23**, 116 (1995).
- [12] Ostrowski W., Barnard E. A., *Biochemistry*, **12**, 3893-3898 (1973).
- [13] Kuciel R., Mazurkiewicz A., Ostrowski W. S., *Int. J. Biol. Macromol.*, **18**, 167-175 (1996).
- [14] Lindquist Y., Schneider G., Vihko P., *J. Biol. Chem.*, **268**, 20744-20746 (1993).
- [15] Van Etten R. L., Ostanin K., Saeed A., *J. Biol. Chem.*, **269**, 8971-8978 (1994).
- [16] Lovelace L., Lewinski K., Jakob C. G., Kuciel R., Ostrowski W. S., Lebioda L., *Acta Biochim. Polon.*, 1997 (in press).
- [17] Vincent J. B., Averill B. A., *FASEB J.*, **4**, 3009-3014 (1991).
- [18] Kim. E. E., Wyckoff H. W., *J. Biol. Chem.*, **218**, 449-464 (1991).

THYMOPOIETIN-LIKE MOTIFS OF PROTEINS

Ignacy Z. Siemion

Faculty of Chemistry, Wrocław University, 50-383 Wrocław, Poland

Abstract: The data collected suggest that the thymopietin-like motifs, which appears in the numerous proteins, originated of the nucleic acid binding domains of archeoproteins. The attention is given to the fact that retro-thymopietin motifs appear within the proteins practically with the same frequency as thymopietin-like motifs. The hypothesis is formulated that both these motifs were accommodated for the interaction with complementary polynucleotide chains of DNA double helix. The appearance of discontinuous thymopietin-like motifs in such protein as actin G, heat shock protein HSC 70, and human leukocyte antigen class II HLA-DQ, is also discussed.

Some years ago we started the search for new peptide immunomodulators, looking for thymopentins-like sequences within the sequences of known regulatory and defence proteins. Thymopentins (TP5) is an active fragment of thymopietin, the polypeptide immunoregulator produced in thymi:

Thymopentins

RKD^{**V**}Y

Thymopietin (human)

GLPKEVPAVLTKQKLKSELVANGVTL**PAGEMRKDVY**VELYLQHLTALH

This approach resulted in a huge variety of immunosuppressive and immunostimulative peptides. We derived their sequences from the sequences of such proteins as human lactoferrin, the proteins of transforming growth factor β (TGF β) family, human leukocyte antigen class II (HLA-DQ), the proteins of interleukin-1 family, human p53 protein, and FKBP - the protein which binds immunosuppressive macrolactam found by Japanese scientists, the substance known as FK-506.

E. g. in HLA-DQ the thymopentins-like fragment RGD^{**V**}Y occupies the positions 167-171. It contains the RGD sequence, which is a centre of adhesion of proteins. We found that some peptides related to this region of HLA-DQ suppress the cellular immune response and show antiadhesive potency [1].

In the case of p53 and FKBP there appear a thymopietin homology:

```
1  VVPYEPPEVGSDCTTIHYNMCNSSCMGGMNRRPILTITLEDSSGNL,  
2  GLPKEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH  
3  GRTFPKRGQTCVVHYTGMLEDGKKFDSSRDRNKPFKFVLGKQEVIRGW  
1-p53 DNA-binding domain; 2-thymopietin; 3-FKBP
```

(Here and in other sequences compared the same and similar amino acid residues are indicated by the bold letters).

A fact that the mentioned (245-249) sequence of p53 belongs to DNA-binding domain of this protein prompted us to examine the thymopietin homology in other nucleic acid binding proteins. A vast number of such homologies was described by us in a

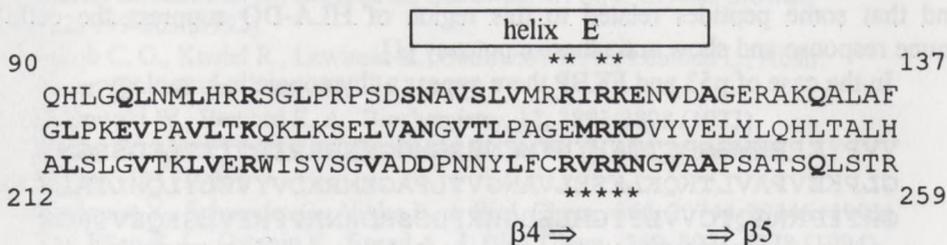
review-article [2]. Such a homology was found. *inter alia*, in histones and archebacterial histone-like proteins:

- 1 GLPKEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH
- 2 SDDARITLAKILEEMGRDIASEAIKLARHAGRRTIKAEDIELAVRRFK
- 3 SDDARIALAKVLEEMGEEIASEAVKLAKHAGRRTIKATDIELARKMFK
- 4 SDDAKETLAKALEEMGEEISRKAVELAKHAGRRTVKATDIEMAAKQL
- 5 SSKAMSIMNSFVNDIFERIAAEASRLAHYNKRSTITSREIQTAVRLLL
- 6 SSKAMGIMNSFVNDIFERIAEASRLAHYNKRSTITSREIQTAVRLLL
- 7 FQSSAVMALQEASEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIR
- 8 TSGLIYEETRGVLKVFLENVIRDAVTEHAKRKTVTAMDVVYALKRQGR

1-human thymopoietin; archebacterial histone-like protein: 2-hmf B, 3-hmf I, 4-hmf A; 5-dhH2B *Drosophila melanogaster*; 6-hH2B human; 7-hH3; 8-hH4.

(The data collected by Hoffmann et al. [3] are used in this comparison.). The homology between the thymopoietin and archebacterial HmfB protein reaches a value of 30%. It diminishes, however, when the sequences of histones are taken into account.

The thymopoietin homologies were found also in TATA binding subunit of TFIID transcription factor, in replication-terminator Tus, Arc-repressor, in the proteins of the family of NFAT transcription factors, in HTH (helix-turn-helix) motifs of many phage transcription factors, in eubacterial homeo-domains, in bacterial recombinases, and in molecules of single-stranded DNA-binding proteins. They appear also in such ribosomal RNA-interacting proteins as L11 protein of *B. stearothermophilus*, S15 protein of chloroplasts of *Zea Mays*, in RNA-binding domain of human nuclear hnRNP A1 protein etc. We found now two new examples of such homology in bacteriophage Cre recombinase and breakage-reunion domain of *E. coli* gyrase. Cre recombinase catalyses the reciprocal exchange of DNA strands. Two thymopoietin-like fragments of this protein comprise the residues 90-137, and 212-259, respectively:



The partial sequences of recombinase, which correspond to thymopentin fragment of thymopoietin, appear here in very different stereochemical environment. The first thymopoietin-like motif is situated in a part of helix E, and the second - in a part of the loop between two β-strands [4]. However, both these fragments belong to the DNA-interacting site of the protein. (The residues involved in the interaction with DNA are indicated on the scheme by asterisks).

Two thymopietin-like fragments appear also in the molecule of bacterial gyrase A. Bacterial gyrase demonstrates the homology with yeast topoisomerase. Both these enzymes regulate the formation of DNA superhelices. The thymopietin-like regions of thymopietin-like fragments of Gyr A occupy the helical segments of the protein [5]:

15	Gyr A	62
	EELKSSYLDYAMSVIVGRALPDVRDGLKPVHRRVLYAMNVLGNDWNKA	
206	Gyr A	253
	EGLMEHIPGPDFPTAAIINGRRGIEEAYRTGRGKVYIRARAEEVEVDAK	
	Thymopietin	
	GLPKEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH	
674	Topo II (yeast)	721
	SDFINLELILFSLADNIRSI PNVLDFKPGQRKVLVYGCFFKKNLKSELK	

The repeat of thymopietin-like motif can be also seen in bacterial GroEL protein which forms the chaperonin complex. These motifs appear in N-terminal, as well as in „equatorial” domain of GroEL. Within the „equatorial” domain such a motif forms part of a nucleotide-binding pocket [5]:

254	N-terminal domain	301
	VEGEALATLVVNTMRGIVKVA AVKAPGFGDRRKAMLQDIATLTGGTVI	
	Thymopietin	
	GLPKEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH	
398	„equatorial” domain	445
	DALHATRAAVEEGVGGVALIRVASKLADLRGQNEQNVGIKVALRA	

The repeats of the thymopietin-like motifs was found by us also in ATP-pyrophosphatase, adenyl cyclase, SH2 domain of tyrosine kinase and many other proteins. All these facts argument strongly for the hypothesis that thymopietin-like motif originated of very old precursor gene, which was in some cases multiplied during the evolution. This gene coded probably the early nucleic acid binding protein.

The thymopietin homology appears also in a vast number of signalling and regulatory proteins. We mentioned above that thymopietin-like motif appears in the molecule of p53 protein, involved in apoptosis. The same situation appears also in the case of the „death domain” of Fas protein:

GLPKEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH
TVAINLSDVDLSKYITTIAGVMTLSQVKG FVRKNGVNEAKIDEIKNDN

The thymopietin homology is visible also in RAIDD adaptor protein which mediates the coupling of Cys-proteinases to the death-signalling pathway; in tumour necrosis receptor factor protein, in ε-subunit of GABA_A receptor, in the proteins Ras/Rap subfamily, etc..

Unexpectedly we have found that in sequences of many proteins there appears a retro-thymopoietin homology. Moreover, frequencies of appearing of thymopoietin or retro-thymopoietin homology in proteins are almost on the same level. A vast number of found examples of such homology was indicated in our review paper [2]. Here I would like to quote only two of them. The first one represents ubiquitin:

N-terminus		C-terminus
	LTKQKLKSELVANGVTLPAGEMRKDVY	thymopoietin
	GGRLRLVLHLTSEKQINYDSLTRGDEL	ubiquitin
C-terminus		N-terminus

The second example is connected to interferon- γ inducing factor:

N-terminus	thymopoietin	C-terminus
	GLPKKEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH	
	VESDKYMYIIILRTQPESASQDIDTMDEFVPPQRKDVFLVQDNINRIVAT	
C-terminus	interferon- γ inducing factor	N-terminus

The authors who cloned and identified interferon- γ inducing factor have stressed that it has no obvious similarity to any peptide in the database. The situation is, however, different, when the retro-sequence of this protein is taken into account. As regards the ubiquitin the presence of retro-sequence of RGD adhesive centre is worth to be mentioned.

The interesting situation takes also place in the case of porcine spermadhesin PSP II [7]. A 9-30 disulphide bridged loop of this protein shows a retro-thymopoietin homology with N-terminal loop of lactoferrin, which is endowed with immunostimulative activity:

9		30
	-CGRVIKDTSGSISNTDRQKNLC-	PSP II
	-DRKLCSVPPGRVKRMNRQWQFC-	human lactoferrin
40		19

It should be noted that in the spermadhesin molecule the thymopoietin homology is clearly expressed.

The common appearance of retro-thymopoietin-like motifs within the proteins suggest that they are of very old origin. It seems possible that both thymopoietin-like and retro-thymopoietin-like motifs have been developed from two DNA binding domains, interacting with two complementary polynucleotide strands in DNA duplex. This hypothesis, however, needs the careful and critical examination.

We found also that in some proteins there appear the discontinuous thymopoietin-like motifs. Such situation takes e. g. place in the very common muscle protein, actin G. The discontinuous thymopoietin-like motif of G actin is composed of

two peptide fragments (96-110, and 276-306) and two amino acid residues (Arg¹⁷⁷ and Asn¹⁶²), located between them.

```

GLPKVEVPAVLTKQKL - K - S - ELVANGVTLPAGEMRKDVYVELYLQHLTALH
VAPEEHPTLLTEAPL  R  N  ETTYNSIMKCDIDIRKDLYANNVMSGGTTMY
96                110 177 162 276                                306

```

The limited changes at the conformation of terminal residues (Leu¹¹⁰ and Glu²⁷⁶), as well as Arg¹⁷⁷, and Asn¹⁶² residues, enable the formation of continuous thymopoietin-like motif within the actin G molecule. According to Goldstein and Schlesinger [8], the systematic release of thymopoietin from thymus produces the phenomena characteristic for the serious neuromuscular disease, *myasthenia gravis*. This observation, taken together with the fact of appearance of thymopoietin-like motif within the molecule of actin G, suggests that both these products (thymopoietin and actin G) could interact with the same cellular receptor. Therefore the known X-ray structure of actin G can serve as the model for the determination of biologically active conformation of thymopoietin. The results of the application of such procedure leading to the conformational proposition for thymopoietin are reported by us elsewhere [9].

The 3D structure of actin G resembles that of ATP-ase domain of heat shock protein HSC 70 [10]. In agreement with this statement the discontinuous thymopoietin-like motif can be also observed in the HSC 70. It comprises the fragments 135-149, and 314-344, and two single residues: Gly²²⁴, and Val²⁰⁷.

```

GLPKVEVPAVLTKQKL - K - S - ELVANGVTLPAGEMRKDVYVELYLQHLTALH
LGKTVTNAVVTVPAY  G  V  LDPVEKALRDAKLDKSIWDIVLVGGSTRIP
135                149 224 207 314                                344

```

The discontinuous thymopoietin-like motif appears also in the molecule of human leukocyte class II antigen HLA-DQ [11]. It is created by two closely situated fragments of the peptide chain (89-108, and 158-181) of this protein. The fragment 158-181 contains the thymopoietin-like sequence, mentioned at the beginning of this text:

```

GLPKVEVPAVLTKQKLKSELVANGVTLPAGEMRKDVYVELYLQHLTALH
TLQRRVEPTVTIISPSRTEAL  VMLEMPQRGDVYTCHVEHPSLQSP
98                108 158                                181

```

The reason for the appearance of thymopoietin-like motifs in these proteins remains to be unclear. The better situation takes place in the case of actin. It is known that some residues of this motif, in actin, participate in the contact formation during the polymerization of the protein. It suggests that thymopoietin may influence this process, inhibiting it. The other possibility is connected to the finding that actin filaments of the cell cytoskeleton are involved in mRNA transport and anchoring [12]. It makes probable the participation of discontinuous thymopoietin-like motif of actin in

RNA binding, in agreement with our postulate that that such a motif is accomodated for recognition and binding of nucleic acids.

References:

- [1] Z. Szewczuk, P. Stefanowicz, A. Wilczyński, I. Z. Siemion, C. S. Cierniewski, U. Kalisz, Z. Wieczorek, *Biopolymers*, **40**, 571 (1997).
- [2] A. Hoffmann, Cheng-Ming Chiang, T. Oelschläger, Xinoling Xie, S. K. Burley, Y. Nakatami, R. G. Roeder, *Nature*, **380**, 356 (1996).
- [3] I. Z. Siemion, I. Strug, Z. Szewczuk, *Current Topics in Peptide Protein Res.*, in press.
- [4] Feng Guo, D. N. Gopaul, G. D. Van Duyene, *Nature*, **389**, 40 (1997).
- [5] J. H. M. Cabral, A. P. Jackson, C. V. Smith, N. Shikotra, A. Maxwell, R. C. Liddington, *Nature*, **388**, 903 (1997).
- [6] H. Okamura et al., *Nature*, **387**, 88 (1995).
- [7] A. Romero et al., *Nature Struct. Biol.*, **4**, 783 (1997).
- [8] G. Goldstein, D. H. Schlesinger, *Lancet*, **2**, 256 (1975).
- [9] I. Z. Siemion, I. Strug, Z. Szewczuk, *Acta Bioch. Polon.*, in press.
- [10] K. M. Flaherty, D. B. McKay, W. Kabsch, K. C. Holmes, *Proc. Natl. Acad. Sci. U. S. A.*, **88**, 541 (1991).
- [11] Z. Szewczuk, personal information.
- [12] C. L. Sundell, E. C. D. Stephenson, *Science*, **253**, 1275 (1991); P. A. Takizawa, A. Sil, J. R. Swedlow, I. Herskowitz, R. D. Vaie, *Nature*, **389**, 90 (1997).

HISTONE H1 - A WELL KNOWN PROTEIN WITH THE UNKNOWN FUNCTION

Marta Prymakowska-Bosak¹, Radosław Tomaszewski⁵, Marcin Przewłoka¹,
Joanna Ślusarczyk², Joanna Dłużniewska², Nicole Chaubet³, Claude Gigot³,
Steven Spiker⁴, Mieczysław Kuraś² and Andrzej Jerzmanowski^{1,5}

Laboratories of ¹ Plant Molecular Biology and ² Plant Morphogenesis, Warsaw University, Pawińskiego 5A, 02-106 Warsaw, Poland; ³ Institute de Biologie Moleculaire des Plantes du Centre National de la Recherche Scientifique, 12 rue du General Zimmer, 67000 Strasbourg, France; ⁴ Department of Genetics, College of Agricultural and Life Sciences, North Carolina State University, Raleigh, NC 27695-7614 and ⁵ Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5A, 02-106 Warsaw, Poland

Histone H1 is a key component of nucleosome though it is not involved in the formation of the nucleosomal histone core. To investigate the local and global functions of this protein we used both the *in vitro* system composed of chromatin reconstituted from purified histones and defined DNA fragments and the engineered transgenic plants that overexpressed or lacked the expression of histone H1 gene. The basic conclusion from these studies is that histone H1 does not function as global transcriptional repressor. However, it can regulate the transcription of specific classes of genes placed in vicinity of DNA sequences with high affinity for H1 and also may play a key role in processes, like meiosis, that require a regular pattern of heterochromatinization in condensed chromosomes.

In the nucleus of a eukaryotic cell, DNA does not exist in a „naked” form but is tightly complexed with proteins. The result of these interactions is a highly compacted form of DNA referred to as chromatin. Proteins that make a major contribution to the architecture of chromatin are histones. The basic chromatin fiber consists of an array of nucleosomes, sometimes compared to the „beads on a string”. The nucleosome is the most fundamental unit of chromatin organization. Each packages around 200 base pairs (bp) of DNA. Out of this 146 bp is wrapped twice (with approximately 80 bp/turn) in a left-handed superhelix around a core of eight histone molecules (an octamer), two each of H2A, H2B, H3 and H4. This structure is referred to as core particle. The remaining part of DNA forms a linker traversing to the next nucleosomal core. A histone of the fifth type, H1, interacts with DNA outside the core particle. In most organisms histone H1 (also known as linker histone) has a distinctive tripartite structure with a central globular domain and two basic unstructured tails forming the N- and C-terminal regions of the molecule. The precise location of H1 has not been firmly established (Pruss et al., 1995), although there is general agreement that it stabilizes the nucleosome and binds at least in part the linker DNA and protects another 20 bp of DNA from nuclease digestion in addition to the 146 bp protected by the core histone octamer. H1 facilitates the folding of the basic nucleosomal threads into 30-nm fibres (see van Holde, 1989 for review). Unlike the core histones, which maintain a high degree of sequence conservation throughout the evolution of eukaryotes,

linker histones show significant variability in sequence and structure, particularly in their unstructured N- and C-terminal tails.

Packaging of DNA into nucleosomes has been shown to interfere with both transcription initiation and elongation (reviewed by Kornberg and Lorch, 1995). Especially in respect to initiation, nucleosomes not only interfere with the binding of the transcription complex to the promoter region, but also mask sites for regulatory factors. The notion that the core histones are indeed serving as general repressors of transcription *in vivo* has gained strong support in the results obtained with budding yeast (reviewed by Grunstein, 1990). The role of H1 is by far less clear. Though the *in vitro* reconstitution of non-specific chromatin templates has led to the conclusion that histone H1 can repress transcription, the studies of the natural chromosomal templates did not substantiate this conclusion (reviewed by Wolffe et al., 1997).

In our laboratory, in order to gain new insight into the function of linker histones, we adopted the strategies using both *in vitro* and *in vivo* model systems.

The *in vitro* system comprises the 5S RNA genes of *Xenopus laevis*. The developmental regulation of transcription of the two types of 5S RNA genes in *X. laevis* has been a most thoroughly documented case of histone H1 involvement in modulation of transcription of a defined set of genes (reviewed by Wolffe, 1994). Two 5S RNA gene families are transcribed in early stages of embryonic development of *X. laevis*: the major oocyte-type, occurring in 20 000 copies per haploid genome, and the somatic type, occurring in 400 copies per haploid genome. From late gastrulation stage transcription of the oocyte-type 5S RNA genes becomes largely repressed whereas that of the somatic-type 5S RNA genes continues unaffected throughout consecutive developmental stages and during the adult life of the frog. Studies on the *in vitro* transcription of *Xenopus* somatic cell chromatin and later works on the elimination of somatic H1 during early embryogenesis using ribozyme strategies (reviewed by Wolffe, 1994) established that histone H1 is necessary to maintain the repression of oocyte-type 5S RNA genes in somatic cells. In our earlier work, based on the results of *in vitro* transcription studies using pure DNA templates, we suggested that the reason for the selective action of H1 *in vivo* may be the difference in base composition of the flanks accompanying the two types of 5S RNA genes (Jerzmanowski and Cole, 1990). We currently extended our studies to chromatin templates. By using fully defined *in vitro* system of chromatin reconstitution on plasmids with cloned oocyte- or somatic-type 5S gene repeats we found that the oocyte repeat which comprises a 120 bp oocyte-type 5S RNA gene placed within the few hundred bp long native AT-rich flanks, but not the somatic repeat (a similar 120 bp somatic-type 5S RNA gene placed within native GC-rich flanks) enables histone H1 to realign the nucleosomal core particles densely packed on plasmid DNA. The realignment results in creation of the nucleosomal repeat unit of about 240 bp and is achieved through complete removal of several core histone complexes from plasmid template with the oocyte-type repeat. This effect of H1 is independent on the plasmid sequences and seems to be solely due to the presence in the oocyte-type repeat of the AT-rich flanks. The effects of H1 are

completely suppressed by distamycin A, a drug that specifically recognizes and binds oligo(dA)-oligo(dT) runs in DNA. The binding of H1 results in increased protection of DNA sites within the AT-rich oocyte-type repeat. In an *in vitro* transcription assay performed with reconstituted chromatin templates containing plasmids with the oocyte- or somatic-type repeats only the transcription of the oocyte-type 5S RNA gene was repressed in the presence of physiological concentration of histone H1 (Tomaszewski and Jerzmanowski, 1997). The above results support the view that the AT-rich flanks of the oocyte-type 5S RNA gene are involved in histone H1 – mediated chromatin reorganisation that results in the transcriptional repression observed *in vivo*.

While the effects of H1 during early *Xenopus* development seem to be rather well proved, they concern a small class of genes flanked by specific DNA sequences with high affinity for H1. The question thus remained of whether other genes (and how many of them) could be similarly regulated, for example during embryogenesis, by the presence of H1? This is directly related to the question about the global role of H1 in the cell. In order to investigate *in vivo* the global role of H1 in transcriptional regulation during development of a multicellular organism, we made transgenic tobacco plants that either overexpressed the gene for *Arabidopsis* histon H1 (Prymakowska-Bosak et al., 1996) or lacked the expression of one of its own major H1 variants (Prymakowska-Bosak et al., unpublished). In all plants that overexpressed H1 the total H1-to-DNA ration in chromatin increased 2.3 – 2.8 times compared with the physiological level. This was accompanied by 50-100% decrease of native tobacco H1. The phenotypic changes in H1-overexpressing plants ranged from mild to severe perturbations in morphological appearance and flowering. No correlation was observed between the extent of phenotypic changes and the variation in the amount of overexpressed H1 or the presence or absence of the native tobacco H1. However, the severe phenotypic changes were correlated with early occurrence during plant growth of cells with abnormally heterochromatinized nuclei. Such cells occurred considerably later in plants with milder changes. Surprisingly, the ability of cells with highly heterochromatinized nuclei to fulfil basic physiological functions, including differentiation, was not markedly hampered. The results support the suggestion that chromatin structural changes dependent on H1 stoichiometry and on the profile of major H1 variants have limited regulatory effect on the activity of genes that control basal cellular functions. However, the H1 – mediated chromatin changes can be of much greater importance for the regulation of specific developmental programs.

The above conclusion was strongly supported by the observed phenotypic changes in transgenic plants with drastically reduced level of native H1. In many of these plants one of the two main variants of chromatin H1 has been completely eliminated. The dominant morphological defects in these plants were the underdevelopment of anthers leading to heterostyly, the delayed and asynchronous meiotic division of microspore mother cells and the absence in the microspore cells of the differentiating division, the so called pollen mitosis I. No defects were seen in the surrounding tapetal cells

indicating that the major defect linked to the decreased level of H1 could be traced to the stage of male meiosis.

In summary, our data indicate that while H1 is not a global transcriptional repressor, it may be critically involved in the regulation of specific class of genes (like the *Xenopus* 5S RNA genes) that are placed in the vicinity of DNA sequences with strong affinity for H1. The effect of H1 on such genes will be exerted through the initiation of the alignment of stable arrays of nucleosomes that can not be easily removed. On the other hand H1 could be of key importance in specific events, like meiosis, that involve pairing of highly condensed chromosomes. One is tempting to suggest that the function of H1 in these events could be linked to the requirement of maintaining a regular pattern of heterochromatinization.

Acknowledgements

This work was supported by Polish Committee of Scientific Research grants 662759201 and 6PO4A02208, M. Skłodowska-Curie Fund grant MEN/NSF 94-194 and by Howard Hughes Medical Institute grant 75195-543403.

References:

1. Grustein, M. (1990) *Annu. Rev. Cell Biol.* **6**, 643-678.
2. Jerzmanowski, A. and Cole, R.D. (1990) *J.Biol.Chem.* **265**, 10726-10732
3. Kornberg, R.D. and Lorch, Y. (1995) *Curr.Opin. Cell Biol.* **7**, 371-375.
4. Pruss, D., Hayes, J.J. and Wolffe, A.P. (1995) *Bioessays* **17**, 161-170.
5. Prymakowska-Bosak, M., Przewłoka, M., Iwkiewicz, J., Egierszdorff, S., Kuraś, M., Chaubet, N., Gigot, C., Spiker, S. and Jerzmanowski, A. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 10250-10255.
6. Tomaszewski, R. and Jerzmanowski, A. (1997) *Nucl. Acids Res.* **25**, 458-465.
7. van Holde, K.E. (1989) *Chromatin* (Springer, New York).
8. Wolffe, A.P. (1994) *Regulation of Chromatin Structure and Function*, (Landes, Austin, TX).
9. Wolffe, AP., Khochbin, S. and Dimitrov, S. (1997) *Bioessays*, **19**, 249-255.

STRUCTURE AND REACTION MECHANISM OF 2-HALOACID DEHALOGENASES

Kenji Soda, Tatsuo Kurihara*, and Nobuyoshi Esaki*

Department of Biotechnology, Faculty of Engineering, Kansai University
3-3-35 Yamate-Cho, Suita, Osaka 564, Japan (soda@ipcku.kansai-u.ac.jp)

*Institute for Chemical Research, Kyoto University, Uji, Kyoto-Fu 611, Japan

Abstract: The reaction mechanism of L-2-haloacid dehalogenase from *Pseudomonas* sp. YL proceeds through the formation of an ester intermediate formed between Asp10 and the α -carbon of substrate (L-2-chloropropionate), and Asp10 and several other catalytically important residues are found in the active site. In contrast, in the DL-2-haloacid dehalogenase reaction, the ester intermediate is not formed.

Halogenated compounds are widely used as herbicides, insecticides, plastics, solvents, and their starting materials. However, they cause environmental pollution as a result of their toxicity, persistence, and transformation into hazardous metabolites. Dehalogenases detoxify various halogen compounds and attract a great deal of attention from viewpoints of environmental technology. A few kinds of 2-haloacid dehalogenases (EC class: 3.8.1.2) catalyzing the hydrolytic dehalogenation of 2-haloacids to produce the corresponding 2-hydroxy acids have been demonstrated (Table 1).

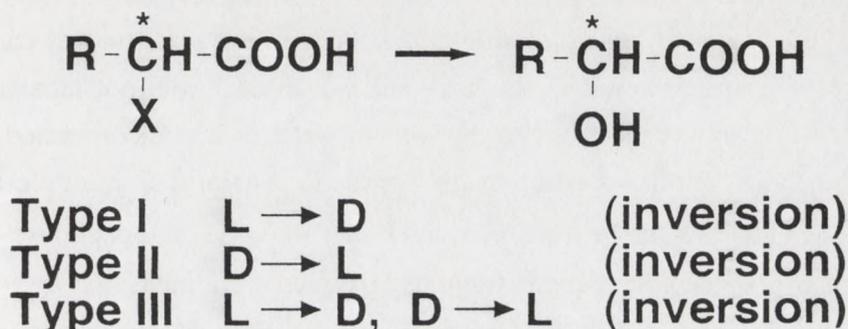


Table 1. 2-Haloacid Dehalogenases

However, catalytic mechanisms of 2-haloacid dehalogenases have not been elucidated. Two possible reaction mechanisms have been proposed as shown in Fig. 1.

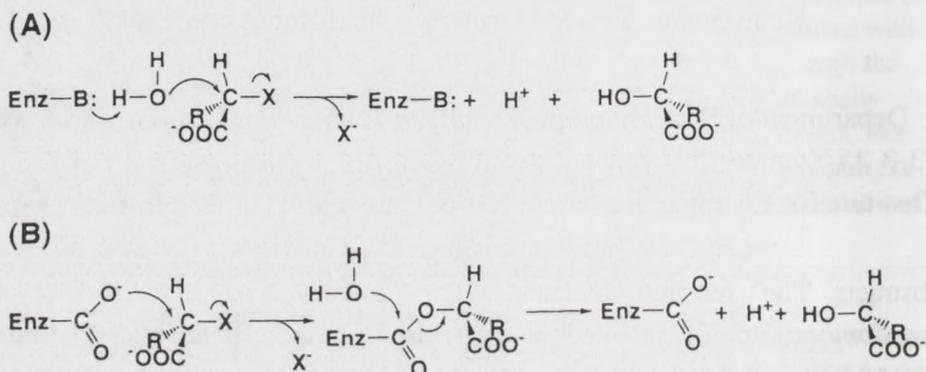


Fig. 1. Proposed mechanisms of DEX reactions.

L-2-Haloacid dehalogenase (L-DEX) catalyzes the hydrolytic dehalogenation of L-2-haloacids with inversion of the C_2 -configuration producing the corresponding D-2-hydroxy acids. We have isolated and purified thermostable L-2-haloacid dehalogenase from a 2-chloroacrylate-utilizable bacterium, *Pseudomonas* sp. YL, cloned its gene, and constructed the over-expression system.

We have analyzed the reaction mechanism of the enzyme from *Pseudomonas* sp. YL, and found that Asp10 is the active site nucleophile. When the multiple-turnover enzyme reaction was carried out in $H_2^{18}O$ with L-2-chloropropionate as a substrate, D-lactate produced was almost fully labeled with ^{18}O . However, when the single-turnover enzyme reaction was carried out by use of a large excess of the enzyme, the product was not labeled. This suggests that an oxygen atom of the solvent water is first incorporated into the enzyme, and then transferred to the product. After the multiple-turnover reaction in $H_2^{18}O$, the enzyme was digested with lysyl endopeptidase, and the molecular masses of the peptide fragments formed were measured by ion-spray mass spectrometer. Two ^{18}O atoms were shown to be incorporated into a

hexapeptide, Gly⁶-Lys¹¹. Tandem mass/mass spectrometric analysis of this peptide revealed that Asp¹⁰ was labeled with two ¹⁸O atoms. Our previous site-directed mutagenesis experiment showed that the replacement of Asp¹⁰ led to a significant loss in the enzyme activity. These results indicate that the carboxylate group of Asp¹⁰ acts as a nucleophile on the α -carbon of the substrate leading to the formation of an ester intermediate, which is hydrolyzed by nucleophilic attack of a water molecule on the carbonyl carbon atom.

We have found that the enzyme is paracatalytically inactivated by hydroxylamine in the presence of substrates such as monochloroacetate and L-2-chloropropionate. We showed by ion-spray mass spectrometry that a molecular mass of the enzyme inactivated by hydroxylamine during the dechlorination of monochloroacetate is higher about 74 Da higher than that of the native enzyme. To determine an increase of the molecular mass more precisely, we digested the inactivated enzyme with lysyl endopeptidase, and measured the molecular masses of the peptide fragments. The molecular mass of the hexapeptide, Gly⁶-Lys¹¹, was shown to increase by 73 Da. Tandem mass/mass spectrometric analysis of this peptide revealed that the increase is due to a modification of the active site Asp¹⁰. When the enzyme was paracatalytically inactivated by hydroxylamine during the dechlorination of L-2-chloropropionate, the molecular mass of the hexapeptide was 87 Da higher. Accordingly, hydroxylamine most probably attacks the carbonyl carbon atom of the ester intermediate to produce the inactivated and modified enzymes. The increments in molecular masses are thought to be caused by the formation of an aspartate β -hydroximate carboxyalkyl ester residue substituting for Asp¹⁰. Thus, we succeeded in trapping the enzyme-substrate complex using hydroxylamine. This is the first evidence for the formation of an ester intermediate in the L-2-haloacid dehalogenase reaction (Fig. 2).

The crystal structure of the homodimeric enzyme from *Pseudomonas* sp. YL has been determined by a multiple isomorphous replacement method and refined at 2.5 Å resolution to a crystallographic R-factor of 19.5%. The subunit

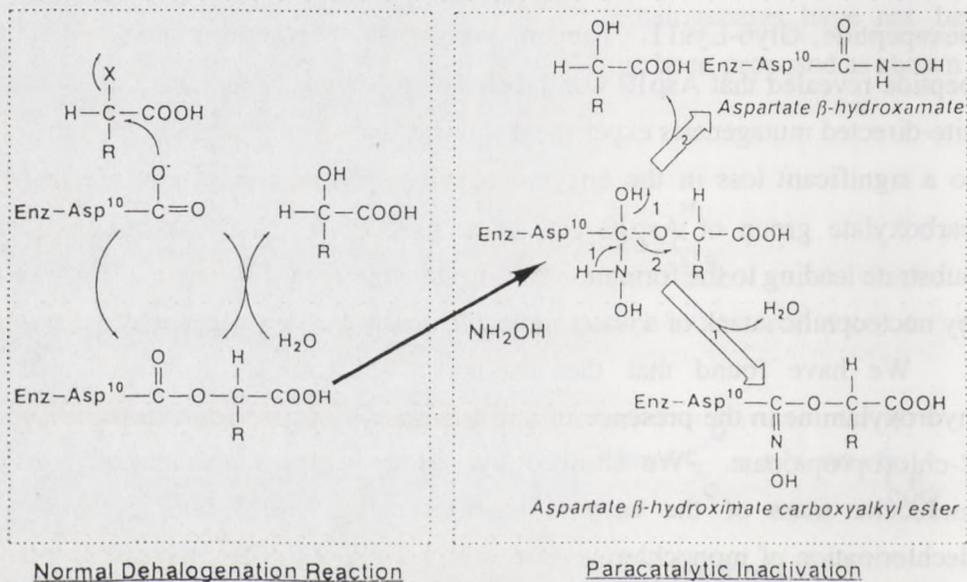
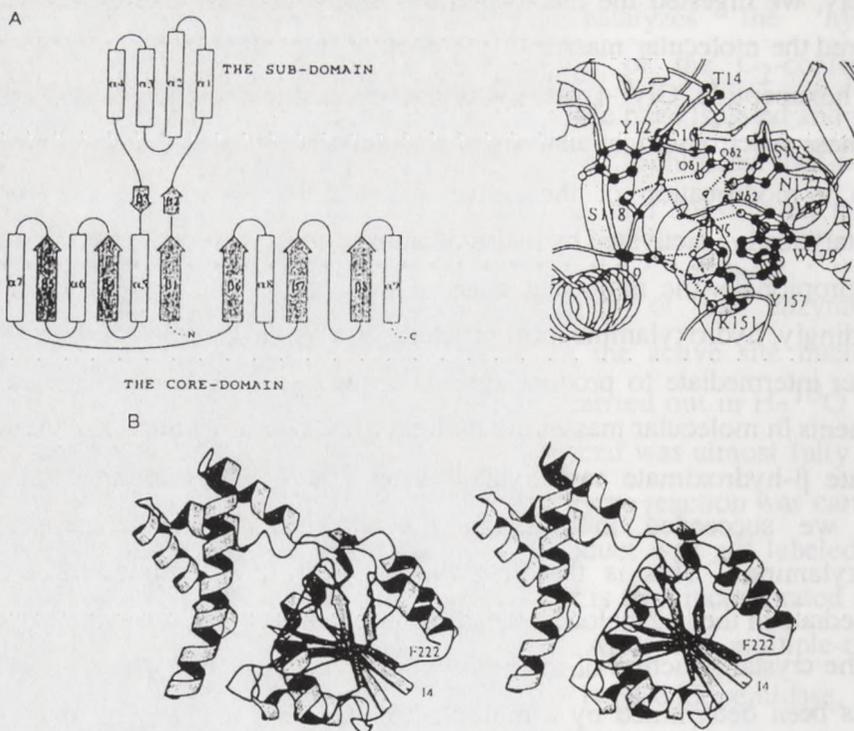


Fig. 2. Mechanism of paracatalytic inactivation of L-DEX.



Secondary structure (A), stereo view (B), putative active site of the L-DEX YL subunit. α -Helices are represented by spirals and β -sheets by arrows.

consists of two structurally distinct domains: the core-domain and the sub-domain. The core-domain has an α/β structure formed by a six-stranded parallel β -sheet flanked by five α -helices. The sub-domain inserted into the core-domain has a four-helical bundle structure providing the greater part of the interface for dimer formation. There is an active site cavity between the domains. An experimentally identified nucleophilic residue, Asp-10, is located on a loop following the *N*-terminal β -strand in the core-domain, and other functional residues, Thr-14, Arg-41, Ser-118, Lys-151, Tyr-157, Ser-175, Asn-177 and Asp-180, detected by a site-directed mutagenesis experiment, are arranged around the nucleophile in the active site. Though the enzyme is an α/β -type hydrolase, it does not belong to the α/β hydrolase fold family, from the viewpoint of the topological feature and the position of the nucleophile.

We purified DL-2-haloacid dehalogenase from *Pseudomonas* sp. 113 and characterized it. The enzyme is constitutively produced, and is a homodimer. It is unique in its stereospecificity: D- and L-haloacids are dehalogenated to form the corresponding L- and D-2-hydroxy acids, respectively. Whether L- or D-2-chloropropionate was used as a substrate of DL-DEX under the single-turnover reaction conditions, about 75% lactate produced contained ^{18}O in contrast to the L-DEX reaction. Thus, the DL-DEX reaction proceeds through the mechanism A in which the ester intermediate is not formed.

References

- Nardi-dei V, Kurihara T, Okamura T, Liu J, Koshikawa H, Ozaki H, Terashima Y, Esaki N and Soda K. *Appl. Environ. Microbiol.* **60**, 3375-3380 (1994).
- Kurihara T, Liu J-Q, Nardi-Dei V, Koshikawa H, Esaki N and Soda K. *J. Biochem.* **117**, 1317-1322 (1995).
- Liu J-Q, Kurihara T, Miyagi M, Esaki N and Soda K. *J. Biol. Chem.* **270**, 18309-18312 (1995).
- Hisano T, Hata Y, Fujii T, Liu J-Q, Kurihara T, Esaki N and Soda K. *J. Biol. Chem.* **271**, 20322-20330 (1996).

course of two-dimensional growth. The growth of the domain is determined by the rate of nucleation and the rate of growth. The nucleation rate is determined by the concentration of the nucleating agent and the rate of growth is determined by the rate of diffusion of the nucleating agent. The growth of the domain is determined by the rate of nucleation and the rate of growth. The nucleation rate is determined by the concentration of the nucleating agent and the rate of growth is determined by the rate of diffusion of the nucleating agent.

We have shown that the growth of the domain is determined by the rate of nucleation and the rate of growth. The nucleation rate is determined by the concentration of the nucleating agent and the rate of growth is determined by the rate of diffusion of the nucleating agent. The growth of the domain is determined by the rate of nucleation and the rate of growth. The nucleation rate is determined by the concentration of the nucleating agent and the rate of growth is determined by the rate of diffusion of the nucleating agent.

References

Hendel, V. (1994) *Journal of Polymer Science*, 33, 1-10.
Tashiro, Y. (1994) *Journal of Polymer Science*, 33, 11-20.
(1994)
Kunze, T. (1994) *Journal of Polymer Science*, 33, 21-30.
Bjork, L. (1994) *Journal of Polymer Science*, 33, 31-40.
Lin, J.-Q. (1994) *Journal of Polymer Science*, 33, 41-50.
1800-1810
Huang, J. (1994) *Journal of Polymer Science*, 33, 51-60.
Chen, Z. (1994) *Journal of Polymer Science*, 33, 61-70.

STRUCTURAL AND FUNCTIONAL ASPECTS OF SMOOTH MUSCLE REGULATION BY ACTIN-ASSOCIATED PROTEINS

Renata Dąbrowska and Robert Makuch

Nencki Institute of Experimental Biology, 3 Pasteur St., Warsaw, Poland

Abstract: Ca^{2+} -regulation of smooth muscle contractile activity involves a dual mechanism. Whereas phosphorylation of myosin regulatory light chains (LC_{20}) is essential for the initiation of muscle contraction, the actin filament-associated proteins, caldesmon and calponin, seem to participate in the control of the relaxation phase. Structure/function analysis of the latter two proteins and their localization in smooth muscle cells leads to the conclusions that caldesmon is directly involved in the regulation of the actin-myosin interaction, while calponin is not confined to the contractile apparatus and if it plays a regulatory role, it does so in an indirect manner.

Introduction

Despite the differences in the structural organization of the contractile apparatus, and of the excitation-contraction coupling pathway of various muscle types, the principal mechanisms of their functioning are the same: (i) force is generated *via* the cyclic weak- and strong-interactions with the actin filament of the myosin heads (cross bridges), which protrude from the myosin filament, and the concomitant hydrolysis of ATP; (ii) the shortening of muscles is achieved by a sliding mechanism through the interdigitation of the two filament types without changing their lengths [1]; (iii) myosin is the motor protein which hydrolyses ATP and the energy of hydrolysis is converted into a mechanical movement by a conformational change within the distal portion of the myosin head when it is strongly bound to actin filaments [2,3]; (iv) the rise and fall of intracellular calcium ions initiate muscle contraction and relaxation, respectively [4].

The proteins that confer Ca^{2+} sensitivity on the actin-myosin interaction and thus the molecular mechanism of its regulation is both different and specific for each muscle type. Biochemical and physiological studies indicate that the primary mechanism of smooth muscle regulation is phosphorylation of myosin light chains (LC_{20}) located in the regulatory domain of the myosin head adjacent to the motor domain. This phosphorylation is catalysed by a specific myosin light chain kinase activated by Ca^{2+} /calmodulin and is necessary and sufficient to remove the inhibition of the motor domain of each head by LC_{20} and to initiate muscle contraction. However, dephosphorylation of LC_{20} is not necessarily followed by muscle relaxation. Many smooth muscles (particularly vascular ones) can maintain tension due to slowly cycling cross-bridges (called latch bridges) when phosphorylation and intracellular Ca^{2+} concentration decline [5]. Therefore, it was postulated that a second regulatory system, which operates through actin filament-associated proteins and regulate the smooth muscle relaxation phase, is required. There were two potential candidates to fulfil this role, both discovered in Japan: first by Kenji Sobue and colleagues [6] at the University of Osaka, called caldesmon, and second, by Katsuhito Takahashi and

co-authors [7] at Ehime University, called calponin.

Despite their completely different structures, caldesmon and calponin share some functional properties. 1) They bind to F-actin, tropomyosin and Ca^{2+} /calmodulin, 2) they inhibit the actin-activated ATPase activity of myosin, sliding of actin filaments over immobilized myosin in *in vitro* motility assays and the active force in muscle fibers (e.g. [8-11]), and 3) the inhibitory effects of both are released by Ca^{2+} /calmodulin (or other Ca^{2+} -binding proteins) or by their phosphorylation mediated by Ca^{2+} /calmodulin-dependent kinase II and protein kinase C [12,13]. However, while all the inhibitory activities of caldesmon are potentiated by tropomyosin, those of calponin appear to be tropomyosin independent. Recent advances in the structure determination of caldesmon and calponin together with protein engineering and kinetic measurements allow us to hypothesize about their involvement in the smooth muscle regulation mechanism.

Structure/function of caldesmon

Smooth muscle caldesmon is a long, 75-80 nm, molecule with a 2 nm diameter and molecular weight of about 89 kDa. It binds actin and tropomyosin in a molar ratio of 1:7:1. All isoforms of caldesmon are encoded by the same gene and generated by alternative splicing [14]. However, in contrast to its non-muscle counterparts, smooth muscle caldesmon possesses 10 repeat motifs of 15 amino acids in the middle of the molecule. The repeat sequences have patterns of acidic and basic residues suggesting a helix stabilized by salt-bridge interactions [15]. This helix links the terminal parts of the molecule that are responsible for other protein interactions.

More than 10 years ago, our laboratory initiated the mapping of the actin- and calmodulin-binding sites of caldesmon. This work indicated that they are localized in the C-terminal domain responsible for the inhibition of actomyosin ATPase [16,17]. The efforts of several groups allowed for its more precise localization. Thus, the actin-binding region of caldesmon comprises three stretches of amino acid residues: 606-625, 690-710 and 713-737. The second region seems to be essential for tropomyosin-potentiated actomyosin ATPase inhibition, but it also requires the presence of at least one of the other flanking sequences [18]. Additionally, a lower affinity, site provides a caldesmon interaction with actin at 483-508 amino acid residues [19] and seems not to participate directly in the actomyosin ATPase inhibition. This site comprises the sequence LKEKQQ (residues 498-503) which has been proposed to be the counterpart of actin-binding motifs present in many other proteins, including myosin heavy chain, tropomyosin and thymosin β 4 [20].

Although the potentiation of caldesmon effect on actomyosin ATPase seems to be dependent on tropomyosin-actin binding rather than a direct caldesmon-tropomyosin interaction [21], the latter probably takes part in defining the location of caldesmon in the thin filament. The tropomyosin multiple contact binding sites of caldesmon were identified, but not precisely localized, within the regions corresponding to residues: 1-175, 199-431 and 606-657 [22-23]. Interactions with the first of these regions are, however, abolished in the presence of actin [24].

Three calmodulin binding sites were mapped to the residues 658-667, 686-

694, and 716-724 in the C-terminal region of caldesmon and named A, B and B', respectively [25]. Involvement of these particular sites in the reversal of actomyosin ATPase inhibition by caldesmon is still a matter of dispute [26].

The N-terminus of the caldesmon molecule contains the main binding site for the neck region of smooth-muscle myosin [27]. An additional region interacting with myosin, located at the C-terminal end of caldesmon (residues 581-657), has also been identified [23]. The physiological role of these interactions are still unclear. It is suggested that it may play a role in organizing the contractile apparatus [28]. Caldesmon binding induces conformational changes at the head/rod junction of myosin [29].

Structure/function of calponin

The calponin molecule is a prolate ellipsoid with molecular weight 31-32 kDa and a length of 16.2 nm and a diameter of 2.6 nm [30]. cDNA sequencing data suggests that there are three calponin genes encoding basic (*h1* and *h2*), and acidic calponins [31,32]. Moreover, *h1* calponin, the major species found in smooth muscle, may exist in the two isoforms α and β as produced by alternative splicing [31]. The cDNA of the acidic calponin isoform is encoded by separate gene isolated from the vascular smooth muscle of rat aorta. Its N-terminus corresponds to the other isoforms but it has an additional strongly acidic C-terminal domain composed of 57 residues. This isoform is also expressed in non-muscle tissues [32]

The N-terminal, mostly helical part of calponin, which is responsible for actin-binding, shares a sequence similarity with other actin-binding proteins. This so-called calponin-homology domain (CH domain) was found in the cytoskeletal proteins: spectrin, filamin, and α -actinin that cause cross-linking of actin filaments as well as in some signalling proteins, such as Ras-GAP (GTP-ase activating protein of Ras) and Vav [33]. The three-dimensional structure of one CH domain originating from human β -spectrin was determined at 2.0 Å resolution by X-ray crystallography [34].

Calponin contains one CH domain, but it also possesses a second actin-binding region in the central part of the sequence (residues 142-184). This region seems to involve multiple contact sites and is similar to the inhibitory segment of troponin I (residues 94-124). The hexapeptide sequence VKYAEK (142-147) in this region seems to be responsible for actomyosin ATPase inhibition whereas the sequence 148-182 anchors calponin to actin [35]. It is worthy of note that this hexapeptide resembles the above mentioned actin-binding motifs present in other actin binding proteins [20], including caldesmon. The presence of two distinct, functionally relevant, actin-binding sites in calponin can be predicted on the basis of its ability to form two functionally different complexes which are in equilibrium and are differently affected by calmodulin and myosin heads [10,36].

The stretch of residues 146-176 in calponin was shown to interact with the neck of smooth muscle myosin [37]. Since it is also responsible for actin binding, it was concluded that it is unlikely that calponin might function as a linker between myosin and actin. Sequences responsible for the interactions with tropomyosin and

calmodulin have been mapped to residues 7-52 and 153-163, respectively [38].

The C-terminal half of the calponin molecule is composed mainly of two or three repeats of 29 amino acid residues each starting with almost identical β -strand segment. Its functional role is so far unknown.

Caldesmon and calponin contact sites on actin

Caldesmon and calponin are competitive for binding to actin. However, the binding sites for these proteins on actin are not identical.

Studies concerning the mapping of caldesmon-contact sites on actin were performed using a variety of approaches [39-44]. The results of all these works demonstrate that caldesmon interacts with subdomain-1, comprising both the N- and C-termini, and part of actin subdomain-2. NMR studies [43] indicate that the caldesmon inhibitory region (residues 597-629) interacts with residues 1-7 of actin whereas the low affinity actin-binding site interacts with actin residues 20-41. Cys580 of caldesmon can be cross-linked to Cys374 in the C-terminus of actin [41]. A recent three-dimensional image reconstruction of the negatively stained F-actin-caldesmon complex from electron micrographs revealed that caldesmon closely approaches the N- and C-terminal amino acids on subdomain-1 of actin, continues up its side past residues 92 to 95 as far as residue 40 on subdomain-2 before connecting to the next actin monomer [44]. In these filaments, strong myosin binding sites on actin are flanked, but not blocked, by both tropomyosin and caldesmon, whereas weak-binding sites are, at least in part, covered by the caldesmon molecule extending longitudinally along the long-pitch actin helix.

Calponin can also be covalently cross-linked to actin C-terminus (residues 326-355) by its N-terminal region (residues 52-168). However modification of N-terminal amino acids on actin, that weakens caldesmon binding, has no effect on calponin binding [45]. Three-dimensional reconstruction of calponin-containing actin filaments shows that it is regularly arranged along actin filaments in such a way that each molecule joins subdomain-1 with subdomain-2 of neighbouring actin protomers [46]. Comparison of the caldesmon-actin and calponin-actin interfaces with that of the myosin head-actin interface revealed that whereas caldesmon interferes with the N-terminal, weak-binding sites of the myosin heads, calponin at equimolar ratio to actin monomer could interfere with strong-binding sites of the myosin heads to actin filaments. However, for full inhibition of the ATPase a molar ratio of 1 calponin per 3 actin monomers is sufficient [36] and in this case the arrangement of calponin on actin filament would be probably different.

Mechanisms of caldesmon and calponin inhibition

To understand how regulatory proteins (caldesmon and calponin) may affect the actomyosin motor it is necessary to consider the mechanism of ATP hydrolysis. During the chemomechanical cycle, i.e. the cycle of ATP hydrolysis coupled with conformational changes in myosin heads causing movement of the actin filament, the affinity of myosin heads to actin is controlled by nucleotides bound to the head [47]. In the absence of nucleotides the tight rigor complex is formed, which dissociates when ATP is bound to the myosin head. Head-bound ATP is rapidly hydrolysed to

ADP and P_i and the energy of this process is stored in a different conformation of the head. Weak-binding of actin accelerates P_i release which restores the strong-binding complex. The transition to strong binding is accompanied by movement of the distal part of the head producing an up to 10 nm stroke. Subsequent release of ADP returns the cycle to a rigor complex.

The inhibitory proteins can affect either binding of myosin head to actin filament or one of the steps of the kinetics in ATP hydrolysis. Analysis of the effects of caldesmon and calponin on the ATP hydrolysis kinetics and smooth muscle fiber tension clearly shows that the inhibitory molecular mechanisms of the of the smooth muscle chemomechanical cycle by caldesmon and calponin are different. Whereas caldesmon causes an inhibition of the weak-binding state of the actomyosin complex, i.e. cross-bridge number [9], calponin is an effective inhibitor of the ATP hydrolysis rate through an inhibition of product release and regulation of the cross-bridge cycling rate [11]. The inhibitory effect of both proteins on the chemomechanical cycle is removed upon an increase in the intracellular Ca^{2+} concentration by calmodulin or protein phosphorylation.

Localization of caldesmon and calponin in smooth muscle

To discuss the role played by caldesmon and calponin *in vivo* it was necessary to determine the composition and distribution of thin filaments in smooth muscle cells. Analysis of the native thin filaments from smooth muscle revealed the presence of calponin and caldesmon, in addition to actin and the two tropomyosin subunits. However, it appeared that caldesmon-rich and calponin-rich filaments could be fractionated and separated from crude muscle extracts by immunoprecipitation using antibodies against these proteins [48]. This strongly suggested that caldesmon and calponin are present in two distinct populations of thin filaments. Immunocytochemical studies confirmed the existence of the two classes of actin filaments in avian and mammalian smooth muscle: one containing caldesmon and smooth muscle γ -actin isoform localized in the contractile domain, and a second containing calponin and cytoplasmic β -actin isoform, present predominantly in the cytoskeletal domain together with intermediate filaments [49,50]. In agreement with these data, biochemical experiments showed a lack of interaction between caldesmon and calponin [51], a competitive binding of both to actin filaments [52] and a direct interaction between calponin and intermediate filament protein - desmin [53].

Concluding remarks

Taken together, these results undoubtedly support participation of caldesmon in the regulation of the actin-myosin interaction in smooth muscle but the role of calponin in this process remains unclear. It can be supposed that calponin as a bridging protein between actin filaments, and between actin filaments and intermediate filaments can in indirect way provide resistance to movement. In view of the results suggesting a redistribution of calponin upon the agonist stimulation of vascular smooth muscle cells [54], the recent data showing a direct inhibition of latch-bridges by calponin [55] seem also to be possible. However, much more evidence is required to substantiate these possibilities.

- [1] Huxley, H.E. (1996) *Annu. Rev. Physiol.* 58, 1-19. [2] Rayment, I., Holden, H. (1994) *Trends Biochem. Sci.* 19, 129-134. [3] Ruppel, K.M., Spudich, J.A. (1996) *Annu. Rev. Dev. Biol.* 12, 543-573. [4] Somlyo, A.P., Himpens, B. (1989) *FASEB J.* 3, 2266-2276. [5] Dillon, P.F., Aksoy, M.O., Driska, S.P., Murphy, R.A. (1981) *Science* 211, 495-497. [6] Sobue, K., Muramoto, Y., Fujita, M., Kakiuchi, S. (1981) *Proc. Natl. Acad. Sci. USA* 78, 5652-5655. [7] Takahashi, K., Hiwada, K., Kokubu, T. (1986) *Biochem. Biophys. Res. Commun.* 141, 20-26. [8] Szpacenko, A., Wagner, J., Dąbrowska, R., Rüegg, J.C. (1985) *FEBS Lett.* 192, 9-12. [9] Malmqvist, U., Arner, A., Makuch, R., Dąbrowska, R. (1996) *Pflügers Arch.*, 432, 241-247. [10] Kolakowski, J., Makuch, R., Stepkowski, D., Dąbrowska, R. (1995) *Biochem. J.*, 306, 199-204. [11] Jaworowski, A., Anderson, K.I., Arner, A., Enström, M., Gimona, M., Strasser, P., Small, J.V. (1995) *FEBS Lett.* 365, 167-171. [12] Dąbrowska, R. (1994) in: *Airways Smooth Muscle* (Raeburn, D., Giembycz, M.A., eds) pp. 31-59, Birkhauser Verlag, Basel. [13] Chalovich, J.M., Pfitzer, G. (1997) in: *Cellular Aspects of Smooth Muscle Function* (Kao, C.Y., Carsten, M.E., eds) pp. 253-287, Cambridge University Press, Cambridge. [14] Haruna, M., Hayashi, K., Yano, H., Takeuchi, O., Sobue, K. (1993) *Biochem. Biophys. Res. Commun.* 197, 145-153. [15] Wang, C.-L.A., Chalovich, J.M., Graceffa, P., Lu, R.C., Mabuchi, K., Stafford, W.F. (1991) *J. Biol. Chem.* 266, 13958-13963. [16] Szpacenko, A., Dąbrowska, R. (1986) *FEBS Lett.* 202, 182-186. [17] Makuch, R., Walsh, M.P., Dąbrowska, R. (1989) *FEBS Lett.* 247, 411-414. [18] Fraser, I.D.C., Copeland, O., Bing, W., Marston, S.B. (1997) *Biochemistry* 36, 18, 5483-5492. [19] Leszyk, J., Mornet, D., Audemard, F., Collins, J.H. (1989) *Biochem. Biophys. Res. Commun.* 160, 1371-1378. [20] Vancompernelle, K., Goethals, M., Huet, C., Louvard, D., Vandekerckhove, J. (1992) *EMBO J.* 11, 4739-4746. [21] Marston, S.B., Redwood, C.S. (1993) *J. Biol. Chem.* 268, 12317-12320. [22] Fraser, I.D.C., Huber, P.A.J., Torok, K., Slatter, D.A., Gusev, N.B., Marston, S.B. (1994) *J. Muscle Res. Cell Motil.* 15, 218. [23] Huber, P.A.J., Fraser, I.D.C., Marston, S.B. (1995) *Biochem. J.* 312, 617-625. [24] Tsuruda, T.S., Watson, M.H., Foster, D.B., Lin, J.J.J.-C., Mak, A.S. (1995) *Biochem. J.* 309, 951-957. [25] Marston, S.B., Fraser, I.D.C., Huber, P.A.J., Pritchard, K., Gusev, N.B., Torok, K. (1994) *J. Biol. Chem.* 269, 8134-8139. [26] Wang, Z., Yang, Z.Q., Chacko, S. (1997) *J. Biol. Chem.* 272, 16896-16903. [27] Ikebe, M., Reardon, S. (1988) *J. Biol. Chem.* 263, 3055-3058. [28] Yamashiro, S., Matsumura, F. (1991) *BioEssays* 13, 563-568. [29] Kulikova, N., Dąbrowska, R. (1996) *Biochem. Biophys. Res. Commun.* 225, 195-202. [30] Stafford, W.F., Mabuchi, K., Takahashi, K. and Tao, T. (1995) *J. Biol. Chem.* 270, 10576-10579. [31] Takahashi, K., Nadal-Ginard, B. (1991) *J. Biol. Chem.* 266, 13284-13288. [32] Applegate, D., Feng, W., Green, R.S., Taubman, M.N. (1994) *J. Biol. Chem.* 269, 10683-10690. [33] Castresana, J. and Saraste, M. (1995) *FEBS Lett.* 374, 149-151. [34] Djinic-Carugo, K., Bañuelos, S., Saraste, M. (1997) *Nature Struct. Biol.* 4, 175-179. [35] EL-Mezgueldi, M., Strasser, P., Fattoum, A., Gimona, M. (1996) *Biochemistry* 35, 3654-3661. [36] Kolakowski, J., Karkucińska, A., Dąbrowska, R. (1997) *Eur. J. Biochem.* 243, 624-629. [37] Szymański, P., Tao, T. (1997) *J. Biol. Chem.* 272, 11142-11146. [38] EL-Mezgueldi, M., Mendre, C., Calas, B., Kassab, R., Fattoum, A. (1995) *J. Biol. Chem.* 270, 8867-8876. [39] Makuch, R., Kolakowski, J., Dąbrowska, R. (1992) *FEBS Lett.* 297, 237-240. [40] Kolakowski, J., Makuch, R., Dąbrowska, R. (1992) *FEBS Lett.* 309, 65-67. [41] Graceffa, P., Adam, L.P., Lehman, W. (1993) *Biochem. J.* 294, 63-67. [42] Crosbie, R.H., Miller, C., Chalovich, J.M., Rubenstein, P.A., Reisler, E. (1994) *Biochemistry* 33, 3210-3216. [43] Mornet, D., Bonet-Kerrache, A., Strasburg, G.M., Patchell, V.B., Perry, S.V., Huber, P.A.J., Marston, S.B., Slatter, D.A., Evans, J.S., Levine, B.A. (1995) *Biochemistry* 34, 1893-1901. [44] Lehman, W., Vibert, P., Craig, R. (1997) 26th European Muscle Conference, Schneverdingen, 21-26 September 1997, Abstr. 75. [45] Miki, M., Walsh, M.P., Hartshorne, D.J. (1992) *Biochem. Biophys. Res. Commun.* 187, 867-71. [46] Hodgkinson, J.L., EL-Mazgueldi, M., Graig, R., Vibert, P., Marston, S.B., Lehman, W. (1997) *J. Muscle Res. Cell Motil.* 18, 197. [47] Spudich, J.A. (1994) *Nature*, 372, 515-518. [48] Lehman, W. (1991) *J. Muscle Res. Cell Motil.* 12, 221-224. [49] Fürst, D.O., Cross, R.A., DeMey, J., Small, J.V. (1986) *EMBO J.* 5, 251-257. [50] North, A.J., Gimona, M., Cross, R.A., Small, J.V. (1994) *J. Cell Sci* 107, 437-444. [51] Czuryło, E.A., Kulikova, N., Dąbrowska, R. (1997) *J. Biol. Chem.*, in the press. [52] Makuch, R., Birukov, K., Shirinsky, V., Dąbrowska, R. (1991) *Biochem. J.* 280, 33-38. [53] Wang, P., Gusev, N.B. (1996) *FEBS Lett.* 392, 255-258. [54] Parker, C.A., Takahashi, K., Tao, T., Morgan, K.G. (1994) *Am. J. Physiol.* 267, C1262-C1270. [55] Malmqvist, U., Trybus, K.M., Yagi, S., Carmichael, J., Fay, S. (1997) *Proc. Natl. Acad. Sci. USA* 94, 7655-7660.

Acknowledgements

This work was supported by grants from the State Committee for Scientific Research (6P04A 056 09 and to the Nencki Institute).

BIOTECHNOLOGICAL IMPLICATION
OF PROTEIN RESEARCH

Abeta
injury
arrest
cells
hydro

prote
in the
my M
demon
depe
phos
GPs
cryst
sub
direct
cont
hydro
dian
point

spec
syst
direct
wher
stap
in ex
hydro

Trans

was
TTR
NIMA
TTRC
TTRD
PTIR

1894

1895

1896

1897

1898

1899

1900

1901

1902

1903

1904

1905

1906

1907

1908

1909

1910

1911

1912

1913

1914

1915

1916

1917

1918

1919

1920

1921

1922

1923

1924

1925

1926

1927

1928

1929

1930

1931

1932

1933

1934

1935

EXPRESSION OF PHOSPHOLIPID HYDROPEROXIDE GLUTATHIONE PEROXIDASE

Kunio Yagi

*Institute of Applied Biochemistry, Yagi Memorial Park,
Mitake, Gifu 505-01, Japan*

Abstract: For the purpose to protect the human body from lipid hydroperoxide-mediated injury, we made an experiment to transfect cells of guinea pig cell line 104C1 with a gene encoding phospholipid hydroperoxide glutathione peroxidase. The transfectant 104C1/O4C cells acquired a strong resistance to both phospholipid hydroperoxide- and linoleic acid hydroperoxide-mediated injury.

The lipid hydroperoxide-decomposing system has a crucial role in protecting the human body from the deleterious effects of lipid hydroperoxides. In this respect, glutathione peroxidase (GPx), which was first discovered in 1957 by Mills [1], is well known to be important in the cellular defense system for decomposing lipid hydroperoxides. Up to now, at least four types of selenium-dependent GPx have been characterized. These are classic GPx [1, 2], phospholipid hydroperoxide GPx (PHGPx) [3-5] or monomeric GPx [6], plasma GPx [7, 8], and GPx GI [9]. Among them, so-called classic GPx is present in the cytosol of various tissues and red blood cells at a high level. Although classic enzyme catalyzes the reduction of free fatty acid hydroperoxides, it does not directly react with phospholipid hydroperoxides in cell membranes [10]. On the contrary, PHGPx or monomeric GPx can react with phospholipid hydroperoxides, and, therefore, this type of GPx seems to be much more crucial than classic GPx in preventing the elevation of lipid peroxide levels. In fact, this point was revealed in selenium-deficient rats [11].

For prevention of the deleterious effects of lipid hydroperoxides, we expected that increasing enzymatic decomposition of the lipid hydroperoxides by introducing the gene of PHGPx would be the most preferable approach, and decided first to express its gene in appropriate cells by transfection to see whether the expression of this enzyme could protect host cells from the injury caused by lipid hydroperoxides. In this paper, expression of human PHGPx gene in cells of a guinea pig cell line and its protective effect against lipid hydroperoxide-mediated cell injury [12, 13] are summarized.

Transfection of cultured guinea pig cells with PHGPx gene

A cDNA having the entire amino acid-coding region of human PHGPx was obtained by PCR amplification with human testis QUICK Clone cDNA. Primers were synthesized on the basis of the reported nucleotide sequence of human testis PHGPx cDNA [14]. The amplified DNA was ligated into the *HincII* site of pUC 19. The insert DNA was excised from the plasmid with *HindIII* and *XbaI*, and inserted into the multicloning site of pRc/CMV, the PHGPx cDNA being positioned downstream of the CMV promoter. The final

Table 1. GPx activity of 104C1 cells transfected with pCMV-hPHGPx

Cells	GPx activity (mU/mg protein)		
	PCOOH	LAOOH	<i>t</i> -ButOOH
Parental cells			
104C1	3.0	61.0	13.8
Transfectants			
104C1/O4C	32.1	265.9	13.2
104C1/O4D	3.4	58.8	15.5

Cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum.

construct (designated pCMV-hPHGPx) was used for transfection of 104C1 guinea pig cells by means of our method using cationic multilamellar liposomes [15]. Transfected cells were selected with G418, and we finally isolated several clones of transfectants. Table 1 shows GPx activity of the parental 104C1 cells and of the transfectants. Since it is known that guinea pig cells scarcely produce selenium-dependent GPx [16], the GPx activity observed in 104C1 cells could be due to glutathione S-transferase. Among the transfectants obtained, the transfectant 104C1/O4C cells had the highest GPx activity toward dilaoleoyl phosphatidylcholine hydroperoxide (PCOOH), whereas such activity was not found in the transfectant 104C1/O4D cells. As is well known, PHGPx cannot react with *t*-butyl hydroperoxide (*t*-ButOOH), and thus it is natural that 104C1/O4C cells did not show any increase in GPx activity toward this substance. The elevation of GPx activity toward linoleic acid hydroperoxide (LAOOH) in the transfectant 104C1/O4C cells was also confirmed.

A band of protein immunoreactive to anti-rat PHGPx antibody was clearly observed with the extract of 104C1/O4C cells, which expressed high GPx activity toward PCOOH. The control 104C1 cells and 104C1/O4D cells had no such immunoreactive protein.

Protection of host cells from lipid hydroperoxide-mediated injury by expression of PHGPx

Since the transfectant 104C1/O4C cells exhibited remarkable expression of PHGPx, the susceptibility of this clone to lipid hydroperoxides was examined in comparison with that of 104C1 cells.

When 104C1/O4C cells were incubated with 0.3 μ mol/ml PCOOH in 0.5 ml RPMI 1640 medium containing 3% fetal bovine serum, the release of lactate dehydrogenase (LDH) activity into the medium 24 hr after the start of incubation was very low, *i. e.*, less than 10%, whereas the parental 104C1 cells released over 70% of their LDH activity into the medium under the same conditions (Table 2). The LDH activity released into the medium from the

Table 2. Protection of cells from PCOOH-mediated injury by expression of PHGPx

Cells	LDH release (% of total)	Cell viability (%)
Parental cells		
104C1	73 ± 4	42 ± 24
Transfectant		
104C1/O4C	8 ± 3	92 ± 3

Cells were incubated with PCOOH for 24 hr at a concentration of 0.3 $\mu\text{mol/ml}$. Data indicate mean \pm SD ($n = 3$ for LDH release and 4 for cell viability). Cell viability was measured by the MTT assay.

transfectant 104C1/O4D cells, which do not express PHGPx, was at the same level as that from the parental cells.

Cell viability was assessed by the MTT assay after the incubation with PCOOH (0.3 $\mu\text{mol/ml}$) for 24 hr. We found that the viability of the transfectant 104C1/O4C cells was 92%, whereas that of the parental 104C1 cells was only 42% (Table 2), indicating that the cell viability of the transfectant was almost 2 times higher than that of the parental cells.

We also examined morphologically the cells by phase-contrast light microscopy. Upon the incubation with PCOOH, the parental cells changed to a round shape, whereas 104C1/O4C cells still maintained their normal shape.

Susceptibility of the transfectant 104C1/O4C cells to lipid hydroperoxide-mediated injury was further studied with LAOOH. When the parental 104C1 cells were incubated with 0.3 $\mu\text{mol/ml}$ LAOOH, the release of LDH into the medium occurred 2 hr after the start of incubation. This release reached nearly 60% after the incubation for 6 hr. When 104C1/O4C cells were incubated under the same conditions, LDH activity released into the medium was very low; *i. e.*, only 10% release was found during a 6-hr incubation period. On the other hand, 104C1/O4D cells released the same level of LDH activity as the parental cells after the 6-hr incubation with LAOOH. After the incubation with LAOOH for 6 hr, cell viability of the transfectant 104C1/O4C cells was 3 times higher than that of the parental 104C1 cells. These results are in principle the same as those obtained with PCOOH.

All these results clearly demonstrate that the transfectant 104C1/O4C cells were extremely resistant to both PCOOH- and LAOOH-mediated injury. Since the transfectant 104C1/O4D cells had no GPx activity toward PCOOH and were susceptible to lipid hydroperoxide-mediated injury as in the case of the parental cells, we consider that the expression of PHGPx in 104C1/O4C cells is certainly responsible for the protection of the cells from the injury. Considering that the increase in lipid hydroperoxide level in the body is causative of age-related diseases or even aging itself, we are now convinced that on the line of this research an efficient method to prevent or cure these diseases or to prevent aging will be found.

REFERENCES

1. Mills, G. C. (1957) *J. Biol. Chem.* 229, 189-197.
2. Flohé, L., Günzler, W. A., and Schock, H. H. (1973) *FEBS Lett.* 32, 132-134.
3. Ursini, F., Maiorino, M., Valente, M., Ferri, L., and Gregolin, C. (1982) *Biochim. Biophys. Acta* 710, 197-211.
4. Ursini, F., Maiorino, M., and Gregolin, C. (1985) *Biochim. Biophys. Acta* 839, 62-70.
5. Maiorino, M., Roveri, A., Gregolin, C., and Ursini, F. (1986) *Arch. Biochem. Biophys.* 251, 600-605.
6. Duan, Y.-J., Komura, S., Fiszler-Szafarz, B., Szafarz, D., and Yagi, K. (1988) *J. Biol. Chem.* 263, 19003-19008.
7. Maddipati, K. R., Gasparski, C., and Marnett, L. J. (1987) *Arch. Biochem. Biophys.* 254, 9-17.
8. Takahashi, K., Avissar, N., Whitin, J., and Cohen, H. (1987) *Arch. Biochem. Biophys.* 256, 677-686.
9. Chu, F.-F., Doroshov, J. H., and Esworthy, R. S. (1993) *J. Biol. Chem.* 268, 2571-2576.
10. McCay, P. B., Gibson, D. D., Fong, K.-L., and Hornbrook, K. R. (1976) *Biochim. Biophys. Acta* 431, 459-468.
11. Guan, J.-Y., Komura, S., Ohishi, N., and Yagi, K. (1995) *Biochem. Mol. Biol. Int.* 37, 1103-1110.
12. Yagi, K., Komura, S., Kojima, H., Sun, Q., Nagata, N., Ohishi, N., and Nishikimi, M. (1996) *Biochem. Biophys. Res. Commun.* 219, 486-491.
13. Sun, Q., Kojima, H., Komura, S., Ohishi, N., and Yagi, K. (1997) *Biochem. Mol. Biol. Int.* 42, 957-963.
14. Esworthy, R. S., Doan, K., Doroshov, J. H., and Chu, F.-F. (1994) *Gene* 144, 317-318.
15. Yagi, K., Noda, H., Kurono, M., and Ohishi, N. (1993) *Biochem. Biophys. Res. Commun.* 196, 1042-1048.
16. Shreve, M. R., Morrissey, P. G., and O'Brien, P. J. (1979) *Biochem. J.* 177, 761-763.

STRUCTURAL AND ENERGETIC ASPECTS OF PROTEIN INHIBITOR - SERINE PROTEINASE RECOGNITION

Jacek Otlewski, Włodzimierz Apostoluk, Olga Buczek, Liliana Cholawska,
Agnieszka Grzesiak, Katarzyna Kościelska, Izabela Krokoszyńska, Daniel
Krowarsch, Damian Stachowiak, Michał Dadlez

Institute of Biochemistry & Molecular Biology, University of Wrocław,
Tamka 2, 50-137 Wrocław, Poland, e-mail: otlewski@bf.uni.wroc.pl
*Institute of Biochemistry & Biophysics, Polish Academy of Sciences,
Pawińskiego 5a, 02-106 Warsaw, Poland, e-mail: michald@ibbrain.ibb.waw.pl

Abbreviations: BPTI, basic pancreatic trypsin inhibitor (Kunitz); CMTI I, *Cucurbita maxima* trypsin inhibitor I; HNE, human neutrophil elastase; OMTKY3, turkey ovomucoid third domain; ASA - solvent accessible surface area; ΔG_{ass} , ΔH_{ass} , ΔS_{ass} and $\Delta C_{p,\text{ass}}$ free energy change, enthalpy change, entropy change and heat capacity change accompanying protein-protein association; ΔG_{den} , ΔH_{den} , ΔS_{den} , $\Delta C_{p,\text{den}}$ and T_{den} free energy change, enthalpy change, entropy change, heat capacity change during protein denaturation; ΔH_{cal} and ΔH_{vH} calorimetric and van't Hoff denaturation enthalpies; K_a , association constant; rmsd, root mean square deviation.

Introduction. Specific protein-protein recognition is a key event in many biological processes (Jones & Thornton, 1996). Formation of the specific complexes between antigen and antibody, hormone and receptor, enzyme and inhibitor are classic examples of highly complementary and specific interactions which are vital to living organisms. Serine proteinases and their canonical protein inhibitors are most intensively studied group of protein-protein complexes. In this paper we present several different approaches which are currently used in our group to understand structural and energetic factors which ensure specific and strong interactions between serine proteinase and protein inhibitor. Our methodology includes: molecular and structural biology, enzymology, structural thermodynamics and kinetics of protein-protein interaction, phage displayed combinatorial libraries, and thus obeys some of major *Trends in Protein Research*. We focused on two protein inhibitors - bovine pancreatic trypsin inhibitor (Kunitz) (BPTI) and *Cucurbita maxima* trypsin inhibitor I (CMTI I) which are small proteins belonging to two different inhibitor families and exhibit different folds.

Three-dimensional structures. Figure 1A presents superimposed three-dimensional X-ray structures of CMTI I (Bode et al., 1989a) and BPTI (Wlodawer et al., 1987). Both proteins exhibit completely different folds: CMTI I is a 29-residue protein containing three disulfide bonds which are the major structural determinants. Its secondary structure is composed of three rather irregular and short β -strands, three β -turns and a short piece of 3_{10} helix. Besides three buried disulfides, CMTI I does not exhibit hydrophobic core. The organization of the secondary structure elements in 58-residue BPTI is completely different. BPTI has a clear hydrophobic core, triple-stranded β -sheet and short α -helices at both N- and C-termini of the molecule. Despite these structural differences both proteins have the proteinase binding loop of similar, so called, canonical conformation. The conformation of the loop is very unusual and very seldom observed in structures of other PDB proteins. The extended conformation of the loop spans from P3 to P3' position (notation of Schechter and Berger (1967) and is defined by the main chain torsional angles or C_{α} - C_{α} distances

(Apostoluk & Otlewski, manuscript submitted). Similar conformation of the loop and, therefore, similar mode of recognition is a common feature of all remaining 18 inhibitor families, including BPTI (Fig. 1B). This strongly points to multiple convergent evolution and biological significance of protein inhibitors (Laskowski & Kato, 1980).

Figure 1A. Three-dimensional structure of CMTI I (black) and BPTI (grey). The structures are superimposed to minimize rmsd of proteinase binding loop segments P3-P3'. **B.** The P4-P4' backbone segment of the canonical proteinase binding loops of CMTI I (black) and BPTI (grey). **C.** Interface region of the CMTI I- trypsin complex. Trypsin is shown in grey and CMTI I in black. **D.** X-ray structure of CMTI I (grey) as determined in the complex with trypsin (Bode et al., 1989a) superimposed on NMR solution structure (black) of the inhibitor (Holak et al., 1989).

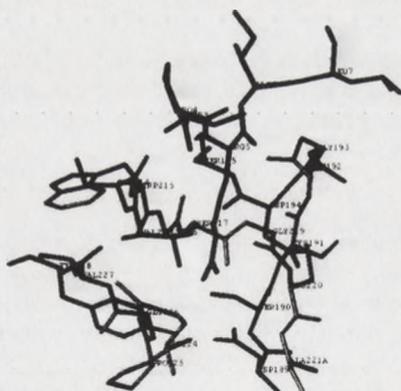
A.



B.



C.



D.

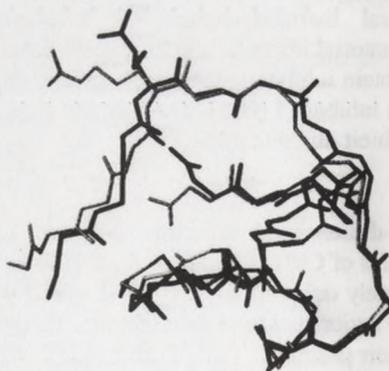


Figure 1C shows a recognition of bovine β -trypsin by CMTI I. The hydrophobic and convex proteinase binding loop of inhibitor is highly complementary to the concave active site of the enzyme. Rigid conformation of the CMTI I loop is maintained via a complicated network of hydrogen bonds from three internal water molecules to carbonyls of P2 and P1' residues. Besides P3-P3' segment, also side chains from surrounding residues and other parts of inhibitor make numerous van der Waals contacts and hydrogen bonds with proteinase. Similar number of contacting residues - about 10-12 on inhibitor side and 20-25 on enzyme side characterizes all inhibitor-proteinase complexes. Typical intermolecular contact area is about 600-900 \AA^2 from each protein. Hydrogen bonds and electrostatic interactions at the interface are well developed. Of particular importance is a short antiparallel β -sheet formed by main chain-main chain hydrogen bonds between P3 and P1 residues and 214-216 segment of the enzyme. Other very important features are: short 2.7 \AA contact between P1 carbonyl carbon and catalytic Ser195 O γ and two hydrogen bonds formed between carbonyl oxygen of P1 and Gly193/Ser195 amides (the oxyanion binding hole). All these above mentioned hydrogen bonds and shape complementarity of interacting areas ensure very similar recognition of different proteinases and inhibitors. The enzyme-inhibitor interaction is particularly rigid and resembles lock-and-key model. High resolution NMR solution structure of free inhibitor is very similar to the complexed inhibitor structure (Fig. 1D) (Holak et al., 1989). The major difference is the dynamics of the canonical loop which is poorly defined in the free state, but becomes particularly rigid when placed at the enzyme-inhibitor interface.

Docking of the P1 side chain inside the S1 binding pocket of the enzyme plays a major role in the energetics of the recognition. To better understand structural and energetic features of this interaction we are currently in progress in collaboration with Prof. Arne Smalas group at the University of Tromsø (Norway) to determine 10 high resolution X-ray structures of the complexes between trypsin and CMTI I/BPTI variants with different amino acid residues at P1 position.

The effect of the P1 position on the interaction with proteinases. Large part of enzyme-inhibitor contacts is made just by the P1 residue side chain, which penetrates deeply into the S1 specificity binding pocket of the proteinase. In case of BPTI-trypsin, interactions from Lys15 (P1) are of almost importance: substitution of the P1 Lys with Gly removes 70% of the total association energy leading to a huge 9 orders of magnitude decrease in association constant. Therefore, to understand how various P1 side chains influence energetics of the interaction, it was important to probe the properties of the S1 pocket. Table 1 presents association constant data which were determined in our laboratory with four different proteinases: bovine α -chymotrypsin, bovine trypsin, human neutrophil elastase (HNE) and human cathepsin G. The study is not yet finished, however, what can be easily recognized even at present stage is the huge dynamic range of K_a values ranging from about 1×10^9 for trypsin to about 5×10^5 for chymotrypsin and elastase. The dynamic range for cathepsin G is much smaller due to low association constants even for optimal Lys and Phe side chains.

Several conclusions about the S1 specificity can be drawn from presented data. Trypsin, due to unique location and orientation of Asp189 at the bottom of narrow and deep pocket, is particularly well studied to interact with basic side chains of Lys and Arg. Other side chains bind much weaker. Phe, which is the next strongest residue, binds over 1×10^5 -fold weaker than Lys. Interaction of hydrophobic aromatic ring of Phe with the hydrophobic walls of the S1 pocket partially compensates for the lack of ion pair. This is also evident in the case of Leu and Trp. Also polar side chains of Ser, Asn, His and Gln bind relatively well, probably due to weak, water-mediated hydrogen bonding interactions with Asp189. β -branched side chains of Ile, Val and Thr bind particularly poorly in the S1 pocket, due to steric clashes with the atoms at the narrow entrance to the pocket. As expected, negatively charged Asp binds 350-fold weaker than isosteric

Asn. Still, rather surprisingly, Asp binds 5-fold better than Gly. Clearly, from energetic point of view, empty pocket produces most deleterious effect.

Table 1. Association constants for the interaction between P1 mutants of BPTI and four serine proteinases in 0.1 M Tris, 20 mM CaCl₂, 0.05% Triton X-100, pH 8.3. HNE buffer additionally contained 0.5 M NaCl.

P1 variant of BPTI	β -trypsin	α -chymotrypsin	HNE	cathepsin G
	Ka (M ⁻¹)	Ka (M ⁻¹)	Ka (M ⁻¹)	Ka (M ⁻¹)
Lys15	1.6 10 ¹³	8.77 10 ⁷	7.80 10 ⁴	1.43 10 ⁶
Arg15	~2.0 10 ¹³	2.53 10 ⁸	1.36 10 ⁵	9.52 10 ⁴
Phe15	1.18 10 ⁸	2.55 10 ⁹	1.06 10 ⁵	1.10 10 ⁶
Ser15	3.9 10 ⁷	2.75 10 ⁵	6.90 10 ⁵	-
Asn15	2.23 10 ⁷	9.17 10 ⁶	1.30 10 ⁵	-
Trp15	7.57 10 ⁶	5.58 10 ⁹	1.10 10 ⁴	-
His15	6.24 10 ⁶	7.37 10 ⁷	3.70 10 ⁴	-
Leu15	5.41 10 ⁶	1.30 10 ⁹	5.71 10 ⁸	1.61 10 ⁵
Gln15	2.34 10 ⁶	6.15 10 ⁶	6.95 10 ⁵	-
Ala15 ^a	7.14 10 ⁵	3.03 10 ⁶	-	-
Thr15	2.94 10 ⁵	2.49 10 ⁶	1.17 10 ⁸	-
Ile15	1.15 10 ⁵	5.87 10 ⁵	1.57 10 ⁹	<5 10 ³
Asp15	6.32 10 ⁴	1.22 10 ⁴	1.05 10 ⁴	-
Val15	4.14 10 ⁴	2.27 10 ⁶	6.22 10 ⁹	<5 10 ³
Gly15	1.47 10 ⁴	7.90 10 ⁴	4.61 10 ⁵	-

^afrom M.J.M. Castro & S. Anderson, *Biochemistry* (1996) **35**, 11435-11443.

Interactions within large and hydrophobic pocket of chymotrypsin are, perhaps, easier to analyze. Energetic effects expressed in $\Delta\Delta G_{\text{ass}}$ (association free energy change referred to Gly) correlate well with the volume of P1 residue side chain (correlation coefficient $r=0.84$) or with its area ($r=0.83$). This suggests that hydrophobic effect is the driving force for the association reaction. Deleterious effect can be easily noticed in case of polar, negatively charged (750-fold effect for Asp-Asn comparison) and β -branched side chains. Side chains of His, Lys and Arg bind surprisingly good, which is explained by a very recent structure of BPTI-chymotrypsin complex (Scheidig et al., 1997). Accordingly, Lys 15 (P1) side chain is in rather unusual conformation and its charged end makes favourable interactions with carbonyl groups at the entrance to the pocket. As in the case of trypsin, empty pocket (Gly) is bad.

The major difference which features the pocket of neutrophil elastase is a clear preference for β -branched side chains of Val and Ile (Ile is favoured 3-fold over Leu, in case of chymotrypsin Leu is favoured 2210-fold over Ile). In principle all other side chains bind much (10^4 - 10^5 -fold) weaker. The data are in agreement with structural organization of the pocket which exhibits significant degree of flexibility (Bode et al., 1989b).

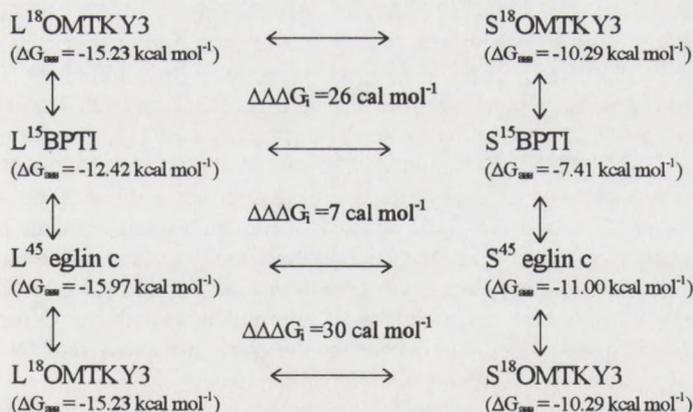
In summary, different P1 side chains introduced at P1 position are valuable tool to probe specificity of the S1 pockets of different proteinases. Our binding data correlate well with substrate specificity indexes for these proteinases expressed in the form of $\log(k_{\text{cat}}/K_m)$. Total pairwise comparison of the binding properties of the S1 pockets can be presented in the matrix form of correlation coefficients (Table 2). Generally, the values are significantly below unity, indicating severe differences among individual pockets.

Table 2. Pairwise correlation of the standard free energy change on association of P1 mutants of BPTI with various enzymes.

β -trypsin	1			
α -chymotrypsin	-0.12577	1		
HNE	-0.15234	-0.1519	1	
cathepsin G	-0.01379	0.107067	-0.53247	1
	β -trypsin	α -chymotrypsin	HNE	cathepsin G

Interscaffolding additivity. OMTKY3 and eglin c are inhibitors similar in size to BPTI and bind to serine proteinases via similar canonical binding loop, yet tertiary structure of these three inhibitors are completely different. Association constants for the interaction of 20 P1 variants of OMTKY3 and 7 P1 variants of eglin C with chymotrypsin and HNE are available (Lu et al., 1997). Extensive comparisons of BPTI, OMTKY3 and eglin c P1 data sets should answer the question, whether introduction of the same mutations in P1 position of binding loops produces similar energetic effects. In these three inhibitor structures P1 side chain is fully exposed to the solvent and makes no interactions with the remaining part of inhibitors. Moreover, χ_1 angle of the P1 side chain adopts virtually identical g^+ (-60°) angle upon complex formation with a proteinase in all crystallographically studied cases. We feel that this test, called the interscaffolding additivity, is a strong test for the validation of protein engineering approaches based on comparative modeling.

Figure 2. Interscaffolding additivity cycle comprising the effect of Leu \Rightarrow Ser substitution at position P1 of three inhibitors (BPTI, OMTKY3 and eglin c) on the interaction with chymotrypsin.



Additivity was calculated using the following equations:

$$\Delta\Delta G_{\text{ass}}(\text{L15} \rightarrow \text{S15})_{\text{BPTI}} = \Delta G_{\text{ass}}(\text{K15L})_{\text{BPTI}} - \Delta G_{\text{ass}}(\text{K15S})_{\text{BPTI}}$$

$$\Delta\Delta G_{\text{ass}}(\text{L45} \rightarrow \text{S45})_{\text{eglin c}} = \Delta G_{\text{ass}}(\text{L45})_{\text{eglin c}} - \Delta G_{\text{ass}}(\text{L45S})_{\text{eglin c}}$$

$$\Delta\Delta G_{\text{ass}}(\text{L18} \rightarrow \text{S18})_{\text{OMTKY3}} = \Delta G_{\text{ass}}(\text{L18})_{\text{OMTKY3}} - \Delta G_{\text{ass}}(\text{L18S})_{\text{OMTKY3}}$$

$$\Delta\Delta\Delta G_i = \Delta\Delta G_{\text{ass}}(\text{L15} \rightarrow \text{S15})_{\text{BPTI}} - \Delta\Delta G_{\text{ass}}(\text{L45} \rightarrow \text{S45})_{\text{eglin c}}$$

$$\Delta\Delta\Delta G_i = \Delta\Delta G_{\text{ass}}(\text{L15} \rightarrow \text{S15})_{\text{BPTI}} - \Delta\Delta G_{\text{ass}}(\text{L18} \rightarrow \text{S18})_{\text{OMTKY3}}$$

$\Delta\Delta G_{\text{ass}}(\text{L15} \rightarrow \text{S15})_{\text{BPTI}}$ - difference in free energy change between variants K15L and K15S M52L BPTI on their interaction with chymotrypsin.

$\Delta\Delta\Delta G_i$ - non-additivity factor.

The graphical presentation and mathematical equations used for data analysis are presented in Figure 2. Here the effect of a particular Leu \Rightarrow Ser substitution at P1 position (residue

18 in OMTKY3, 15 in BPTI and 45 in eglin c) is checked for its interscaffolding additivity. Exact additivity means that the energetic effect of the substitution is the same in all three thermodynamic cycles presented, and non-additivity factor ($\Delta\Delta\Delta G_I$) equals zero. This occurs very seldom and practically in all studied cases non-additivity term appears as a result of either experimental error or as an effect of real non-additivity. We assume that for $\Delta\Delta\Delta G_I$ in the range: -800 cal/mol to 800cal/mol interscaffolding cycle is additive.

Table 3. Interscaffolding additivity involving P1 substitutions in BPTI, OMTKY3 and eglin c.

$\Delta\Delta\Delta G_I$	BPTI - OMTKY3 (α -chymotrypsin)		BPTI - OMTKY3 (HNE)		BPTI -eglin c (α -chymotrypsin)		BPTI -eglin c (HNE)	
	No.	%	No.	%	No.	%	No.	%
< -1600	9	9.89	1	1.10	0	0.00	0	0.00
-1599 to -1200	6	6.59	1	1.10	0	0.00	0	0.00
-1199 to -805	9	9.89	8	8.79	0	0.00	1	10.00
-805 to -400	12	13.19	7	7.69	1	10.00	0	0.00
-399 to 0	11	12.09	8	8.79	6	60.00	0	0.00
1 to 400	9	9.89	14	15.38	3	30.00	1	10.00
401 to 805	9	9.89	10	10.99	0	0.00	0	0.00
806 to 1200	3	3.30	17	18.68	0	0.00	2	20.00
1201 to 1600	4	4.40	7	7.69	0	0.00	2	20.00
> 1601	19	20.88	18	19.78	0	0.00	4	40.00
total:	91	100	91	100	10	100	10	100

Table 3 combines all currently possible comparisons between three inhibitors. There are 91 possible single amino acid substitution comparisons between BPTI and OMTKY3 in their interaction with chymotrypsin and HNE and 10 comparisons between BPTI and eglin c in their interaction with chymotrypsin and HNE. Substitutions of the same amino acid residues at position P1 produces generally interscaffolding additivity effects with chymotrypsin: 41 of 90 i.e. 45.6% of cycles are additive. Similarly 43% of cycles fulfil additivity assumption when the same inhibitors are compared in their interaction with HNE. Introduction of the same substitutions into P1 position of BPTI and eglin C in their interaction with chymotrypsin are fully additive. However, for the interaction of these inhibitors with HNE only 1 out of 10 substitutions is additive. Generally, the presented data show that in many cases introduction of the same mutations into P1 position produces very similar energetic effect in different inhibitor structures.

Specific inhibitors. Wild types of CMTI I and BPTI inhibitors are strong inhibitors of some serine proteinases. The respective values of association constants are given in Table 4. One of our goals was to convert them to even stronger inhibitors using site directed mutagenesis, molecular modeling and literature data on substrate hydrolysis kinetics. Table 4 presents some of our improvements.

The strongest effects were obtained in case of HNE. Introduction of Val or Ile at position P1 of BPTI yielded inhibitor with K_a 79750- and 20100-fold higher than wild type inhibitor. In case of other proteinases: factors Xa and XIIa, cathepsin G and chymotrypsin the increase of association constant was lower even after introduction of multiple substitutions. The mutations were almost always introduced at the positions which contact the enzyme.

Table 4. The strongest variants of CMTI I and BPTI inhibitors designed to interact with several serine proteinases. The data in italics are for native inhibitors and the number in bold is the variant/wild type ratio.

Mutation(s)	human factor X_a	α -chymotrypsin	cathepsin G	HNE	human factor XII_a
CMTI I_{wt}	4.08×10^3	4.04×10^4	9.05×10^7		5.3×10^6
Ile6Val, Met8Leu	6.14×10^4 <i>4.08×10^3</i> 15				
Pro4Gly, Met8Leu	2.11×10^5 <i>4.08×10^3</i> 52				
Val2Ala,Pro4Thr, Arg5Leu, Ile6Asp, Met8Leu		8.07×10^5 <i>4.04×10^4</i> 20			
Arg5Leu, Met8Arg		1.10×10^6 <i>4.04×10^4</i> 27			
Ala18Gly, Met8Leu		1.85×10^5 <i>4.04×10^4</i> 4.6	9.45×10^8 <i>9.05×10^7</i> 10.4		
Glu9Lys					3.5×10^8
BPTI_{wt}		8.77×10^7	1.43×10^6	7.81×10^4	
Lys15Phe		2.55×10^9 <i>8.77×10^7</i> 29		1.06×10^5 <i>7.81×10^4</i> 1.36	
Lys15Trp		5.58×10^9 <i>8.77×10^7</i> 63			
Lys15Ile				1.57×10^9 <i>7.81×10^4</i> 20102	
Lys15Val				6.22×10^9 <i>7.81×10^4</i> 79750	

Stability of P1 variants of BPTI. We have already described the huge P1 residue effects on association energy with different serine proteinases. Has this exposed residue any effect on the thermodynamic stability of inhibitor? To answer this question we determined stability parameters ΔH_{den} , ΔS_{den} , ΔG_{den} , $\Delta C_{p,den}$ and T_{den} for a set of 10 P1 mutants of BPTI using highly sensitive differential scanning calorimeter (DSC). BPTI denatures according to a simple two-state model, what can be easily recognized from the $\Delta H_{cal}/\Delta H_{vH}$ (ratio of calorimetric to van't Hoff enthalpies)(Table 6). The denaturation parameters data are presented in Table 6 and typical

denaturation plot is shown on Figure 5. It can be noticed that BPTI mutants differ significantly in their stability, as can be judged from difference in their T_{den} values. The range of T_{den} values is 11.2° C at pH 2.0. Data points for all mutants at three different pH values are on the single line of ΔH_{vH} versus T_{den} plot, what indicates that all mutants have very similar $\Delta C_{p,den}$ values. In other words, difference in T_{den} values reflects real differences in their stabilities.

The interpretation of differences in stabilities is not straightforward. It is generally accepted that residues which are buried significantly influence protein stability. This effect on protein stability is due to breaking of many van der Waals and hydrogen bonding interactions together with transfer of non-polar buried area from protein interior to contact with solvent which accompany protein denaturation. Therefore, it is usually anticipated that solvent exposed side chains (as in the case of P1 residue) will not influence protein stability. As shown on Figure 5 and in Table 6 this is not the case here. Difference in stability correlate with different properties of amino acids. The highest correlations were found for: amino acid hydrophobicity (anticorrelation), β -turn propensity and P1 stability effect determined for variants of OMTKY3. The most probable interpretation is that different amino acids to different extent influence the main-chain entropy of the P1 residue which is in very similar (turn-like torsion angles) in BPTI and OMTKY3. Another

explanation is that hydrophobic side chain main weak interaction in denatured state of inhibitor leading to its apparent stabilization and (therefore) destabilization of the native state.

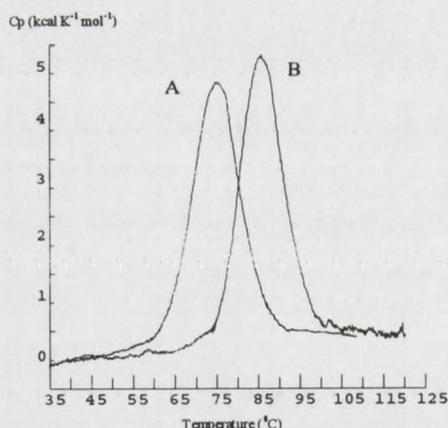


Figure 5. Denaturation curve recorded on Nano Differential Scanning Calorimeter (CSC Corp., USA) of two P1 mutants of BPTI in 10mM glycine buffer, pH 2.0. A 189 µg Lys15Trp BPTI, B - 193 µg Lys15His BPTI.

Table 6. Thermodynamic denaturation parameters for temperature unfolding of P1 variants of BPTI determined in 10 mM glycine-HCl buffer, pH 2.0.

P1 variant of BPTI	T_{den} (°C)	ΔG_{den}^a (kcal mol ⁻¹)	ΔH_{cal} (kcal mol ⁻¹)	ΔH_{vH} (kcal mol ⁻¹)	$\Delta H_{cal}/\Delta H_{vH}$	ΔS_{den} (kcal K ⁻¹ mol ⁻¹)	$\Delta C_{p,den}$ (kcal K ⁻¹ mol ⁻¹)
Phe	77.2	-0.372	65.24	67.61	0.96	0.19	0.24
Gly	81.5	0.321	67.75	68.25	0.99	0.19	0.24
His	86.7	1.401	69.90	69.28	1	0.19	0.33
Ile	76.5	-0.583	65.12	67.61	0.96	0.19	0.36
Arg	83.6	0.746	68.44	67.85	1	0.19	0.31
Ser	81.4	0.365	68.49	69.71	0.98	0.19	0.45
Val	77.9	-0.414	66.91	68.56	0.97	0.19	0.29
Asp	82	0.512	67.85	70.71	0.96	0.19	0.33
Asn	82.3	0.464	67.99	70.47	0.96	0.19	0.36
Trp	75.6	-0.755	65.12	65.84	0.99	0.19	0.22

^a ΔG_{den} determined at 80°C. Values of other thermodynamic parameters are given at T_{den} .

Phage display. The phage display strategy enables the presentation of large libraries of peptides or small proteins on the surface of phage particles followed by rapid selection of molecules of desired activity (Barbas & Burton, 1996). The great advantage of this method is a direct linkage between an observed phenotype and encapsulated genotype, which allows for fast determination of selected sequences. To address the problem of efficiency of selection procedures we constructed a (very) small library of BPTI fully randomized only at P1 position. The genetic system which we applied for monovalent presentation is based on phagemid pComb3H containing BPTI gene fused to phage coat protein III. The selection was performed on five serine proteinases (trypsin, chymotrypsin, pancreatic and neutrophil elastases, azurocidin) for which high quality thermodynamic data are available on binding of all 20 coded P1 amino acids (see above). We also confirmed an expected monovalent display of BPTI with 1.25 inhibitor molecule per phage particle by comparing the inhibitor concentration with the phage titer. The library was designed to contain all 20 P1 mutants of BPTI. Library construction was performed by double PCR. Diversity of the pool was verified by DNA sequencing of randomized pComb3H+BPTI that showed equimolar distribution of introduced nucleotides with trace of wild-type AAA lysine codon. Sequencing of individual clones from population of the initial library and pools after selection on the target proteins (together over 200 clones) proved that all amino acid residues except for proline were represented. The absence of proline at the P1 position of sequenced clones is probably caused by incorrect folding of this mutant. The screening of the library involved up to 3 rounds of enrichment on a target protein immobilized on Sepharose 4B. After each cycle of enrichment approximately 15 individual clones were sequenced to monitor the changes in amino acid occurrence at P1 position.

Each of the chosen target enzymes trypsin, chymotrypsin, PPE and HNE has different binding pocket and prefers substrate/inhibitor P1 residue of clearly different character. Because the selection was performed under equilibrium conditions, it should proceed in the direction determined by differences in association constant values for all P1 variants and a given proteinase. The amino acids found at the randomized P1 position after each round of selection on different targets are summarized in Table 7. The results show that each enrichment proceeds in the expected direction. In the case of trypsin, the first round resulted in the selection of both positively charged amino acids at P1 position. P1 Lys and Arg BPTI variants are known to interact with trypsin at least 10^6 -fold stronger, compared to other P1 mutants. The second round of enrichment led to the consensus of Lys. Here, lack of Arg is unexpected, although 3-fold stronger binding of Lys compared to Arg side chain to bovine trypsin has been reported (Otlewski & Zbyryt, 1994). For PPE we expected the selection of aliphatic P1 residues. The first round of enrichment did not give any observed selective pressure, while the second and the subsequent third one resulted in the selection of such residues as leucine, threonine with some appearance of alanine, valine and isoleucine.

Compared to trypsin, chymotrypsin does not exhibit as strong preference for a particular type of amino acid chain. In agreement, after 2 rounds of selection 5 different amino acids (Trp, Leu, Met, Lys, and Gln) have been observed. All of them bind strongly to chymotrypsin. Perhaps a bit surprising is lack of Phe and Tyr, which were present after first round of selection. We suppose that in case of enzymes with broader specificity like chymotrypsin, all amino acids with strong or moderate binding can be selected. This is further confirmed in case of PPE. Here, after 3 rounds of selection 5 amino acids could be found (Leu, Thr, Ala, Ile, Gln) all but Gln with strongest affinity for PPE. Again, no amino acids which bind weakly to PPE were found. However, compared to trypsin or chymotrypsin, first round of enrichment on PPE was very poorly selecting and many fairly non-optimal side chains for small, hydrophobic and rigid pocket of elastase (Tyr, Lys, Arg, Trp, Phe) have been found.

Next, sensitivity of the selection was tested using screening on HNE. We hoped to see differences in selection, since, compared to PPE, HNE prefers β -branched side chains of Ile and

Val. The consensus sequences have been already obtained after two selection rounds and included Val, Ile and Leu which are the three amino acids exhibiting strongest binding to the enzyme (see above). The results obtained for HNE highly differed from those for PPE and proved that the system nicely differentiates between even very similar enzymes. The positive results on 4 proteinases encouraged us to probe the specificity of azurocidin, which is known to bind wild type (P1 Lys) of BPTI. Due to lack of enzymatic activity, determination of its specificity using substrate mapping is impossible. Two round selection of P1 BPTI library showed its strong affinity for P1 Lys but also for moderately large uncharged P1 amino acids (Leu, Thr, Met, Gln). In case of chymotrypsin and azurocidin Lys was dominating P1 amino acid after final round of selection. To verify these results, for both proteins we performed the selection starting from the library after one cycle of enrichment on PPE. This pool of phages seems to represent all amino acids quite well, but has clearly reduced level of amino acids, which exhibit very low affinity for PPE, including Lys. After one (azurocidin) or two (chymotrypsin) additional rounds of selection the results obtained after direct selection could be confirmed. In case of azurocidin the exactly the same amino acids occur with rather similar frequencies. In case of chymotrypsin similar situation can be noticed, however, reverse of the Lys to Leu ratio after preselection on PPE is clear.

Table 7. P1 residues found in the P1 residue library of BPTI variants from sequencing of individual clones after selection on respective target proteins.

Target proteinase	round	observed amino acids at P1 residue
bovine trypsin	1	10Lys, 3Arg, 1Cys, 1Gln, 1Leu
	2	10Lys
porcine pancreatic elastase	1	2Phe, 2Arg, 1Trp, 1Thr, 1Tyr, 1Ser, 1Leu, 1Met, 1Lys
	2	8Leu, 2Thr, 1Val, 1Ile
	3	4Leu, 3Thr, 1Ala, 1Ile, 1Gln
human leukocyte elastase	1	5Leu, 4Ile, 3Val, 2Arg, 1Met, 1Thr, 1Tyr
	2	7Val, 7Ile, 3Leu
chymotrypsin	1	11Lys, 2Leu, 2Gln, 2Tyr, 1His, 1Phe, 1Met, 1Arg
	2	14 Lys, 3Trp, 2Gln, 1Met, 1Leu
	2*	19Leu, 6Lys, 3Trp, 3Gln, 2Met
azurocidin	1	9Lys, 4Leu, 2Thr, 1Gln
	2	10Lys, 1Met, 1Leu, 1Gln
	1*	6Lys, 5Leu, 1Met, 1Thr, 1Gln

Rounds denoted as 1*, 2* are the screening performed on given target with the pool obtained after one round of enrichment on porcine pancreatic elastase.

Acknowledgements. Supported by KBN grant 6 PO4B 002 10.

References

- Barbas, C. F. & Burton, D. R. (1996) *TiBTech* **14**, 230-234.
- Bardi, J.S., Luque, I. & Freire, E. (1997) *Biochemistry* **36**, 6588-6596
- Bode, W., Greyling, H.J., Huber, R., Otlewski, J. & Wilusz, T. (1989a) *FEBS Lett.* **242**, 285-292.
- Bode, W., Meyer, E. Jr. & Powers, J.C. (1989b) *Biochemistry* **28**, 1951-1963
- Castro, M.J.M. & Anderson, S. (1996) *Biochemistry* **35**, 11435-11446
- Holak, T.A., Bode, W., Huber, R., Otlewski, J. & Wilusz, T. (1989) *J. Mol. Biol.* **210**, 649-654
- Jones, S. & Thomson, J.M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13-20.
- Laskowski, M., Jr. & Kato, I. (1980) *Annu. Rev. Biochem.* **49**, 593-626.
- Lu, W., Apostol, I., Qasim, M.A., Wame, N., Wynn, R., Zhang, W.L., Anderson, S., Chiang, Y.W., Rothberg, L., Ryan, K. & Laskowski, M. Jr. (1997) *J. Mol. Biol.* **266**, 441-461
- Otlewski, J. & Zbyryt, T. (1994) *Biochemistry* **33**, 200-207.
- Schechter, I. & Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157-162.
- Scheidig, A.J., Hynes, T.R., Pelletier, L.A., Wells, J.A. & Kossiakoff, A.A. (1997) *Protein Sci.* **6**, 1806-1824
- Wlodawer, A., Nachman, J., Gilliland, G.L., Gallagher, W. & Woodward, C. (1987) *J. Mol. Biol.* **198**, 469-480

DESIGNING OF PROTEIN PROTEINASE INHIBITORS IN THE ASPECT OF BOTH PROTECTION AND NUTRITIONAL IMPROVEMENT OF PLANTS.

Tadeusz Wilusz, Antoni Polanowski

Institute of Biochemistry and Molecular Biology, University of Wrocław, ul. Tamka 2, 50-137 Wrocław, e-mail: apolan@bf.uni.wroc.pl

Plasnts synthesize many toxic substances of which protein inhibitors of serine proteinases are considered to be the most important antinutritional factors. These properties of proteinase inhibitors are manifested by quenching the activity of proteolytic enzymes that leads to pathological changes and growth retardation. On the other hand, these inhibitors are believed to enhance resistance of plants to insect pests, therefore, research has been conducted on introducing foreign genes encoding desired inhibitors to improve this resistance. The inhibitors however, should meet special requirements not to impair the safety and quality of foods and feeds.

Most proteins in an organism are subjected to continuous degradation and synthesis. For example the healthy organism of human adult synthesizes about 400g of protein per day. Such a high rate of turnover of these fundamental compounds in animal organisms brings about the necessity for constant amino acid supply. The main source of proteins are plants.

It is well known that Nature has provided the plants with ability to synthesis of many biologically active compounds, some proteins including, which, when utilized in nutrition, may cause harmful effect on the animal and human organism [1].

Among these noxious substances one can find: a) lectins, b) tannins, c) saponnins and d) inhibitors of serine proteinases, which are considered to be the most important antinutritional factors found in plants.

The antinutritional properties of proteinase inhibitors, present in the foodstuffs are being manifested in different manners:

- a) as factors blocking or making digestion of proteins difficult, resulting in growth retardation due to lowered bioavailability of amino acids,.
- b) by quenching the activity of digestive proteolytic enzymes, the level of trypsin and chymotrypsin in duodenum decreases, what in turn, due to the feed-back mechanism increases their synthesis in pancreas; this results in pathological enlargement of the gland (hypertrophy).
- c) the constrained synthesis of required proteases, in deficiency or inaccessibility of exogenous proteins, is carried on at the expence of degradation of endogenous proteins as the only source of necessary amino acids, particularly sulfurous ones, used in an overproduction of enzymes.

Plant origin serine protease inhibitors are usualy proteins with molecular mass values in the range of 3 to 21 kDa and majority of them inhibits trypsin and/or chymotrypsin. Some inhibitors inhibit only one digestive enzyme, for example trypsin inhibitors from *Cucurbitaceae* family seeds (2).

Protein serine proteinase inhibitors can be divided into several families (4). The classification is based on their sequence homology, topology of the disulfide bridges, and the reactive site(s) assignment. Cucurbita family inhibitors, for the first time isolated and described in our institute (5), can be given as an example of such a family. These, the smallest among known protein inhibitors of serine proteinases, built only of about 30 amino acids residues and of precisely elucidated structure (6) belong to the most potent bovine trypsin inhibitors (2).

There are many contradictable data concerning the toxicity of proteinase inhibitors. Divergence is due to the fact that:

- a) misdefined types of inhibitors have been used in animal feeding experiments and
- b) lack of knowledge on the interaction of inhibitors with digestive proteases of various animal species used in conducted experiments.

Ad a) In soybeans, the main ingredient of foodstuffs, one can find two different types of serine proteinase inhibitors. One of them, Kunitz-type inhibitor (STI) built of 181 amino acid residues, inhibits mainly bovine trypsin and to a lesser extent also chymotrypsin. It contains only two disulfide bridges and has one reactive site identified as the Arg64-Ile65 peptide bond. It can be denatured in high temperatures and is susceptible to slow pepsin cleavage under acidic conditions (7). Second inhibitor, Bowman-Birk inhibitor (BBI) of 8 kDa, consists of 70 amino acid residues, with seven disulfide bridges. The inhibitor contains two reactive sites, one responsible for trypsin inhibition was identified as Lys16-Ser17 and the second, active against chymotrypsin, was assigned to the Leu43-Ser44 peptide bond. BBI, when compared to STI, is more active towards both trypsin and chymotrypsin and is more resistant to denaturing agents.

In over 95% of conducted research on nutrition, soybean inhibitors were used as model proteins. High cost of purchase or preparation of purified BBI and STI inhibitors tended researchers to employ samples of only partially purified inhibitors, so called 'inhibitor concentrates'. These samples, although exhibiting high antitrypsin activity, were in fact a mixture of unknown ratio of two inhibitors with different properties.

From the experience we have gained during setting up the methods for the soybean trypsin inhibitor separation, it is obvious that depending on extraction procedures (salt or acetone fractionation), the obtained final product, although of the same activity, contains different amounts of two inhibitors.

Ad b) Published data suggest that the same inhibitor exhibit different activity towards enzymes of different origin. For example, the basic pancreatic trypsin inhibitor that effectively inhibits bovine chymotrypsin reveals no activity towards human enzyme (9). The same inhibitor also works well on porcine pancreatic elastase but shows no activity against hen pancreatic elastase (10). Another example of the same sort comes from the comparison of inhibitory activities toward human and rat trypsin. The cationic human trypsin is much weaker inhibited by both BBI and hen ovomucoid than rat enzyme, whereas hen ovomucoid is equally active towards trypsin of either species (11). It has been found that one of human trypsin isoform (mezotrypsin) retains full activity in the presence of known trypsin inhibitors (12) e.g. Kunitz-type inhibitor

from soybean (STI), basic pancreatic trypsin inhibitor (BPTI) and even very specific for trypsin, pancreatic secretory trypsin inhibitor (PSTI).

It is important to point out that the antitrypsin activity of preparations used in feeding research concerns rather the cationic either bovine or porcine trypsin and obtained results are referred only to this form of enzyme. Most vertebrates however, synthesizes anionic forms of trypsin as well, which in fact, due to instability under acidic conditions, are very difficult to purify, and actually they are inaccessible on the market. It has been proved that anionic forms of trypsin exhibit different sensitivity to inhibitors than cationic ones.

Consecutive very important problem concerned with the inhibitors as noxious substances is connected with more and more commonly conducted research on transgenic plants. Higher plants due to „settled” manner of life, are very susceptible to phytopathogens, however, like animals, they are able to perceive the environmental stimuli and to respond to them with appropriate biochemical reactions. In response to wounding caused by herbivorous insect attack or phytopathogenic fungi, plants activate defence genes bringing about the synthesis and rapid accumulation of defensive proteins, among others the inhibitors of proteolytic enzymes. Proteinase inhibitors by inhibiting the activity of proteinases in insect guts can lead to protein malnutrition, reduced growth, and in some cases, death (13). A search for a wound-inducible signal molecules involved in the transcriptional activation of the defensive genes has resulted in the identification of an 18-amino-acid polypeptide, called systemin (15).

Proteinase inhibitors from plants are also consider as a novel class of fungicides. Some trypsin and chymotrypsin inhibitors are able to block the synthesis of chitin in cell wall, thus weakening the fungal hyphae. This is a result of the inhibition of a proteinase converting the chitin synthetase zymogen into an active form (14).

To enhance resistance to herbivorous insect pests or pathogenic fungi, introducing foreign genes encoding desirable proteinase inhibitors into many crop plants are now becoming quite common (16). However, there are limitations to the efficacy of these procedure. Surprisingly, it was found that insects, to overcome the proteinase inhibitors in their host plants, have developed adaptive strategies by synthesizing inhibitor resistant enzymes. Thus, it may be required to introduce not only one but rather several genes for multiple inhibitors, matched to the complement of enzymes in the insect's midgut. This however, may generate a real problem in reasonable use of transformed plants as foods and feeds. Transgenic plants, bearing genes of various inhibitors, are indeed more resistant to insect pests and phytopatogens but their nutrient properties certainly remain questionable.

In case of scale up of agronomically desirable transgenic plants unforeseen problems may come into sight. For instance, Malone et al. (17) have found that trypsin/chymotrypsin inhibitor added to the bee feed in amount of 1% appeared to be toxic for honey bees. One can expect of similar effect of high-inhibitory transgenic plants on other useful insects. Therefore, the goal of plant genetic engineering should not be limited only to the introduction into the crop plants any gene(s) encoding an inhibitor improving their resistance to insect pests. The inhibitor(s) chosen for this purpose

should be precisely design to meet very special demands in terms of specificity, selectivity and susceptibility to inactivation.

There are several possibilities nowadays to obtain proteolytic enzyme inhibitors of desired properties. One can take advantage of naturally occurring iso inhibitors and their homologs or use designed molecule of defined activity and specificity. In our research, conducted in cooperation with the groups of Professors M. Laskowski (Perdue University, U.S.A.), U. Ragnarsson (Uppsala University, Sweden) and Z. Kupryszewski (Gdańsk University, Poland), both approaches have been applied. From comparison of three inhibitors separated from the seeds of *Cucurbitaceae* family in our laboratory results that the substitution of a single amino acid residue in any iso inhibitor considerably affects its activity towards cognate enzymes (Fig. 1). For example, CMTI-I (*Cucurbita maxima* trypsin inhibitor) having glutamic acid residue in the position 9 inhibits factor XIIa (a proteinase initiating the blood coagulation cascade) over 60-fold weaker than inhibitor CMTI-III in which lysin residue occupies this position. On the other hand, the presence of arginin residue in position 5 of CMTI-III evokes almost a 1000 fold increase of its activity against kallikrein as compare to CPTI-II (*Cucurbita pepo* trypsin inhibitor) which has lysin in that position. It is worth noting that in spite of great differences in activity against factor XIIa or kallikrein all these three inhibitors are equally active towards bovine trypsin (18).

Fig. 1. Specificity of some *Cucurbitaceae* family trypsin iso inhibitors

		K_a (M^{-1})	
		XIIa	kal.
CMTI I	R V C P R I L M E C K K D S D C L A E C V C L E H G Y C G	5.3×10^6	
CMTI III	R V C P R I L M K C K K D S D C L A E C V C L E H G Y C G	3.3×10^8	1.3×10^6
CPTI II	R V C P K I L M K C K K D S D C L A E C I C L E H G Y C G		2.0×10^3

CMTI - *Cucurbita maxima* trypsin inhibitor, CPTI - *Cucurbita pepo* trypsin inhibitor, XIIa - factor XIIa, kal. - kallikrein

Employing either chemical synthesis or molecular techniques of mutagenesis it is possible to produce, in an efficient way, inhibitory mutants of desired activity. Replacement of Arg 5 at P_1 position of the reactive site in CMTI-III for Val totally abolishes the antitrypsin activity, however, such a mutant gains activity against human leukocyte elastase (HLE) (19). Replacement Arg5Ala yields good inhibitor of pancreatic elastase, and Arg5Phe, as could be expected, brings about the change of the inhibitory specificity from antitrypsin to antichymotrypsin. Though, the activity of the latter mutant against chymotrypsin is rather low ($K_a=10^6 M^{-1}$). Replacing the several amino acid residues in the reactive site loop of the inhibitor (Gly2, Thr4, Phe5, Glu6, Tyr7, Arg8,) enabled to obtaine one of the strongest known chymotrypsin inhibitor of $K_a=6 \times 10^{11} M^{-1}$ (Fig.2) (20).

Fig. 2. Synthetic analogues of *Cucurbita maxima* trypsin inhibitor (CMTI III).

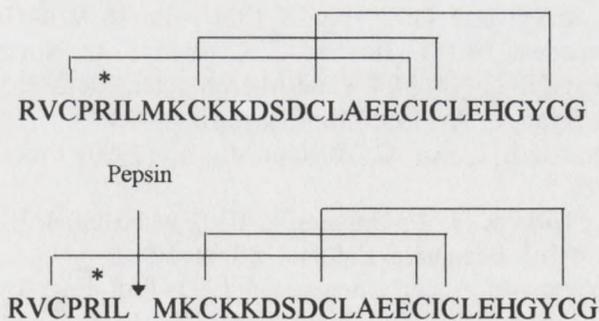
	*	$K_a (M^{-1})$
CMTI III	R V C P R I L M K C K K D S D C L A E C V C L E H G Y C G	6.8×10^{11}
HLEI	R V C P V I L M K C K K D S D C L A E C V C L E H G Y C G	6.8×10^{11}
ChTI	R V C P F I L M K C K K D S D C L A E C V C L E H G Y C G	8.0×10^6
ChTI	R G C T F E Y R K C K K D S D C L A E C V C L E H G Y C G	6.0×10^{11}

CMTI III - *Cucurbita maxima* trypsin inhibitor, HLEI - human leukocyte elastase inhibitor, ChTI - bovine chymotrypsin inhibitor, * - reactive site.

On the other hand, elongation of one of the squash inhibitors at C-terminal by a tetrapeptide: Pro-Tyr-Val-Gly allowed to produce a double-headed inhibitor of both the antitrypsin and anticarboxypeptidase A activities (21)

Producing transgenic plants of enhanced resistance to insect pests, by means of introducing the foreign genes encoding proteinase inhibitors of required activity and specificity, their harmlessness to animals and men should also be taken into consideration. One of the possibilities is designing inhibitors which are susceptible to pepsin digestion. Such inhibitors would be inactivated in stomach therefore, they would not be affected by trypsin, chymotrypsin and elastase activities. Squash inhibitors are again a good example (22), since under physiological conditions (pH 1.8), pepsin selectively hydrolyzes in these molecules a single peptide bond: Leu6-Met7 causing their inactivation (Fig. 3).

Fig. 3. Pepsin inactivation of CMTI III.



Another example of inhibitor being inactivated by pepsin is pancreatic secretory trypsin inhibitor (PSTI).

Consequently, inhibitors applicable in creation of transgenic plants resistant to insect pests should fulfill the following requirements: a) to be very strong towards digestive enzymes of defined insect, b) to be highly selective to proteinases of insect pests of given plant species; it was found, for example, that one of inhibitors isolated

from rice seeds appeared to be very effective against proteinases of *Meloidogyne hapla* whereas, enzymes of other two species of the same genera *M. incognita* and *M. javanica* were only slightly inhibited, and c) to be digestable by a proteinase(s) of animal and human origin.

References:

1. Grant, G., Dorward, P., Buchan, W. C., Armour, J. C. & Pusztai, A. (1995) *British J. Nutr.* **73**, 17-29.
2. Wiczorek, M., Otlewski, J., Cook, J., Parks, K., Leluk, J., Wilimowska-Pelc, A., Polanowski, A., Wilusz, T. & Laskowski, M., Jr. (1985) *Biochem. Biophys. Res. Commun.* **126**, 646-652.
3. Birk, Y. (1985) *Int. Pept. Prot. Res.* **25**, 113-131.
4. Laskowski, M., Jr (1986) In: *Nutritional and toxicological significance of enzyme inhibitors* (M. Friedman ed.) Plenum Publ. Corp. pp.1-17.
5. Wilusz, T., Wiczorek, M., Polanowski, A., Denton, A., Cook, J. & Laskowski, M., Jr (1983) *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 93-95.
6. Bode, W., Greyling, H. J., Huber, R., Otlewski, J. & Wilusz, T. (1989) *FEBS Lett.* **242**, 285-292.
7. Kassell, B., & Laskowski, M., Jr (1956) *J. Biol. Chem.* **219**, 203-219.
8. Kunitz, M. (1947) *J. Gen. Physiol.* **30**, 291-310.
9. Coan, M. H. and Travis, J. (1970) In: *Proc. Int. Res. Conf. Proteinase Inhibitors* (Fritz, H. & Tschesche, H. eds) de Gruyter, Berlin, N.Y. pp. 294-298.
10. Polanowski, A. and Wilusz, T. (1996) *Acta. Biochim. Polon.* **43**, 445-454.
11. Hanlon, M.H. and Liener, I. R. (1986) *Comp. Biochem. Physiol.* **84B**, 53-57.
12. Nyaruhucha, C. N. M., Kito, M. & Fukuoka, S. I. (1997) *J. Biol. Chem.* **272**, 10573-10578.
13. Mallory, P. A. & Travis, J. (1975) *Am. J. Clin. Nutr.* **28**, 923-930.
14. Lorito, M., Broadway, R. M., Haynes, C. K., Woo, S. L., Noviello, C., Williams, D. L. and Harman, G. E. (1994) *Mol. Plant-Microb. Interactions* **7**, 525-527.
15. Schaller, A. & Ryan, C. A. (1995) *BioEssays* **18**, 27-33.
16. Jonson, R., Narvaez, J., An, G. & Ryan, C. A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9871-9875.
17. Malone, L. A., Giacon, H. A., Burgess, E. P. J., Maxwell, J. E., Christeller, J. T. & Laing, W.A. (1995) *J. Economic. Entomol.* **88**, 46-50.
18. Wilusz, T., Polanowski, A. and Otlewski, J. (1995) *Prot. Eng.* **8**, 55.
19. Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Wilusz, T. & Polanowski, A. (1989) *Biol. Chem. Hoppe-Seyler* **370**, 499-502.
20. Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Krokoszynska, I. & Wilusz, T. (1991) *Biol. Chem. Hoppe-Seyler* **372**, 63-68.
21. Le-Nguyent, D., Matras, H., Coletti-Previero, M.-A. & Castro, B. (1989) *Biochem. Biophys. Res. Commun.* **162**, 1425-1430.
22. Otlewski, J., Dryjański M., Bułaj, G. & Wilusz, T. (1994) *Biochemistry* **33**, 208-213.

Heterologous gene expression system developed in the methylotrophic yeast, *Candida boidinii*, for useful enzyme production

Nobuo Kato and Yasuyoshi Sakai

Division of Applied Life Sciences, Graduate School of Agriculture,
Kyoto University

Kitashirakawa Sakyo-ku Kyoto 606-01, Japan
e-mail: nkato@kais-u.ac.jp

Abstract: The alcohol oxidase gene (*AOD1*) of the methylotrophic yeast, *Candida boidinii*, is efficiently expressed during the growth on methanol. Diverse heterologous genes (or cDNA) were expressed in *C. boidinii* under the control of the *AOD1* promoter. The gene expression system of *C. boidinii* was applicable to production of useful enzymes, without regard to the localization of the protein, for example, cytosolic *Saccharomyces cerevisiae* adenylate kinase, secretory *Rhizopus oryzae* glucoamylase, and peroxisomal *Penicillium janthinellum* fructosyl amino acid oxidase.

Since the first discovery of the methylotrophic eukaryote, *C. boidinii* [1], the methylotrophic yeast has been used extensively both in academic and applied fields. In 1980s, the overall pathway for methanol dissimilation and assimilation was elucidated by purifying each enzyme responsible for C₁-metabolism, and some of them were found to be compartmented into single-membrane bound organelles, peroxisomes. As methanol is a cheap carbon source, production of SCP (single cell protein) attracted much attention in the early 1970s, and these interests resulted in the establishment of a high-cell density cultivation method. Several processes have also been developed with *C. boidinii* for the production of useful chemicals through unique C₁ metabolic functions [2]. In 1980s, recombinant DNA technology expanded and invaded into this field of studies. A heterologous gene expression with a methylotrophic yeast now become a basic technology in molecular biology and for the production of pharmaceutical proteins. And in the last decade, the mechanism of peroxisome assembly and protein sorting has begun emerge at molecular level using these methylotrophic yeasts as model organisms (Subramani). These studies have medical importance in connection to the human peroxisomal genetic disorder (e.g. Zellweger syndrome).

Here we summarize our recent studies with the methylotrophic yeast, *C. boidinii*, based on a concept "application of cellular functions to produce useful enzymes". These cellular functions include 1) metabolism, 2) gene expression, and 3) posttranslational processes, e.g. organelle assembly and protein translocation.

Methanol metabolism of yeasts

Methanol is oxidized to formaldehyde by alcohol oxidase, and hydrogen peroxide formed in this reaction is scavenged by catalase. Formaldehyde is fixed with xylulose 5-phosphate by dihydroxyacetone synthase, leading to form glyceraldehyde 3-phosphate and dihydroxyacetone, which is phosphorylated by

dihydroxyacetone kinase to dihydroxyacetone phosphate. Both glyceraldehyde 3-phosphate and dihydroxyacetone phosphate are condensed to fructose 1,6-bisphosphate by an aldolase, then introduced to synthesis of the cell-constituents. The methanol oxidation and formaldehyde fixation take place in the peroxisome. Further oxidation of formaldehyde to CO₂ occurred in the cytosol, where two steps of NAD-linked dehydrogenations catalyzed by formaldehyde and formate dehydrogenases provide energy for the growth.

Gene expression system using *C. boidinii* AOD1 promoter

The yeast gene expression system of *C. boidinii* is considered advantageous in the following respects:

- 1) The yeast is easy to handle and to scale up to an industrial scale.
- 2) Eucaryotic intracellular structures of the yeast are suitable for folding, modification and secretion of proteins from higher organisms.
- 3) Methylophilic yeasts have strong promoters such as promoter of alcohol oxidase gene (*AOD1*), and this permits a high-level expression from a stable expression cassette on chromosome.
- 4) *C. boidinii* is well suited for high-cell density culture.
- 5) As methanol is a cheap substrate, low cost for cultivation and induction can be achieved.
- 6) When glycerol and methanol are used for growing and inducing substrates, one step fermentation can be performed.
- 7) One step gene disruption system can easily be constructed a desirable mutant for protein production.
- 8) Peroxisome is greatly induced with methanol in *C. boidinii*. Peroxisome is an attractive organelle "to pack" a heterologous proteins produced. In the organelle, proteins are protected from the degradation by cytosolic protease.

There are some demerits of this expression system; such as the deference in glycosylation of secretory proteins between the yeast and animal cells, and no information on classical genetics have been accumulated.

An expression vector pNoteI for *C. boidinii*, which was constructed from pUC18, has the *URA3* gene as the selection marker, and the *Not I* site between the *C. boidinii* *AOD1*-promoter and terminator sequences [3]. This vector was designed to be integrated into the *URA3* locus at a unique restriction site on the *C. boidinii* *URA3* gene. A desirable gene is introduced into *Not I* site of pNoteI by blunt-end ligation, and is efficiently expressed during the growth on methanol but repressed on glucose or ethanol. Heterologous gene expression in *C. boidinii* used the *AOD1*-promoter was investigated. Our system is the third example of the expression system under the promoter of alcohol oxidase gene, the precedents being the systems of *Pichia pastoris* and *Hansenula polymorpha*. Among them, our system is characterized by the high cell-density cultivation through simultaneous use of two carbon sources, glycerol and methanol. Two examples of our results of gene expression are shown as follows.

1) *Saccharomyces cerevisiae* adenylate kinase[3,4]: Adenylate kinase gene was expressed in *C. boidinii*. Methanol-induced transformants has 10,000-fold enhanced levels of adenylate kinase activity which was found in the cytosol, and produced 23-fold more ATP from adenosine when compared to the parent strain of *C. boidinii*. In a pH controlled reaction system with successive adenosine-feeding, the ATP concentration in the reaction mixture reached 117 g/liter over 45 h.

2) *Rhizopus oryzae* glucoamylase [5]: Using *R. oryzae* glucoamylase (Gluc1) as a target protein for industrial production, we studies this issue using *C. boidinii* as

an expression host. cDNA of Gluc1 was inserted into the *Not* I site of pNoteI. A transformant integrated with a single-copy expression cassette to the chromosome produced the enzyme into the medium to a high amount when the cells were grown on methanol plus glycerol as carbon source. The transformant *C. boidinii* cells were grown up to ca. 95 g dry cell weight/liter medium, and the concentration of glucoamylase in the medium reached 3.4 g/liter. This showed that the signal sequence from the enzyme functioned very effectively in *C. boidinii*. The production of *C. boidinii* was efficiently achieved with glycerol and methanol as carbon sources where both growth and expression were rapid and strong as is the case for *H. polymorpha*. This enabled us to minimize the cultivation time required to achieve high cell density to attain a high yield. In *P. pastoris*, glycerol or glucose could not be added to the methanol-induced expression. *R. oryzae* Gluc1 cDNA was expressed in *S. cerevisiae* only at ca 300 mg/liter medium. Since the third letter in the preferred codon on *C. boidinii* is A or T. *R. oryzae* Gluc1 has AT-preferred codon usage, *C. boidinii* may be the more suitable host for the production of this enzyme. To our knowledge, this productivity is of the highest value among previously reported secretory enzyme production systems using the yeast system as an expression host.

Posttranslational processes involved in the expression of peroxisomal enzyme activity

Recently, molecular bases for peroxisomal protein transport began to appear with the use of yeast genetics. Methylotrophic yeasts, *P. pastoris*, *H. polymorpha*, and *C. boidinii*, are currently used as model systems in these peroxisomal studies because of their ability to proliferate massive peroxisomes. Since peroxisomes of *C. boidinii* can metabolize fatty acids, D-amino acids as well as methanol, proliferation of peroxisomes in the yeast are easily switched on and off by carbon sources. Therefore, *C. boidinii* should support an understanding of the mechanism of peroxisome assembly at the molecular level, an essential process for the metabolism of various substrates [6-8].

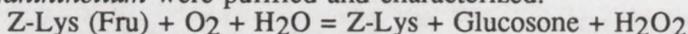
Intracellular production of enzyme protein is the basic technology to prepare biocatalytic cells for production of useful chemicals and to produce useful enzyme at large scale. Under highly methanol-induced conditions, much portion of intracellular volume (up to 80%) is occupied by peroxisomes. From a biotechnological point of view, peroxisome is an attractive organelle "to pack" the produced heterologous proteins where they are protected from the degradation by cytosolic proteases. Most peroxisome matrix proteins are destined for peroxisome by a 3-amino acid sequence, -SKL and its derivatives (Peroxisomal Targeting Signal 1: PTS1), located at an extreme carboxyl end. And so, an enzyme protein can be easily targeted to peroxisome just by adding 3 amino acids at C-terminus, when the addition does not affect its activity.

Characterization of FAOD found in diverse fungi

Reducing sugars such as glucose attach to the amino groups of proteins by forming Schiff's bases. Subsequently, the adduct undergoes Amadori rearrangement to form a stable ketoamine product. The non enzymatic reaction is called glycation, in order to distinguish it from the enzymatic glycosylation of proteins. In diabetic patients whose blood glucose levels are high, the glycation of blood proteins, e.g. hemoglobin and albumin, is enhanced. The amounts of these glycated proteins reflect the level of blood glucose in periods corresponding to the half-life of the protein (14-20 days for albumin and 1-3 months for hemoglobin). The prevalent methods used for the determination of glycated

proteins in serum are high-performance liquid chromatography and the fructosamine methods. The specificities of these methods are relatively low and/or the assays are somewhat tedious when examining a great number of samples. To overcome these demerits, we are developing a new enzyme method which is highly specific and easy to conduct. The most commonly glycosylated site of albumin is the ϵ -amino group of lysine residue, and that of hemoglobin A_{1c} is the N-terminal valine. Therefore, *N*^ε-fructosyl *N*^α-benzyloxycarbonyl-lysine [Z-Lys(Fru)] and *N*-fructosyl valine [Fru-Val] are taken to be model compounds of glycosylated albumin and glycosylated hemoglobin, respectively.

We have successfully used these model compounds for the screening of microorganisms which produced fructosyl amino acid oxidase (FAOD) [9-12]. FAODs from *Aspergillus terreus*, *Fusarium oxysporum*, *Gibberella fujikuroi*, and *Penicillium janthinellum* were purified and characterized.



The equation indicates the reaction catalyzed by FAOD, when [Z-Lys(Fru)] was used as the substrate. The enzyme activity is measured spectrophotometrically with peroxidase-coupled reaction system and also applicable to the determination of the amount of substrate added. The FAODs from *A. terreus*, *F. oxysporum* and *G. fujikuroi* showed high activities toward Z-Lys(Fru), and one from *P. janthinellum* exhibited high activities toward Fru-Val. The former enzymes are expected to be applicable in the enzymatic measurement of glycosylated albumin, and the latter for enzymatic measurement of hemoglobin A_{1c}. Glycosylated albumin could be determined by the use of *A. terreus* FAOD. Since the enzyme was unable to use whole glycosylated albumin, a protease-digestion was needed prior to application of FAOD reaction. A linear relationships between the concentration (or glycosylation rate) of glycosylated human serum albumin and the absorbance after the reaction with FAOD.

cDNA coding for FAODs were cloned from cDNA libraries of *A. terreus* and *P. janthinellum*. The coding region for both fungal FAODs consisted 1314 bp encoding 437 amino acids. Between the sequences of the two enzymes, 35% of the amino acid identity and 60% of the similarity were found. When the amino acid sequences were compared with those of other proteins using the FASTA and BLAST programs, there was no protein for which the amino acid sequence showed remarkable similarity to that of either of the FAODs. However, the sequence of a dinucleotide-binding motif, GXGXXG, was in the deduced N-terminal region and the regions close to the C-terminus were found to show a striking similarity to the amino acid sequences of several bacterial sarcosine oxidase. The region contained a cysteine residue, that was predicted to be the active site of sarcosine oxidase on site-directed mutagenesis. This same region of FAOD was expected to be involved in the enzyme reaction. Sarcosine oxidase catalyzes the oxidative cleavage of secondary alkyl amine linkage. Although neither of the FAODs was active toward sarcosine, the reaction catalyzed by FAOD was similar at the point of the oxidative cleavage at the C-N bond. The tripeptide sequence at C-terminus was SKL for *A. terreus* and AKL for *P. janthinellum*. Both of these sequences belong to PTS1 necessary for peroxisomal protein import. Both FAOD cDNAs were expressed in an active form in *E. coli*, and each recombinant FAOD had indistinguishable enzymatic and physicochemical properties (molecular mass, N-terminal amino acid sequence, and substrate specificity) from each fungal FAOD. The FAOD activities in the cell-free extract of each *E. coli* transformants was almost same level (0.16 units/mg) as in its original host of fungus. In order to use FAOD for diagnostic analysis of glycosylated proteins, the much higher productivity is required.

Therefore, we have attempted to express FAOD genes in *C. boidinii*. This becomes the first example of heterologous protein production, which is exclusively accumulated in the peroxisome.

High-level accumulation of *P. janthinellum* FAOD in peroxisome of *C. boidinii*

Localization of foreign proteins are able to be packed into peroxisomes of *C. boidinii* by the use of PTS1 pathway for protein targeting into the organelle, which is highly induced with methanol. As mentioned above, the peroxisome is still occupied with alcohol oxidase, whose prosthetic group is FAD. When FAOD is highly expressed, FAOD and AOD become to competed with each other for FAD. In order to prevent the FAD-deficient, we constructed *AOD1* disruptant (*AOD1Δ*) by the one-step gene disruption vector. The strain *AOD1Δ* impair the grown on methanol but not on other carbon sources. The ability for enzyme induction with methanol in *AOD1Δ* was same as that in the parent strain. From the nucleotide sequence, the third letter in the preferred codon of FAOD cDNA of *P. janthinellum* is found to be G or C, while that of *C. boidinii* is A or T. On the basis of the findings, DNA corresponding to the FAOD cDNA was synthesized by the used of AT-preferred codons, and inserted into the expression vector, pNotel.

When FAOD cDNA was used, the transformant of the parent strain, *C. boidinii* TK62, produced FAOD at the same level of the original fungus cells (0.79 units/mg protein). The productivity was increased 2.6- and 7.5-fold by the application of synthetic DNA and *AOD1Δ* as the host strain, respectively. Furthermore, in the transformant of *AOD1Δ* in which 5 copies of the expression cassette were integrated in the chromosome, FAOD activity was 30 times higher than that of the transformant (TK62) with the cDNA. Localization of FAOD expressed in each transformant was investigated by sucrose gradient fractionation of the cell-free extracts. In every transformant, FAOD was found in the peroxisome fraction.

Conclusion: The strong and inducible gene expression system of *C. boidinii* can be directly applied for the production of useful proteins. When the molecular breeding system is coupled with gene expression, the yeast can be considered a novel bioconversion system. In other words, intracellular overexpression means amplification of the reaction involved in bioconversion, and by-product formations can be eliminated by the disruption of the specific gene. Using the yeast cells, we will be able to "optimize" the localization of foreign protein or to pack foreign proteins into peroxisome. On the basis of our results, we are proposing a new technological concept for gene expression, "**sorting engineering**", that means a heterologous protein is allow to accumulation in a specific organelle of the yeast through the molecular mechanism of the cellular function. The produced enzyme is able to maintained at stable state in the yeast cell by the use of this technique, and this makes possible for large scale production of useful protein and construction of high performance biocatalyst for production of useful chemicals.

References

1. Ogata, K., Nishikawa, H., and Ohsugi, M.: A yeast capable of utilizing methanol, *Agric. Biol. Chem.*, **33**, 1517-1520 (1969).
2. Kato, N., Tani, T., and Yamada, H.: Microbial utilization of methanol: Production of useful metabolites, p171-202. In A. Mizrahi and A.V. van

- Wezel (eds.), *Advances in Biotechnological Processes*, Alan R. Liss, Inc., New York (1983).
3. Sakai, Y., Rogi, T., Takeuchi, R., N. Kato, and Tani, Y.: Expression of *Saccharomyces* adenylate kinase gene in *Candida boidinii* under the regulation of its alcohol oxidase promoter. *Appl. Microbiol. Biotechnol.*, **42**, 860-864 (1995).
 4. Y. Sakai, Rogi, T., Yonehara, T., Kato, N., and Tani, Y.: High-level ATP production by a genetically-engineered *Candida* yeast. *Bio/Technology*, **12**, 291-293 (1994).
 5. Sakai, Y., Akiyama, M., Kondoh, H., Shibano, Y., and Kato, N.: High-level secretion of fungal glucoamylase using the *Candida boidinii* gene expression system. *Biochim. Biophys. Acta*, **1308**, 81-87 (1996)
 6. Sakai, Y., Matsuo, H., Ze, H. K., Saiganji, A., Yurimoto, H., Takabe, K., Saiki, H., and Kato, N.: Isolation and characterization of mutants of the methylotrophic yeast, *Candida boidinii* S2 that are impaired in growth on peroxisome-inducing carbon sources. *Biosci. Biotech. Biochem.*, **59**, 869-875 (1995).
 7. Sakai, Y., Marshall, P.A., Saiganji, A., Takabe, K., Saiki, H., Kato, N., and Goodman J.M.: The *Candida boidinii* peroxisomal membrane protein Pmp30 has a role in peroxisomal proliferation and is functionally homologous to Pmp27 from *Saccharomyces cerevisiae*. *J. Bacteriol.*, **177**, 6773-6781 (1995).
 8. Sakai, Y., Saiganji, A., Yurimoto, H., Takabe, K., Saiki, H., and Kato, N.: The absence of Pmp47, a putative yeast peroxisomal transporter, causes a defect in transport and folding of a specific matrix enzyme. *J. Cell Biol.*, **34**, 37-51 (1996)
 9. Sakai, Y., Yoshida, N., Isogai, A., Tani, Y., and Kato, N.: Purification and properties of fructosyl lysine oxidase from *Fusarium oxysporum* S-1F4. *Biosci. Biotech. Biochem.*, **59**, 487-491 (1995).
 10. Yoshida, N., Sakai, Y., Isogai, A., Tani, Y., and Kato, N.: Distribution and properties of fructosyl amino acid oxidase in fungi. *Appl. Environ. Microbiol.*, **62**, 4487-4489 (1995).
 11. Sakai, Y., Yoshida, N., Tani, Y., and Kato, N.: Production of fructosyl lysine oxidase from *Fusarium oxysporum* S-1F4 on autoclave-browned medium. *Biosci. Biotech. Biochem.*, **60**, 150-151 (1996).
 12. Yoshida, N., Sakai, Y., Isogai, A., Fukuya, H., Yagi, M., Tani, Y., and Kato, N.: Primary structures of fungal fructosyl amino acid oxidases and its application to the measurement of glycated proteins. *Eur. J. Biochem.*, **242**, 499-505 (1996)

PATHOGENESIS - RELATED PROTEIN GENES STRUCTURE AND EXPRESSION

Michał M. Sikorski, Andrzej B. Legocki

Institute of Bioorganic Chemistry, Polish Academy of Sciences
Poznań, Poland

Abstract:

Two homologous acidic polypeptides of M_r 17000 which are constitutively expressed in roots of yellow lupin were identified. The amino acid composition of these proteins and chemical properties classified them as intracellular pathogenesis-related proteins of PR10 class. Our data on protein and DNA level indicate that their expression is susceptible to regulation during symbiotic nitrogen fixation.

Plant Response to Stress

Terrestrial plants have evolved an extraordinary genomic and metabolic plasticity enabling them to survive in different environmental stress conditions. The commonly encountered plant stresses are: pathogen invasion, mechanical damage (wounding) and chemical pollutants.

In response to stress, plants induce a number of defence reactions. The overall plant defence mechanism involves the expression of transcriptionally activated defence genes. The protein products of these genes are involved in the synthesis of phytoalexins (enzymes of phenylpropanoid metabolism), reinforcement of cell walls by deposition of polysaccharides (callose), lignin, structural proteins (hydroxyproline-rich proteins, glycoproteins), accumulation of hydrolytic enzymes (β -1,3-glucanases and chitinases), proteinase inhibitors and intracellular pathogenesis-related proteins (IPRs) of PR10 class.

Intracellular pathogenesis-related proteins have been shown to be ubiquitous in the plant kingdom (Walter *et al.*, 1990). They are structurally related to tree-pollen allergens and to major food allergens from celery and apple (Breiteneder *et al.*, 1989; Vanek-Krebitz *et al.*, 1995). Although the exact cellular localisation of PR10 proteins has not been determined, the absence of apparent signal peptides and the identity of cDNA-predicted amino acid sequences to those obtained from protein sequencing classify them as cytosolic proteins (Walter *et al.*, 1990; Awade *et al.*, 1991). It was reported by several laboratories that PR10 proteins accumulate around the sites of pathogen invasion or wounding (Warner *et al.*, 1994; Pinto & Ricardo, 1995; Breda *et al.*, 1996).

There are suggestions that PR10 proteins play an important function in the plant development. They have been identified in seeds (Warner *et al.*, 1994), developing

roots, senescent leaves (Crowell *et al.*, 1992) and senescent nodules (Sikorski *et al.*, 1989, 1996; Legocki *et al.*, 1997), stems (Warner *et al.*, 1994) and different parts of flowers (Breiteneder *et al.*, 1989; Constabel & Brisson, 1995). It has recently been shown that birch pollen allergen *Bet v 1* belonging to the PR10 protein class revealed RNase activity *in vitro* (Bufe *et al.*, 1996; Swoboda *et al.*, 1996). On the basis of a high amino acid sequence homology and similarity of expression pattern to ginseng ribonuclease, their ribonuclease activity in the defence reaction was suggested (Moiseyev *et al.*, 1994). Due to their structural similarity, PR10 proteins have been classified as ribonuclease-like PR proteins (Van Loon *et al.*, 1994). Recently, the X-ray and NMR structure of birch pollen allergen *Bet v 1* was determined (Gajhede *et al.*, 1996). Proteins of PR10 class have been found in both dicotyledonous and monocotyledonous plants. It may suggest that they evolved from a common ancestor and possess similar functions.

Genes encoding *Lupinus luteus* PR10 proteins

We have identified in yellow lupine two homologous proteins of PR10 class - *L/PR10.1A* (156 amino acids, Mr 16859) and *L/PR10.1B* (156 amino acids, Mr 16655) (Sikorski *et al.* 1989, Legocki *et al.* 1997). The corresponding full-length cDNA clones encoding these proteins were selected from lupine cDNA expression library (in XR UNI-ZAP vector) based on poly A⁺ RNA of non infected plant roots (Sikorski *et al.* 1996). The predicted amino acid sequences of the two protein homologues exhibit 76% identity (90% similarity) indicating that genes encoding these proteins belong to a multigene family. The entire genomic clones *LlYpr10.1a* and *LlYpr10.1b* were purified from EMBL3-phage genomic library, using homologous cDNA probe fragments. Fig. 1 shows structure and organization of genomic clones *LlYpr10.1a* and *LlYpr10.1b* encoding lupin PR10 proteins. They are composed of two exons interrupted by one intron. The alignment of the amino acid sequences of both lupin PR10 proteins is shown in Fig. 2.

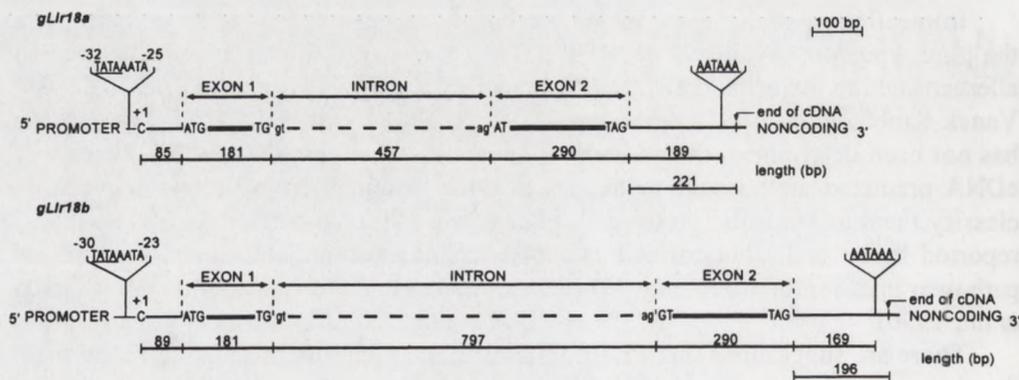


Fig.1. Structure and organization of genomic clones *LlYpr10.1a* and *LlYpr10.1b* encoding PR10 proteins from yellow lupin.

L1R18A	MGIFAFENEQSSLVAPAKLYKALTKDSDEIVPKVIEPIQNVEIVEGNGGP	50
L1R18B	MGVFAFEDEHPSAVAQAKLFKALTKDSDDIIPKVIEQIQSVEIVEGNGGP	50
L1R18A	GTIKKIIAIIHDGHTSFVLHKLDAIDEANLTYNYSIIIGGEGLDDESLEKISY	100
L1R18B	GTVKKIIASHGGHTSYVLHKIDAIDEASFENYSIVGGTGLDESLEKITF	100
L1R18A	ESKILPGPDGGSIGKINVKFHTKGDVLSQAVRQAKFKGLGLFKAIEGYV	150
L1R18B	ESKLLSGPDGGSIGKIKVKFHTKGDVLSDAVREEAKARGTGLFKAVEGYV	150
L1R18A	LAHPDY	156
L1R18B	LANPNY	156

Fig.2. Amino acid sequence comparison of *LIPR10.1A* and *LIPR10.1B* proteins (76% identity, 90% similarity)

Northern hybridization analysis of *Lupinus luteus* PR10 mRNA transcripts

It was shown by Northern hybridization, that the genes coding for *LIPR10* proteins are constitutively expressed in roots of uninoculated plant and their expression is down-regulated during the nodule development after lupin plant is inoculated with symbiotic bacteria - *Bradyrhizobium*, *sp.* (*lupini*). The expression pattern of *LIPR10.1A* gene is shown in Fig. 3. The transcription regulation of *LIPR10.1B* gene expression is similar.

The transcripts of both lupine genes were also detected in mature leaves after infiltration with pathogenic bacteria (Sikorski, *unpublished*). These results implied that both lupine PR10 protein homologues are involved in a more general plant developmental program as well as in the defence mechanism.

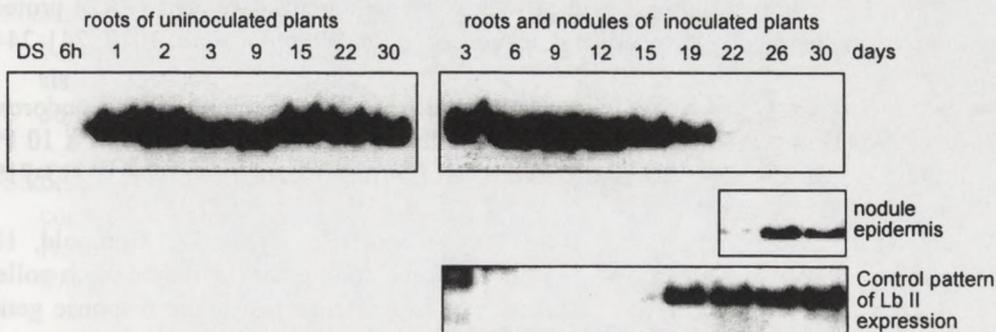


Fig. 3. Northern hybridization analysis of mRNA transcripts encoding *LIPR10.1A* protein.

Analysis of *LlYpr10.1a* promoter in transgenic plants

In order to study the organ specific expression of lupine PR10 genes, we have made a fusions consisting *EcoRI/HindIII* promoter part of *Ypr10.1a* or *Ypr10.1b* genes (-2088/+18 fragment with TATA-box at the position -32/-25) and *gus/int* gene, using a binary plasmid pPR97 (Szabados *et al.*, 1995). These vectors were used for *Agrobacterium*-mediated transformations of *L. corniculatus*. Preliminary results indicate that promoters of both lupin PR-type genes direct the expression of the reporter gene in the tissue-specific manner (Sikorski *et al.* unpublished).

Expression of PR10 proteins in pathogenic and symbiotic interactions

Despite the numerous similarities between pathogenic and symbiotic modes of infection and considering the fact that elements of the molecular recognition might be common in both cases almost all plants exhibit the ability to distinguish between pathogenic and symbiotic interactions. It has recently been suggested, that many aspects of legume root nodulation mechanism can be explained by a simple model based on the known plant defence reactions (Mellor and Collinge, 1995). The question whether an active suppression of plant defence by the invading microsymbiont is a general requirement for the establishment of symbiosis still needs further elucidations.

Acknowledgements

This work was supported by the grant No. 6P04B011-10 from the State Committee for Scientific Research.

REFERENCES

- Awade, A., Metz-Boutique, M.H., Le Ret, M., Genot, G., Amiri, I. & Burkard, G. (1991) The complete amino acid sequence of pathogenesis-related (PR2) protein induced in chemically stressed bean leaves. *Biochim. Biophys. Acta.* **1077**, 241-244.
- Breda, C., Sallaud, C., El-Turk, J., Buffard, D., de Kozak, I., Esnault, R. & Kondorosi, A. (1996) Defence Reaction in *Medicago sativa*: A Gene Encoding a Class 10 PR Protein Is Expressed in Vascular Bundles. *Mol. Plant-Microbe Interact.* **9**, 713-719.
- Breiteneder, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H., Scheiner, O. & Breitenbach, M. (1989) The gene coding for the major birch pollen allergen *Bet v 1* is highly homologous to a pea disease resistance response gene. *EMBO J.* **8**, 1935-1938.
- Bufe A., Spangfort, M.D., Kahlert, H., Schlaak, M. & Becker W.-M. (1996) The major birch pollen allergen, *Bet v 1*, shows ribonuclease activity. *Planta.* **199**, 413-415.

- Constabel, C.P. & Brisson, N. (1995) Stigma- and vascular-specific expression of the PR10a gene of potato: a novel expression of a pathogenesis-related gene. *Mol. Plant-Microbe Interact.* **8**, 104-113.
- Crowell, D.N., John, M. E., Russel, D. & Amasino, R.M. (1992) Characterisation of stress-induced, developmentally regulated gene family from soybean. *Plant Mol.Biol.* **18**, 459-466
- Gajhede, M., Osmark P., Poulsen, F. M., Ipsen, H., Larsen, J. N., Joost van Neerven, R. J., Schou, C., Lovenstein, H., Spangfort, M. D. (1996) X-ray and NMR structure of Betv1, the origin of birch pollen allergy. *Nature Structural Biol* **12**:1040-1045
- Legocki, A. B., Karłowski, W. M., Podkowiński, J., Sikorski, M. M. & Stępkowski T. (1997) Advances in molecular characterization of the Yellow lupin-*Bradyrhizibium* sp. (*Lupinus*) symbiotic model; in *Biological Fixation of Nitrogen for Sustainable Agriculture* (Legocki, A. B., Bothe, H. & Puchler A., eds.), Springer Verlag, *NATO ASI Series*. Vol. **G39**, 262-266.
- Mellor, R. B., Collinge, D. B. (1995) A simple model based on known plant defence reactions is sufficient to explain most aspects of nodulation. *J Exp Bot* **282**: 1-18
- Moiseyev, G. P., Beintema, J. J., Fedoreyeva, L. I. & Yakovlev, G. I. (1994) High frequency similarity between a ribonuclease from ginseng calluses and fungus-elicited proteins from parsley indicates that intracellular pathogenesis-related proteins are ribonucleases. *Planta (Heildelb.)* **193**, 470-472.
- Pinto, P. M. & Ricardo, C. P. P. (1995) *Lupinus albus* L. Pathogenesis-Related Proteins That Show Similarity to PR10 Proteins. *Plant Physiol.* **109**, 1345-1351.
- Sikorski, M. M., Szybiak-Stróżycka, U., Stróżycki, P., Golińska, B., Mądrzak., C. J., Kamp, R. M., Wittmann-Liebold, B. & Legocki, A. B. (1989) Coordinated synthesis of leghemoglobin and root protein R18 in yellow lupin. *Acta Biochim. Polon.* **36**, 63-72.
- Sikorski, M. M., Szlagowska, A. E. & Legocki A. B. (1996) cDNA sequences encoding for two homologues of *Lupinus luteus* (L.) IPR-like proteins (Accession Nos X79974 and X79975 for LIR10A and LIR18B mRNA's respectively). *Plant Physiol. PGR* **95**-114.
- Swoboda, I., Hoffman-Sommergruber, K., O'Riordain, G., Scheiner, O., Heberle-Bors, E. & Vicente, O. (1996) *Bet v 1* proteins, the major birch pollen allergens and members of a family of conserved pathogenesis-related proteins, show ribonuclease activity *in vitro*. *Physiol. Plant.* **96**, 433-438.

- Szabados L, Charrier B, Kondorosi A, de Bruijn, FJ, Ratet P (1995) New plant promoter and enhancer testing vectors. *Mol Breeding* **1**, 419-423.
- Van Loon, L. C., Pierpoint, W. S., Boller, T. & Conejero, V. (1994) Recommendations for naming plant pathogenesis-related proteins. *Plant Mol. Biol. Rep.* **12**, 245-264.
- Vanek-Krebitz, M., Hoffman-Sommergruber, K., Machado, M.L.D., Susani, M., Ebner, C., Kraft, D., Scheiner, O. & Breiteneder, H. (1995) Cloning and sequencing of *Mal d 1*, the major allergen from apple (*Malus domestica*), and its immunological relationship to *Bet v 1*, the major birch pollen allergen. *Biochem. Biophys. Res. Commun.* **214**, 538-551.
- Walter, M. H., Liu, J.-W., Grand, C., Lamb, C. J. & Hess D. (1990) Bean pathogenesis-related (PR) proteins deduced from elicitor-induced transcripts are members of a ubiquitous new class of conserved PR proteins including pollen allergens. *Mol. Gen. Genet.* **222**, 353-360.
- Warner, S. A. J., Gill, A. & Draper, J. (1994) The developmental expression of the asparagus intracellular PR protein (AoPR1) gene correlates with sites of phenylpropanoid biosynthesis. *Plant J.* **6**, 31-43.

PROTEIN CONFORMATION AND MODELLING

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

100000

1995) *Plant Cell Growth and Div. Rev.* 1, 415-425.

1994) *Plant Cell Growth and Div. Rev.* 1, 415-425.

1994) *Plant Cell Growth and Div. Rev.* 1, 415-425.

1994) *Plant Cell Growth and Div. Rev.* 1, 415-425.

1994) *Plant Cell Growth and Div. Rev.* 1, 415-425.

STRUCTURAL BASIS FOR RESTRICTED SUBSTRATE SPECIFICITY AND HIGH CATALYTIC POTENCY OF A LYSYLENDOPEPTIDASE, *ACHROMOBACTER* PROTEASE I

Fumio Sakiyama, Shigemi Norioka, Shigeki Ohta, Seong-II Lim, Kentaro Shiraki, Naruto Katsuragi, Takeharu Masaki*, Yukiteru Katsube, Yasuyuki Kitagawa and Kazumichi Sashi

Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565 and

* Faculty of Agriculture, Ibaraki University, Ami-machi, Inashiki-gun, Ibaraki 300-03, Japan

Abstract: *Achromobacter* protease I (API) is a trypsin-type serine protease with a low degree of sequence similarity to bovine trypsin. API exhibits the restricted substrate specificity for lysine and holds a higher lysylendopeptidase activity than the mammalian enzyme. Site-directed mutagenesis studies and crystallographic analysis revealed that Asp 225 in a shallow S1-pocket is responsible for the lysine specificity and the aromatic-aromatic interaction between Trp 169 and His 210 locating close to Asp 113, a constituent of catalytic triad, is closely associated with the enhanced catalytic potency of API.

API, an extracellular serine protease of a gram-negative soil bacterium, *Achromobacter lyticus*, was isolated by Masaki et al.(1,2). The protease consists of a single polypeptide chain of 268 residues with three disulfide bonds including two novel ones (Fig.1). Although amino acid sequence homology between API and bovine trypsin is as low as 20%, key constituents for catalytic triad (His 57, Asp 113 and Ser 194) and substrate binding subsites S1, S2 and S3 (His 210, Gly 211 and Gly 212, respectively) were predicted and the bacterial enzyme was classified as a member of the trypsin family (3) (Fig.1). API holds four characteristics distinct from bovine trypsin; 1) restricted substrate specificity for lysine, 2) several times higher activity than trypsin, 3) a broad optimal pH range (pH 8.5-10.5) and 4) stability against urea and SDS (4). Based on a comparison of the primary structure for API and trypsin, it has been thought that Glu 190 or Asp 225 is a candidate amino acid for primarily determining lysine specificity and His 210 is possibly associated with a higher

enzyme activity than that of trypsin. Then we have begun studies to dissect structural features for these particular properties of API. This paper describes the result of our recent investigations on the structural basis for lysine specificity

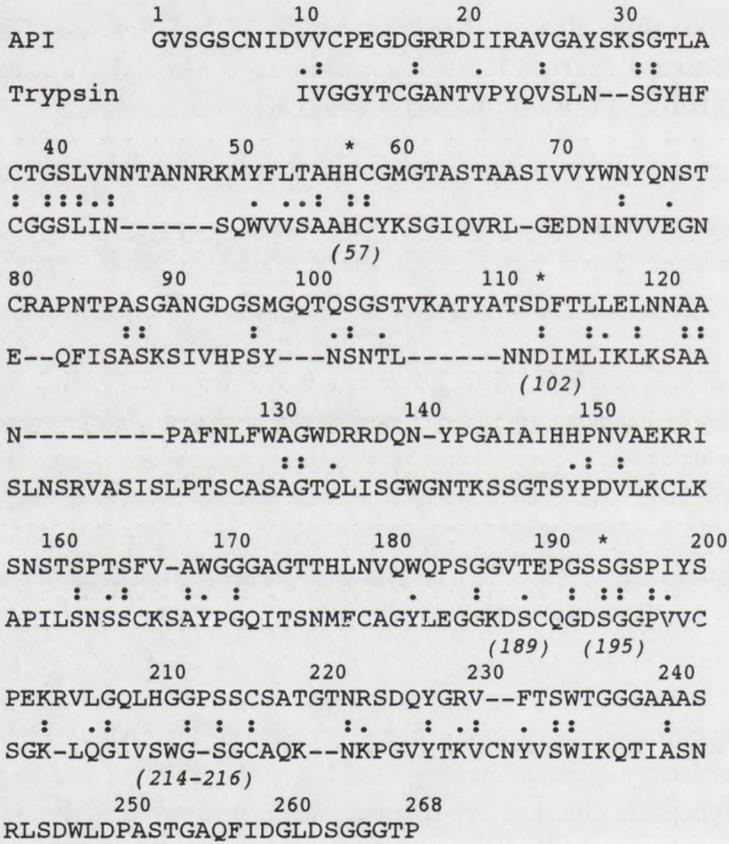


Fig.1. Amino acid sequences of API and bovine trypsin. Numbers in italics for trypsin are based on chymotrypsinogen numbering. Constituent amino acids of catalytic triad are marked (asterisks). Disulfide bonds are present at Cys 6-Cys 216, Cys 12-Cys 80 and Cys 36-Cys 58 in API.

and high catalytic potency by site-directed mutagenesis experiments and crystallographic analysis.

Asp 225 is essential for the lysine specificity and high activity of API

To identify the acidic residue to which the lysine substrate binds, Glu 190 and Asp 225 were individually replaced by other amino acids using site-directed mutagenesis. All Glu 190 mutants tested retained the lysylendopeptidase activity (5). However, the Asp 225→Glu mutant was a sole active enzyme secreted to the periplasm. Since API is synthesized as an inactive preproprotein of 653 residues and activated by subsequent selective hydrolysis of a single peptide bond, the Lys -1-Gly 1 bond (6), the proprotein can not be converted to the active form unless the processed protein is active. The Asp 225 mutants having no negative charge at position 225 were synthesized as inactive precursor proteins, but could not be activated, showing that their putative mature proteins are inactive. Even the Asp225→Glu mutant was much less active than the wild-type enzyme. These results suggested that Asp 225 serves as the negative charge that binds the lysine side chain and is essential for highly active API.

Three dimensional structure, active site and TLCK-API

The three dimensional structure was solved at 1.2 Å resolution for API crystals obtained at pH 8.0. The main chain folds to a tertiary structure similar to that of trypsin though similarity is low at the primary structure level. Two domains (Gly 1-Phe 127 and Arg 137-Ala 255) are formed and Asn 128-Arg 136 constitutes an anti-parallel β -structure with Gln 256 - Ser 263, possibly stabilizing the two domain structure. As expected, an ion-pair corresponding to the Ile (16) - Asp (194) pair of trypsin is absent and instead the Cys 6-Cys 216 bond bridges the N-terminus to the active-site region. The α -amino group of Gly 1 is fixed by hydrogen bonding to Tyr 75, Asn 91 and Glu 154. The main chain folding of the active site is quite similar for API and trypsin so as to be roughly superimposed. The remarkable difference is the position of a negative charge in the S1-pocket: in API, Asp 225 is located at a 1.5 Å shorter position from the Ser 194 hydroxyl group when compared with Asp (189) in trypsin. Accordingly, it is difficult for the arginine side chain to reside in such a shallow S1-pocket of API. This is a major cause for restricting the substrate specificity of the enzyme to lysine. Another structure distinct from the S1-pocket of trypsin is the presence of a hydrogen bond network involving one bound water molecule (W420) fixed by several hydrogen bonds. Amino acid residues involved in the network are Asp 225, Ser 214, Thr 189

and Trp 182, constituents of the bottom part of the S1-pocket (Fig.2).

Outside the S1-pocket, the arrangement of aromatic rings is also different from that found for most members of trypsin family. The most remarkable is an imidazole-indole stack near Asp 113, a member of catalytic triad. The role of this aromatic-aromatic interaction will be discussed later.

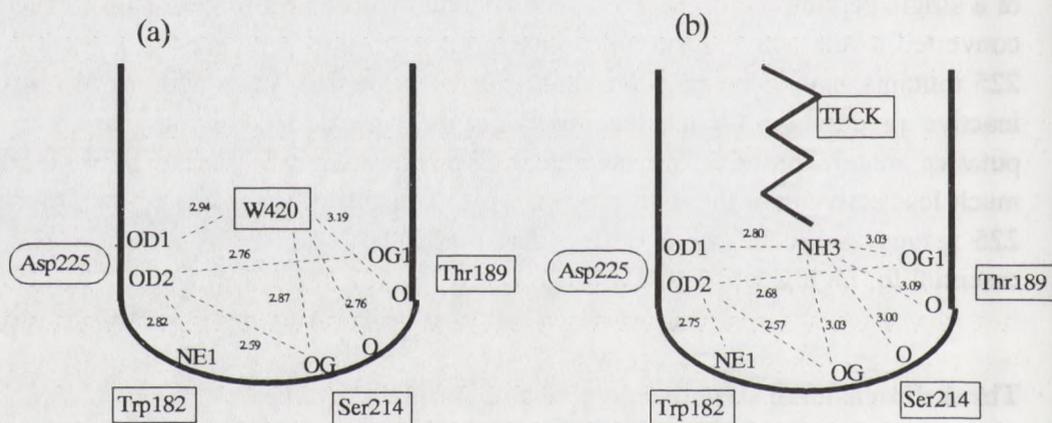


Fig.2. Schematic representation of a hydrogen bond network in the bottom part of API (a) and TLCK-API (b). Hydrogen bond distance is noted (\AA).

To obtain information concerning the detail of the lysine-aspartic interaction in the bound lysine substrate, the TLCK-API complex was prepared and its crystal structure was solved at 2.8 \AA resolution. The tertiary structure of API in the complex was identical with that of the unmodified enzyme except for the active site region where TLCK is covalently bound to His 57. As expected, the side chain of lysine in the inhibitor is situated in the S1-pocket and its ϵ -amino group locates close to Asp 225 carboxyl by occupying the position that W420 was present in the free enzyme. Namely, the ϵ -amino nitrogen firmly binds to the S1-pocket by both electrostatic interaction with Asp 225 and integration into the hydrogen network in place of W420 (Fig.2). Atoms of the bound lysine side chain thus filled the S1-pocket and no space exists for the binding of additional atoms. It is impossible for the arginine side chain to be accepted in the S1-pocket in the same manner. It appears that the S1-pocket of API can accept the side chain of lysine at the best position for

the efficient hydrolysis of its peptide bond at the C-terminal side, implying that API is a particular endopeptidase specific for the lysyl peptide bond. Since the binding of a substrate molecule determines not only substrate specificity but also orientation of the scissile bond toward the catalytic serine hydroxyl group, the binding mechanism as described here may constitute a key factor for substrate binding by which API can act as a powerful lysylendopeptidase.

High peptidase activity and Asp 113-His 210-Trp 169 triad

Subsite mapping analysis (7) and amino acid sequence alignment suggested that API has three substrate binding subsites, S1, S2 and S3, composed of His 210, Gly 211 and Gly 212. The presence of histidine at subsite S1 is novel and possible contribution to the enhanced peptidase activity of API has been proposed (3). As mentioned earlier, crystallographic analysis revealed that the imidazole ring of His 210 locates close to the carboxylate of Asp 113 at one side of ring surface and another side faces the surface of the indole ring of Trp 169 to form a stacked structure of two aromatic rings (Fig.3). Distance to either the putative surface composed of the β -carboxylate of Asp 113 or

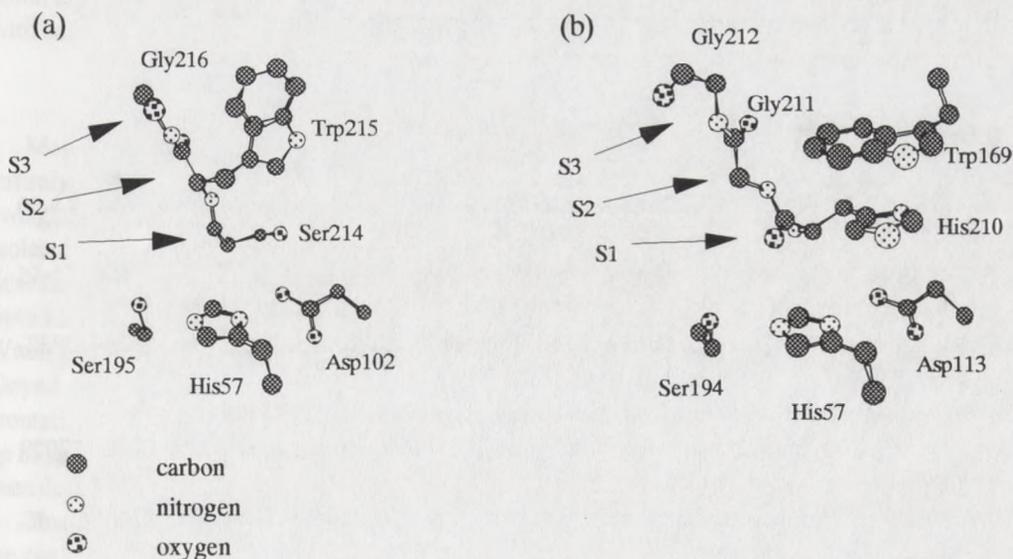


Fig.3. Aromatic rings in the active site of API (a) and bovine trypsin (b).

the Trp 169 indole ring is around 3.5 Å. So far, this type of interaction involving the Asp-His-Trp triad in the active site has not been reported for serine proteases.

Since the Asp 113-His 210-Trp 169 interaction was thought to contribute to the enhancement of peptidase activity, His 210 and Trp 169 mutants were prepared and their catalytic potency was examined. It was shown that removal of the imidazole ring from His 210 affected both K_M and k_{cat} . The effect of mutation on the enzyme activity differed depending on the nature of substrate: a hydrophobic tripeptide substrate (Boc-Val-Leu-Lys-MCA) was almost insensitive to the loss of the imidazole. In contrast, the substitution of Trp 169 for non-aromatic residues led to a significant decrease in the enzyme activity even for the hydrophobic tripeptide substrate. These results suggest that His 210 plays a dual role in the functionality of the active site. One role is to act as a subsite S1 that binds the main chain part of a lysine substrate at subsite P1 to properly orient the scissile bond. Another one is to participate in cooperation of the substrate binding site with the catalytic site through the interaction with Asp 113. By shielding the imidazole ring from surrounding water molecules, Trp 169 is likely to play a critical role in the stabilization and/or maintenance of this His 210-Asp 113 pair to potentiate the function of the active site. It appears reasonable that API is a new type of highly active, lysine-specific serine protease which has a catalytic quadruplet (Ser 194-His 57-Asp 113-His 210) and a built-in potentiator (Trp 169) for the quadruplet.

References

- 1) T.Masaki, M.Tanabe, K.Nakamura and M.Soejima, *Biochim.Biophys. Acta*, **660**, 44-50 (1981)
- 2) T.Masaki, T.Fujihashi, K.Nakamura and M.Soejima, *Biochim.Biophys. Acta*, **660**, 51-55 (1981)
- 3) S.Tsunasawa, T.Masaki, M.Hirose, M.Soejima and F.Sakiyama, *J.Biol.Chem.*, **264**, 3832-3839 (1989)
- 4) F.Sakiyama and T.Masaki, *Methods in Enzymology*, **244**, 126-137 (1994)
- 5) S.Norioka, S.Ohta, T.Ohara, S-I. Lim and F.Sakiyama, *J.Biol.Chem.*, **269**, 17025-17029 (1994)
- 6) T.Ohara, K.Makino, H.Shinagawa, A.Nakata, S.Norioka and F.Sakiyama, *J.Biol.Chem.*, **264**, 20625-20631 (1989)
- 7) F.Sakiyama, M.Suzuki, A.Yamamoto, I.Aimoto, S.Norioka, T.Masaki and M.Soejima, *J.Protein Chem.*, **9**, 297-298 (1990)

MODELLING ENZYME ACTION

CLARE E. SANSOM^{1,2} AND MARIUSZ JASKÓLSKI^{1,3}

¹ Center for Biocrystallographic Research, Polish Academy of Sciences, Poznań, Poland

² Department of Crystallography, Birkbeck College, University of London, UK

³ Department of Crystallography, Adam Mickiewicz University, Poznań, Poland

The term "molecular modelling" is generically applied to any use of computers to model the structure, properties, function and interaction of molecules. It can be conveniently divided into molecular simulation which describes the properties of molecules, such as energy and geometry, in mathematical terms, molecular visualisation which uses advanced graphics, and various analytical techniques. These techniques are increasingly being applied to physiologically important molecular systems – typically proteins complexed with ligands or with DNA – as a part of rational, or structure-assisted, drug design. As the second of those terms implies, molecular modelling is dependent on accurate and detailed knowledge of molecular structures, most often obtained using X-ray crystallography. The first X-ray structure of a protein – myoglobin – was solved by John Kendrew in 1957 [1]. The next thirty years produced no more than a few hundred protein structures. However, the number of available macromolecular structures is now growing very rapidly: the October 1997 version of the Protein DataBank [2] contains 6514 total entries, of which 6015 are structures of proteins or of protein – nucleic acid complexes. The proportion of those structures which were obtained using nuclear magnetic resonance is growing and is now over 15%. In this introductory paper we show, using two case histories, how one molecular simulation technique, molecular mechanics, is applied to the study of enzymes and their interactions with substrates and inhibitors.

MOLECULAR MECHANICS

Molecular mechanics is probably the most widely used molecular simulation technique. It is entirely based on classical Newtonian mechanics. A molecule is represented as a number of "weights" (i.e. atomic centres) connected by "springs" (i.e. bonds). The potential energy of the molecule in a particular conformation is calculated using a simple, empirically-defined, mathematical function termed a force field. Force fields typically contain terms representing deviations from ideal bond lengths, bond angles and torsion angles, as well as those representing Van der Waals and electrostatic interactions between non-bonded atoms. Different parameters are employed to represent atoms in different chemical environments: for example, carbon atoms in aromatic hydrocarbons, in aliphatic hydrocarbons and in carbonyl groups. The force field is made up of the energy equation and the parameters representing each atom type. Some force fields have been developed for use with specific classes of molecules: most typically, proteins or nucleic acids.

The potential energy of a molecule is a function of the conformation of the molecule. Altering the conformation to give the lowest energy, as defined by the force field, is termed energy minimisation or geometry optimisation. This should result in a conformation close to a real physical structure. However, application of this technique will only determine a low-energy conformation

close to the particular starting conformation, which is not necessarily (and, in practice, will very rarely be close to) the global energy minimum. One way of exploring larger regions of conformational space is using molecular dynamics, or simulated molecular motion. Atoms are initially assigned random velocities and the dynamic trajectory of the molecule is calculated from the forces on the individual atoms. Both molecular mechanics and molecular dynamics techniques are often applied to the simulation of complexes between enzymes and substrates or inhibitors.

Enzymes are divided into six classes – oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases – according to the reaction catalysed. Hydrolases catalyse the transfer of a water molecule from a donor to an acceptor. The enzymes highlighted in the two case studies following – HIV protease and L-asparaginase – fall into this class. All proteases catalyse the cleavage of a peptide bond with the addition of a water molecule; asparaginases catalyse the hydrolysis of L-asparagine into L-aspartate and ammonia.

CASE HISTORY 1: HIV PROTEASE

HIV protease is an aspartic protease. It is inhibited by the renin inhibitor, pepstatin, and is most active at acid pH. The HIV protease sequence contains the triad “DTG” which is characteristic of aspartic proteases. All known non-retroviral aspartic proteases contain two such sequences. Crystal structures of HIV protease [3-5] confirmed the earlier prediction that it would form a dimer [6], and that the active site would contain the DTG triads from both monomers. A substrate or inhibitor molecule is bound tightly into a long, deep groove between the monomers, with a “flap” from each monomer closed over it. Peptide-mimetic inhibitors bind in an extended conformation and are held in place by a network of main chain – main chain hydrogen bonds with the enzyme. One tetrahedrally-coordinated water molecule, known as the “structural” or “flap” water bridges the inhibitor and the flaps with hydrogen bonds. Three subsites to accommodate substrate or inhibitor side chains (S1-S3 and S1'-S3') are well defined on each side of the scissile bond or its surrogate.

An early comparison of the interaction energies between HIV protease and three inhibitors - JG365, MVT-101 and U85548e - using energy minimisation did not find a correlation between the interaction energy and the binding constants [7]. However, the contribution of the “transition-state analogue” hydroxyl group, present only in JG365 and U85548e, to the total binding energy was consistently calculated to be approximately 1.5 kcal/mol. These calculations were performed using the molecular mechanics program Discover [8] (Biosym; now MSI), with the CVFF force field which is parametrised for amino acids and proteins.

A more detailed study of the same set of inhibitors, and comparing the Discover protocol with the program Brugel, was later performed in order to determine general rules for substrate and inhibitor binding [9]. Both protocols employed a “bulk solvent” model, with a distance-dependent dielectric constant, but including the “structural” water molecule explicitly. Both force fields included explicit parametrisation for hydrogen bonds. In the calculations using Discover, one of the catalytic aspartic acid residues was protonated, and a full minimisation was carried out. By contrast, the Brugel protocol used charge screening to simulate the electrostatic environment of the active site and only limited minimisation was performed to keep the simulated protease conformation close to that of the crystal structure. However, both these protocols emphasised the importance of the main chain – main chain hydrogen bonds in enzyme-inhibitor binding. These

interactions were predicted to contribute between 56% and 68% of the total binding energy for all complexes and using both protocols. Thus, altering the sequences of peptide-mimetic HIV protease inhibitors is likely to have only a limited effect on inhibitor potency. The relative contribution to the total interaction energy was calculated to be largest for the residue in subsite S2. A comparison of these structures with 15 other published structures of HIV protease with different inhibitors also revealed that the network of main chain – main chain hydrogen bonds was conserved in subsites S3 to S3'. Hydrogen bonds between the flexible flaps, the “structural” water molecule, and the inhibitors were short and very well conserved.

Since 1995, four HIV protease inhibitors have been licenced by the U.S. Food and Drug Administration for use as therapies against AIDS. Three of these, including Invirase, the first such drug to be approved, are peptide-mimetic inhibitors; molecular simulation techniques such as these have helped in the process of their development.

CASE STUDY 2: L-ASPARAGINASE

Two bacterial L-asparaginases – those isolated from *Escherichia coli* (EcA) and from *Erwinia chrysanthemi* (ErA) – are very potent anti-leukaemic agents [10]. Asparaginase treatment can be associated with severe side effects, which are believed to be related to their glutaminase activity [11]. In order to be able to “engineer out” the unwanted glutaminase activity, and increase asparaginase affinity (which is a hallmark of therapeutic value in bacterial asparaginases), it is necessary to understand fully the structure and mechanism of this enzyme. The precise mechanism of action of these enzymes is not completely clear, although hydrolysis is known to proceed in two steps via a β -acyl-enzyme intermediate formed through nucleophilic attack by the enzyme. The residue which acts as the primary nucleophile for this reaction is still unknown.

The crystal structures of these bacterial L-asparaginases are known [12,13]. The asparaginase monomer has a two-domain fold. Both domains fit into the α/β class; the larger N-terminal domain contains an unusual left-handed β - α - β crossover similar to that found in flavodoxin, which forms a “cradle” for the active site. The enzyme is active as a tetramer. This tetramer consists of a pair of dimers, each with an extensive dimer interface: it can be described as a dimer of “intimate dimers”. Each “intimate dimer” contains two active sites, and each active site contains some residues from both monomers. The crystal structure of EcA [12] contains four molecules of L-aspartate (the product of the “forward” reaction, or the substrate for the “reverse” reaction), one bound into each active site. Two threonine residues – T12 and T89 in the *E. coli* enzyme – which are conserved throughout the L-asparaginase family, and are known to be essential for activity, lie close to the bound aspartate. These threonines are the most likely residues to act as the primary nucleophile. Removing either of these threonines by site-directed mutagenesis reduces asparaginase activity; removing both destroys all activity.

A molecular dynamics simulation of a single intimate dimer of *E. coli* L-asparaginase was performed in order to determine which threonine residue was more likely to act as the primary nucleophile. A molecule of the product, L-aspartate, was bound into each of the two equivalent active sites (as observed in the crystal structure). As with some of the calculations on HIV protease described above, the program Discover was used with the CVFF force field, and solvent was modelled using a distance-dependent dielectric constant. In each active site, the water molecule directly hydrogen-bonded to the aspartate was included in the model. After an initial

2500 cycles of energy minimisation, 25ps of molecular dynamics simulation were carried out at 300K, starting from the minimised structure.

During the dynamics run, as expected, most of the scaffold of the active site remained fairly inflexible. A loop, made up of residues 12 to 25, which forms a "cap" over the active site was seen to "open" and "close" over the active site cavity during the simulation. After minimisation, the side chain carboxylate group of each aspartate remained in close contact with the side chains of both T12 and T89 in the respective monomer, implying strong polar interactions. During the dynamics simulation, these potential aspartate-threonine hydrogen bonds were in rapid exchange, with each potential hydrogen bond breaking and re-forming over a timescale of a few picoseconds. In one active site, the side chain carboxylate group remained bound to the active site threonines throughout the simulation, although it was bonded to T89 for a higher percentage of the time than to T12. However, in the other equivalent active site the aspartate molecule "flipped over" during the simulation, so that the main chain carboxyl group occupied an equivalent position to the side chain carboxylate. The fact that this gross change in orientation was possible without a significant increase in energy indicates the relatively large size of the active site and the number of quasi-equivalent low energy positions that it must be possible for the substrate to adopt. The fact that glutamine (with one more CH₂ group than asparagine) can also act as a substrate of this enzyme is another indication of the large size of the active site cavity.

As the two active site threonine residues behaved in an equivalent manner throughout the simulation, it was not possible to determine which is more likely to act as the primary nucleophile. We are currently repeating these calculations using the structures of two mutant enzymes, T89V and T12V, to simulate the behaviour of the enzyme if either potential nucleophile is removed. In the recently solved crystal structure of the T89V mutant [14] the aspartate ligand was modelled to be covalently bound to T12 through its side chain carboxy group, in a structure which may resemble an acyl-enzyme intermediate.

Acknowledgments

We would like to acknowledge the contributions of Dr. A. Wlodawer and several of the members of his group at the Macromolecular Structure Laboratory, Frederick, Maryland, USA to the crystal structures described in this paper, and to acknowledge the collaboration of Drs. I. Weber, R. Harrison and M. Prvost in the molecular modelling of HIV protease. Our collaboration is supported by grants from the British Council, the Royal Society, the Polish Academy of Sciences, the Polish Cultural Institute and the Open Society Institute. The research of Mariusz Jaskolski was supported in part by an International Research Scholars' award from the Howard Hughes Medical Institute.

References

- [1] Kendrew, J. C. (1961). *Sci. Am.* 205, 96-101.
- [2] Bernstein, F.C. et al. (1977) *J. Mol. Biol.* 112, 535-547.
- [3] Navia, M.A. et al. (1989) *Nature* 337, 615-620.
- [4] Wlodawer, A. et al. (1989) *Science* 245, 616-621.
- [5] Lapatto, R., et al. (1989) *Nature* 342, 299-302.
- [6] Pearl, L.H. and Taylor, W.R. (1987) *Nature* 329, 351-354.
- [7] Sansom, C.E. et al. (1992) *Protein Eng.* 5, 659-67.
- [8] Insight II and Discover (1989-1997) Molecular Simulations Inc. (formerly Biosym). San Diego, CA, USA.
- [9] Gustchina, A. et al. (1994) *Protein Eng.* 7, 309-317.
- [10] Hill, M. et al. (1967) *J. Am. Med. Assoc.* 202, 882-888.
- [11] Distasio, J.A. et al. (1982) *Int. J. Cancer* 30, 343-347.
- [12] Swain, A.L. et al. (1993) *Proc. Nat. Acad. Sci. USA* 90, 1474-1478.
- [13] Miller, M. et al. (1993) *FEBS Letts.* 328, 275-279.
- [14] Palm, G.J. et al. (1996). *FEBS Letts.* 390, 211-216.

BINDING OF DIFFERENT DIVALENT CATIONS TO THE ACTIVE SITE OF ASV INTEGRASE

Grzegorz Bujacz^{1,2}, Jerry Alexandratos² and Alexander Wlodawer²,
George Merkel³, Mark Andrade³, Richard Katz³ and Anna Marie Skalka³

¹*Faculty of Food Chemistry and Biotechnology, Technical University of Łódź,
Łódź, Poland*

²*Macromolecular Structure Laboratory, NCI-Frederick Cancer Research
and Development Center, ABL-Basic Research Program, Frederick, MD*

³*Institute for Cancer Research, Fox Chase Cancer Center,
Philadelphia, PA*

Retroviral integrases (IN) contain two known metal-binding domains. The N-terminal domain includes a Zn-finger motif and has been shown to bind Zn^{2+} , while the central catalytic core domain includes a triad of acidic amino acids that bind Mn^{2+} or Mg^{2+} , the metal cofactors required for enzymatic activity. The integration reaction occurs in two distinct steps; the first is a specific endonucleolytic cleavage step, called "processing," and the second is a polynucleotide transfer, or "joining" step. Our previous results showed that the metal preference for *in vitro* activity of avian sarcoma virus (ASV) IN is $Mn^{2+} > Mg^{2+}$, and that a single cation of either metal is coordinated by two of the three critical active site residues (Asp-64 and Asp-121) in crystals of the isolated catalytic domain. Here, we report that Ca^{2+} , Zn^{2+} , and Cd^{2+} , can also bind in the active site of the catalytic domain. Furthermore, two Zn and Cd cations are bound at the active site, with all three residues of the active site triad (Asp-64, Asp-121, and Glu-157) contributing to their coordination. These results are consistent with a two metal-mechanism for catalysis by retroviral integrases. We also show that Zn^{2+} can serve as a cofactor for the endonucleolytic reactions catalyzed by either the full length protein, a derivative lacking the N-terminal domain, or the isolated catalytic domain of ASV IN. However, polynucleotidyl transferase activities are severely impaired or undetectable in the presence of Zn^{2+} . Thus, although the processing and joining steps of integrase employ a similar mechanism and the same active site triad, they can be clearly distinguished by their metal preferences.

The retroviral integrase (IN) is a virus-encoded enzyme that catalyzes integration of viral DNA into host DNA (1-3). As DNA integration is an essential step in the virus life cycle, IN is an important target for the design of drugs that will block the replication of pathogenic retroviruses such as the human immunodeficiency virus (HIV). The integration reaction occurs in two distinct steps. First, IN nicks the viral DNA near the 3' ends of both strands (the "processing" reaction); it then inserts these ends into host target DNA (the "joining" reaction). Both reactions comprise a nucleophilic attack by an hydroxyl oxygen on a phosphorous atom in the DNA backbone; the hydroxyl is derived from a water molecule in processing, and from the newly formed 3'-OH at the end of the viral DNA in joining. Divalent cations, Mn^{2+} or Mg^{2+} , are known to be required as cofactors. *In vitro*, the reactions are most efficient in the presence of Mn^{2+} , but as Mg^{2+} is more abundant in living cells, it is generally presumed to be the physiologically relevant cation.

Retroviral integrases contain approximately 300 amino acids and are composed of three domains (4). The first two domains are highly conserved, and both include metal binding sites. The N-terminal domain (amino acids ~1-50) contains a Zn finger-like motif, HHCC. Binding of Zn^{2+} at this site stabilizes the structure of HIV-1 IN and enhances multimerization and activity (5-7). The central, catalytic domain (amino acids ~50-200) is characterized by a triad of invariant acidic amino acids (Asp-64, Asp-121, and Glu-157, in ASV IN), the last two separated by 35 amino acids, comprising the D,D(35)E motif. These three acidic residues are essential for both processing and joining activity, and have been proposed to bind the divalent metal cofactors during catalysis (8).

Solution of the crystal structures of the isolated catalytic core domains of HIV-1 (9,10) and avian sarcoma virus (ASV) IN (11,12) have revealed that these retroviral enzymes belong to a superfamily of nucleases and polynucleotidyl transferases, all of which contain a cluster of conserved acidic amino acids at their presumed active sites. HIV-1 reverse transcriptase (RT) ribonuclease H (RNase H) domain, another member of this superfamily, was shown to bind two divalent cations in this site (13), prompting the suggestion (14) that all members of this family may use a two-metal catalytic mechanism like that deduced for the 3'-5' exonuclease of *E. coli* DNA polymerase I (15,16). We have shown that side chains of two of the acidic triad residues in the D,D(35)E motif in ASV IN also form a metal binding pocket. A single ion, Mn or Mg, is complexed to Asp-64 and Asp-121 when crystals of the isolated catalytic domain of ASV IN are soaked in metal-containing solutions (12). We hypothesized that a second metal might be bound between the Asp-64 and Glu-157 in the full length protein or in the presence of substrate.

Here we show that additional divalent cations can also bind in the active site of crystals of the isolated catalytic core domain of ASV IN. Moreover, in the case of Zn^{2+} and Cd^{2+} , two ions are complexed by side chains from all three of the acidic amino acids of the D,D(35)E motif (17). To investigate the significance of these observations, we measured enzymatic properties in the presence of these metals with the isolated catalytic core, full length ASV IN protein, and an N-terminal deletion derivative that lacks the Zn-finger motif but retains both processing and joining activities. The results of the structural analysis are consistent with a two-metal reaction mechanism. The biochemical studies show that while Zn^{2+} is a cofactor for the hydrolytic, nicking activities, it provides little or no detectable support for the polynucleotidyl transferase activities of IN.

Mn^{2+} , Mg^{2+} and Ca^{2+} bind to a single site in ASV catalytic domain. In our earlier studies we soaked crystals of the isolated ASV IN catalytic domain with the two known divalent cation co-factors, Mn^{2+} and Mg^{2+} (12). We observed coordination of both cations between Asp-64 and Asp-121 of the catalytic triad, but no participation of its third member, Glu-157. To further investigate the possible participation of Glu-157 in metal binding, we carried out a systematic analysis encompassing a variety of divalent cations and a wider range of concentrations.

Crystals of the catalytic domain of ASV IN (amino acids 52-207) were soaked in 2 mM to 500 mM solutions of five divalent cations: Mg^{2+} , Mn^{2+} , Zn^{2+} , Ca^{2+} , and Cd^{2+} . Structures of the metal-soaked crystals were solved and refined at moderately high resolution. The electron density map corresponding to the structure of the Mg^{2+} complex obtained at the highest concentration of its salt is exceptionally clear and shows a cation bound between the carboxylates of Asp-64 and Asp-121 of the D,D(35)E motif (denoted

site I). However, no indication of binding of a second cation could be found in this map. Less than 30% occupancy was observed at a lower concentration of Mg^{2+} (20 mM), indicating that binding was much weaker. A similarly positioned, single metal ion was detected at full occupancy after soaking the crystals in an even lower (10 mM) concentration of $MnCl_2$. No additional metal binding sites were observed even in 500 mM $MnCl_2$. A higher apparent affinity for Mn^{2+} is consistent with observations that ASV IN is approximately 30 fold more active in Mn^{2+} than Mg^{2+} .

Crystals soaked in 100 mM $CaCl_2$ were also found to bind a divalent cation at site I, with interactions between the Ca ion and the active site residues similar to those observed with Mg^{2+} and Mn^{2+} . With Ca^{2+} , coordination is an incomplete octahedron in appearance, a square pyramid with the metal in the center of the square base, and three water molecules in the coordination sphere.

Zn^{2+} and Cd^{2+} bind at multiple sites in the ASV IN catalytic domain. Unexpectedly, we found that Zn^{2+} binds in four separate sites on the surface of the isolated catalytic domain of ASV IN, with partial occupancy observed at a concentration as low as 2 mM, and full occupancy of the sites at 100 mM. At the higher concentration of Zn^{2+} , two of the ions bind in the catalytic center, one at site I between Asp-64 and Asp-121, and a second, denoted site II, coordinated by Asp-64 and Glu-157. The distance between the Zn cations in the active site is 3.62 Å. Site III is in a loop defined by amino acids 92-107, with the ion directly coordinated to His-103, coordinated to His-93 via a water molecule, and liganded with two other water molecules. The fourth Zn ion is bound in the vicinity of the C terminus of the catalytic domain, and is coordinated by residues His-198 and Tyr-194 (site IV).

Zn ions bound at sites III and IV show clear tetrahedral coordination, while the cations in sites I and II are coordinated to two oxygens of the carboxylates and one water molecule each. One additional water molecule may complete a tetrahedral coordination by cations in the catalytic center by bridging both Zn ions. With the crystal soaked in 2 mM Zn^{2+} , we observed a pattern of metal binding similar to that at 100 mM. In this case, both catalytic binding sites showed less than complete occupancies, with site II lower than site I. Temperature factors are correspondingly higher for site II than for site I. However, the structure with 100 mM Zn^{2+} had complete occupancy and relatively low temperature factors for both cations located in the active site.

Two Cd^{2+} ions were visible in the catalytic center after soaking in 100 mM $CdCl_2$, again coordinated to all three carboxylate residues of the essential triad. However, there was lower occupancy of site II and higher temperature factor for this ion than for the ion in site I. This is similar to binding observed at the lowest concentration (2 mM) of Zn^{2+} and again suggests that the cations in site II may be bound less strongly than those occupying site I. The distance between the Cd ions in the catalytic center is 4.06 Å. The Cd^{2+} in site I has an almost perfect octahedral coordination sphere, very similar to the coordination of Mn^{2+} and Mg^{2+} . The Cd^{2+} in site II has deformed octahedral coordination, sharing a water ligand with the first cation. Singularly in this case, both Glu-157 carboxyl oxygens coordinate with this one cation. The most solvent-accessible water molecules liganded to each metal, bound opposite to the Asp-64 ligand, have slightly longer hydrogen bonding distances and higher temperature factors.

The atomic coordinates of the three crucial carboxylic acids are only slightly affected when one or more cations are bound. There is practically no change in the position of the side chain of Asp-64 (mean shift of the atomic positions from their average values, calculated for all metal structures, is 0.125 Å), a very slight movement of the side chain of Asp-121 (0.187 Å), and a more pronounced difference in location of the side chain of Glu-157 (0.857 Å shift). The minimal deviation in carboxylate positioning upon metal binding is consistent with the observation that even single conservative substitutions in these residues drastically reduced activity of both ASV and HIV-1 IN (8).

Activity of the ASV IN catalytic domain with various divalent cations. Having observed Ca^{2+} , Zn^{2+} , and Cd^{2+} occupancy of site I (or I and II) in the active site of catalytic domain, we asked whether any of these cations could function as cofactors for enzymatic activity of IN. The isolated catalytic core domain of ASV IN displays two metal-dependent activities, a DNA endonuclease "nicking" activity, and a DNA cleavage-ligation "disintegration" activity (11,18). DNA nicking by the ASV IN catalytic domain is quite efficient and similar in specific activity to that of the full length protein. Therefore we first assayed for this nicking activity, which is characterized by preferred cleavage between the C and A of the conserved CA dinucleotide near the viral DNA termini (the "-3" site). This endonucleolytic activity is distinct from "-2" processing and its biological relevance is not yet understood. The catalytic core domain was incubated with a short DNA duplex substrate that represents a viral DNA end, in the presence of varying concentrations of ZnCl_2 , ZnSO_4 , CaCl_2 and CdSO_4 , as well as the known metal cofactors, MgCl_2 and MnCl_2 . Of the new metals tested, only Zn^{2+} supported significant activity. The Zn^{2+} -dependent activity shows a sharp peak at approximately 2 mM ZnSO_4 ; similar results were obtained with ZnCl_2 . At this peak, activity is approximately half of that observed with the same concentration of MnCl_2 . However, significantly less activity is observed at higher ZnSO_4 concentrations. The optimum concentration of MnCl_2 is approximately 10 mM, with little change in activity up to 25 mM. The decreased activity at ZnSO_4 concentrations higher than 2 mM is not due to the anion, as similar results were observed with the chloride and sulfate salts. We conclude that the ASV IN catalytic domain displays significant nicking activity with Zn^{2+} as a cofactor at a concentration in which site I is likely to be fully occupied and site II at least partially occupied. Higher concentrations of Zn^{2+} are inhibitory. Although Mg^{2+} and Ca^{2+} could bind to site I, and Cd^{2+} could bind to site I and II in the crystal, no nicking activity could be detected in the presence of a broad range of concentration of these metals.

The ASV catalytic domain was also assayed for disintegration activity, which proceeds at 0.2% of the rate observed with the full length protein. In the presence of 10 mM MnCl_2 disintegration is clearly detectable with the catalytic domain, as described previously (20). However, no disintegration activity was observed with 2 mM ZnSO_4 , even after prolonged exposure of the autoradiogram. We conclude that Zn^{2+} is unable to support significant disintegration activity of the isolated catalytic domain under these conditions.

Zn^{2+} can serve as a cofactor for the processing activity of full length ASV IN. We next asked if Zn^{2+} , or the other previously untested divalent metals, could function as a cofactor for the processing and joining activities of full length ASV IN. Our initial survey revealed no significant activity with full length IN in a range of concentrations of

Ca^{2+} or Cd^{2+} . Cation Zn^{2+} did support activity, with optimal concentration 2 mM. We then compared the activities in the presence of 2 mM ZnSO_4 or 10 mM MgCl_2 . The full length IN showed significant processing activity (-2 nicking) in the presence of Zn^{2+} ; this activity is slightly reduced compared to that observed with Mg^{2+} , but both are approximately ten-fold lower than that observed with Mn^{2+} as the cofactor. Joining activity can be detected as insertion events into the viral DNA substrate which produce a ladder of products that are longer than the substrate. As expected, joining activity is readily detected in the presence of Mg^{2+} ; however, no significant joining activity was observed with Zn^{2+} .

Many enzymes active in the nucleic acid metabolism have an absolute requirement for divalent cations. However, details of the arrangements of such cations have been reported in only a few of the published structures. In the case of ASV IN, we previously identified binding of Mn^{2+} and Mg^{2+} to a single site between Asp-64 and Asp-121 (site I). Here we report the binding of two Zn ions and two Cd ions to sites I and II. Site II ligands are Asp-64 and Glu-157, suggesting a role for this third member of the invariant triad in the D,D(35)E motif. Two more Zn^{2+} cations are located away from the active site of the enzyme. It is possible that metal ions bound to one or both of these sites could play a role in structural stabilization and/or activity control, as reported for other enzymes (19).

The structural features of Zn ions in the active site agree well with the description of other cocatalytic sites in multi Zn^{2+} , or Mg^{2+} plus Zn^{2+} enzymes discussed by Vallee and Auld (20). As is commonly the case, both ions are close to each other (the distance is 3.62 Å in the high-occupancy structure), and they are bridged by the carboxylate group of an aspartic acid (*e.g.* Asp-64). Similar arrangements have been reported in the past for other enzymes, such as phospholipase C (21) and nuclease P1 (22). In common with these structures, there is also indication of a shared water molecule bridging the two cations, although this putative water is not well ordered in ASV IN.

Our biochemical analyses show that although Zn^{2+} is less effective than Mn^{2+} , it can serve as a cofactor for nicking activity of the isolated catalytic core domain of ASV IN, whereas there is no detectable activity with Mg^{2+} , Ca^{2+} or Cd^{2+} . Comparisons of nicking activities in mixtures of Mn^{2+} and the other cations suggest that both Zn^{2+} and Cd^{2+} bind with higher affinity than Mn^{2+} , and that Mg^{2+} and Ca^{2+} bind with lower affinity. This is consistent with the occupancies of these metals observed in our structural analyses of the catalytic domain. We also observed a sharp peak for the optimal concentration of Zn^{2+} at 2 mM; higher concentrations were inhibitory. The reason for this decreased activity at higher Zn^{2+} concentrations is not yet apparent. However, the most striking result from these studies was the observation that Zn^{2+} supported the endonuclease activity of the catalytic domain whereas Mg^{2+} , the presumed physiologically relevant cation, did not. This observation prompted us to ask whether Zn^{2+} or any of the other catalytic domain-binding cations could serve as cofactors for processing and joining by the full length enzyme. This is the first report that Zn^{2+} can also act as a cofactor for catalysis by a retroviral integrase.

Further structural analyses should help us to understand the basis for these distinct metal preferences. Lastly, the fact that Zn^{2+} binds tightly to the active site but can only support one step in integration may be relevant to the design of active-site inhibitors of retroviral integrases.

REFERENCES

1. Katz, R. A. and Skalka, A. M. (1994) *Annu. Rev. Biochem.* **63**, 133-173
2. Goff, S. P. (1992) *Annu. Rev. Genet.* **26**, 527-544
3. Vink, C., et al. (1993) *Nucleic Acids Res.* **21**, 1419-1425
4. Andrade, M. D. and Skalka, A. M. (1996) *J. Biol. Chem.* **271**, 19683-19686
5. Lee, S. P., et al. (1997) *Biochemistry* **36**, 173-180
6. Lee, S. P. and Han, M. K. (1996) *Biochemistry* **35**, 3837-3844
7. Zheng, R., et al. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13639-13664
8. Kulkosky, J., et al. (1992) *Mol. Cell. Biol.* **12**, 2331-2338
9. Dyda, F., et al. (1994) *Science* **266**, 1981-1986
10. Bujacz, G., et al. (1996) *FEBS Lett.* **398**, 175-178
11. Bujacz, G., et al. (1995) *J. Mol. Biol.* **253**, 333-346
12. Bujacz, G., et al. (1996) *Structure* **4**, 89-96
13. Davies, J. F., II, et al. (1991) *Science* **252**, 88-95
14. Yang, W. and Steitz, T. A. (1995) *Structure* **3**, 131-134
15. Beese, L. S. and Steitz, T. A. (1991) *EMBO J.* **10**, 25-33
16. Rice, P., Craigie, R., and Davies, D. R. (1996) *Curr. Opin. Struct. Biol.* **6**, 78-83
17. Bujacz, G., et al., *The Journal of Biological Chemistry.* **272**, 18161-18168
18. Kulkosky, J., Katz, R. A., Merkel, G., and Skalka, A. M. (1995) *Virology* **206**, 448-456
19. Kim, J. L., et al. (1996) *Cell* **87**, 343-355
20. Vallee, B. L. and Auld, D. S. (1993) *Biochemistry* **32**, 6493-6500
21. Hough, E., et al. (1989) *Nature* **338**, 357-360
22. Volbeda, A., Lahm, A., Sakiyama, F., and Suck, D. (1991) *EMBO J.* **10**, 1607-1618

IDENTIFICATION OF THE MINIMUM SEGMENT IN WHICH THREONINE²⁴⁶ RESIDUE IS PHOSPHORYLATED BY PROTEIN KINASE A FOR THE LUKS-SPECIFIC FUNCTION OF STAPHYLOCOCCAL LEUKOCIDIN

Hirofumi Nariya, Akihito Nishiyama, and Yoshiyuki Kamio*

Department of Applied Biological Chemistry, Faculty of Agriculture, Tohoku University, 1-1
Amamiya-machi, Tsutsumi-dori, Aoba-ku, Sendai 981, Japan

*Corresponding author, Tel: +81-22-717-8779; Fax: +81-22-717-8780;
e-mail: ykamio@biochem.tohoku.ac.jp

Abstract: Staphylococcal leukocidin and γ -hemolysin consist of LukF and LukS for leukocidin and LukF and Hlg2 for γ -hemolysin. In this report, we identify the minimum segment responsible for the LukS specific function of leukocidin. After chemical analysis and homology study of the amino acid sequence of the C-terminal region between LukS and Hlg2, we found a unique 5-residue sequence [242 K²⁴³R²⁴⁴S²⁴⁵T²⁴⁶ in LukS in which 4-residue KRST is identical with that of phosphorylated segment of a protein phosphorylated by protein kinase A. To elucidate whether the 5-residue segment is essential for the LukS function, we created plasmids containing a series of mutant genes corresponding to the 5-residue and expressed them in *Escherichia coli*. The mutant proteins were purified and assayed for their leukocytolytic activity with LukF. The mutant MLS-TS in which the T²⁴⁶ in the 5-residue was replaced by S residue showed leukocidin activity 10 times higher than that of the intact LukS. However, neither mutant MLS-TY nor MLS-TA in which T²⁴⁶ was replaced by Y or A residue, respectively, showed leukocidin activity. The 5-residue segment was found to be deleted in Hlg2. The mutant of Hlg2 in which the 5-residue segment was inserted at the position that the segment is deleted, showed leukocidin activity. The boiled LukS, MLS-TS, and MHS-Z were strongly phosphorylated by [γ -³²P] ATP in the presence of protein kinase A in cell-free system. Thus, we conclude that the 5-residue segment [242 K²⁴³R²⁴⁴S²⁴⁵T²⁴⁶ is pivotal segment of LukS responsible for the LukS-function of the staphylococcal leukocidin.

Staphylococcal leukocidin has been isolated as a bi-component leukocidin from the culture fluids of *Staphylococcus aureus*. It consists of LukS of 32kDa and LukF of 34 kDa, which cooperatively lyse human and rabbit polymorphonuclear leukocytes [1]. We have demonstrated in the previous studies that leukocidin shares one component with the staphylococcal bi-component hemolytic toxin, γ -hemolysin, which consists of Hlg1 and Hlg2 (i.e., Hlg1 is identical with LukF), and that cell specificities of leukocidin and γ -hemolysin are decided by LukS and Hlg2, respectively [1-3]. The deduced amino acid sequences from the genes encoding LukS and Hlg2 indicated 72% identity between them (Fig. 1) [3]. The specificities of the two toxins towards the target cells raise the question of what region(s) of LukS or Hlg2 plays a pivotal role in the LukS- and Hlg2-specific function for leukocidin or

Luk S 1' ANDTEDIGKGS^{*}DI^{*}EIIKRTEDKTSNKGWGTQNIQDFVKDKTKYNKDALILKMQGF^{*}ISSRT
Hlg 2 1' ENKIEDIGGGA--EIIKRTQDITSKRLAITQNIQDFVKDKKYNKDALVVKMQGF^{*}ISSRT
 61' TYNYKKTNHVKAMRWP^{*}FQYNI^{*}GLKTN^{*}DKYVSLIN^{*}YLPK^{*}NK^{*}IE^{*}STN^{*}V^{*}SQ^{*}TL^{*}GYNI^{*}G^{*}GN^{*}FQ
 59' TYS^{*}DLK^{*}KYPY^{*}IKRMI^{*}WPFQY^{*}NI^{*}SLK^{*}TK^{*}D^{*}SN^{*}VD^{*}LIN^{*}YLPK^{*}NK^{*}IDS^{*}ADV^{*}SQ^{*}KL^{*}GYNI^{*}G^{*}GN^{*}FQ
 121' SAPSLGNGS^{*}FNYSK^{*}IS^{*}ISY^{*}TQ^{*}QNYV^{*}SEVE^{*}EQ^{*}NSK^{*}SVL^{*}WG^{*}VK^{*}ANS^{*}FAT^{*}ESG^{*}QK^{*}SA^{*}FD^{*}SDLF
 119' SAPSIGGSGS^{*}FNYSK^{*}IS^{*}ISY^{*}NQ^{*}KNY^{*}VT^{*}EVSQ^{*}NSK^{*}GVK^{*}WG^{*}VK^{*}ANS^{*}FAT^{*}PNG^{*}QV^{*}SAY^{*}DQ^{*}YLF
 181' VGYKPHSKDPRDYFVPDSELPLLVQSGFNPSFIATVSHEKGS^{*}SDTSEFEITYGRNMDVTH
 179' AQ-DPTGPAARDYFVPDNQLPPLIQSGFNPSFITLSHERGKGDKSEFEITYGRNMDATY
 fragment A (D²³⁷-N²⁸⁶)
 241' A IKRST HYGNSYLDGHRVHNAFVN^{*}RNY^{*}TVK^{*}YEV^{*}NW^{*}KTHEIKV^{*}KGQN 286
 238" A ----- YVTRHRLAVDRKHDAFKNRNVTVK^{*}YEV^{*}NW^{*}KTHEVK^{*}IKSITPK 280
 fragment B (D²³⁵-K²⁸⁰)

Fig. 1. Comparison of the deduced amino acid sequence of LukS with that of Hlg2. Gross dissimilarity between them is indicated by box. Identical and related residues are indicated by stars and dots, respectively. Dashed lines indicate the deleted amino acids. Arrowheads indicate the cleavage site of CNBr.

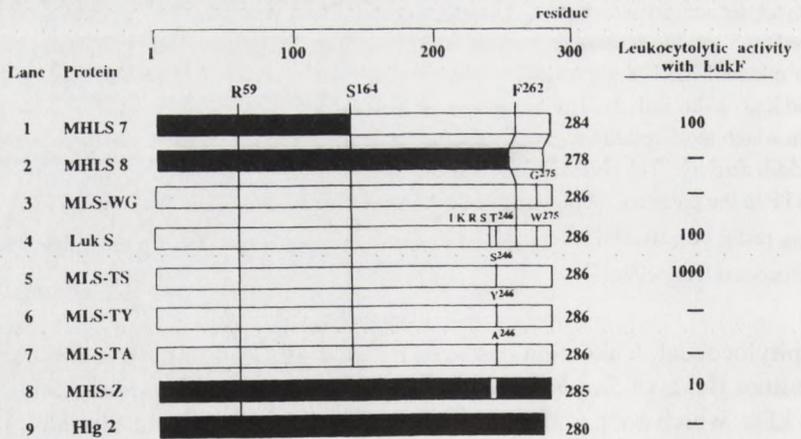


Fig. 2. Schematic representation of LukS, Hlg2 and mutant proteins, and the resulting leukocidin activity. White and black boxes indicate the LukS- and Hlg2-segments, respectively. White box in MHS-Z (lane 8) represents the 5-residue IKRST segment inserted. Percentage leukocytolysis indicates the activity compared with that of intact LukS (lane 4). A minus indicates no detectable activity.

γ -hemolysin activities. To answer this question, previously, we created a series of chimeric genes (*lukS/hlg2*) and expressed them in *Escherichia coli*. From the results obtained, we concluded that there is an essential region for LukS-specific function within C-terminal 122-residue segment (between S¹⁶⁴ and the C-terminus) of LukS [4] (see Fig. 2, lane 1). Here, we identify the minimum amino acid residues in the C-terminal 122-residue segment of LukS responsible for the leukocytolytic activity. This report indicates that the 5-residue segment I²⁴²KRST²⁴⁶, in which T residue was phosphorylated by protein kinase A in cell-free system, is pivotal region of LukS responsible for the LukS-specific function of the staphylococcal leukocidin.

After homology study of the deduced amino acid sequence from the structural genes for LukS and Hlg2, we found a unique 5-residue sequence I²⁴²KRST²⁴⁶ of which 4-residue KRST is identical with a recognition site of protein phosphorylated by protein kinase A. This segment was deleted in Hlg2 (see box in Fig. 1). To obtain the direct evidence for the presence and absence of the 5-residue segment in LukS and Hlg2, respectively, LukS and Hlg2 were degraded with CNBr by the method described previously [5] and a [Asp²³⁷-Asn²⁸⁶] fragment (fragment A) in LukS and [Asp²³⁵-Iys²⁸⁰] fragment (fragment B) in Hlg2 were isolated by using HPLC. The N-terminal 12-residue of the fragments A and B were determined to be D²³⁷VTHAIKRSTHY²⁴⁸ and D²³⁵ATVAYVTRHR²⁴⁶, respectively. The data reconfirmed that the 5-residue segment in LukS is deleted in Hlg2.

To study whether or not the 5-residue segment (IKRST) is essential for the LukS-function, we created a series of mutant genes in the segment by the overlapping-extension method [8], and inserted them into pUC119 vector DNA. The three pairs of 53- or 54-oligonucleotides for replacing T²⁴⁶ residue in the segment by S, Y, or A residue, [TS-1 and TS-2], [TA-1 and TA-2], and [TY-1 and TY-2], respectively, were synthesized and used as primers for amplifying the segments using PCR.

(i) TS-1, 5'-GCCATTAAAAGATCATCACATTATGGCAACAGTTATTTAGACGGACATAGAGTC-3' and TS-2, 5'-ACTGTTGCCATAATGTGATGATCTTTTAAATGGCATGAGTGACATCCATGTTTC-3' (double underlined nucleotide sequences in both primers correspond to the 5-residue, IKRSS in the segment; (ii) TA-1, 5'-GCCATTAAAAGATCAGCCATTATGGCAACAGTTATTTAGACGGACATAGAGTC-3' and TA-2, 5'-ACTGTTGCCATAATGGGCTGATCTTTTAAATGGCATGAGTGACATCCCATGTTTC-3' (double underlined nucleotide in both primers are corresponding to the 5-residue, IKRSA), and (iii) TY-1, 5'-GCCATTAAAAG-ATCATATCATTATGGCAACAGTTATTTAGACGGACATAGAGTC-3' and TY-2, 5'-ACTGTTGCCATAATGATATGATCTTTTAAATGGCATGAGTGACATCCATGAATC-3' (double underlines indicate the nucleotide sequences corresponding to 5-residue, IKRSY). The resulting 1.5 Kbp [*Hind*III-*Hind*III] fragments were ligated into the *Hind*III site of pUC119. The mutagenesis and

orientation of the fragments was confirmed by DNA sequencing and restriction endonuclease analysis, respectively. As a result of these manipulations, three different plasmids that contained mutant genes corresponding to the 5-residue segment were obtained and they were designated as pMLS-TS, pMLS-TA, and pMLS-TY. The mutant proteins expressed in *E. coli* harboring the appropriate plasmid were prepared from the sonicated extract of the cells from one liter of culture and purified to the degree of electrophoretic homogeneity by the methods described previously [6], and the leukocytolytic activity for human polymorphonuclear leukocytes was measured according to the methods described previously [7, 8]. After measuring leukocytolytic activity of the purified mutant components in the presence of LukF, the following findings became evident: (I) Mutant proteins MLS-TS, showed 10 times higher leukocytolytic activity than that of intact LukS (Fig. 2, lanes 4 and 5). (II) Neither mutant protein MLS-TA nor MLS-TY has any leukocidin activity (Fig. 2, lanes 6 and 7). The findings clearly indicate that the T²⁴⁶ residue of the 5-residue segment of LukS is pivotal for the LukS function. It is known that, for the recognition site-motifs of the protein phosphorylated by protein kinase A, serine residue is more suitable than threonine as a final amino acid residue in the 4-residue, R-R/K-X-S/T [9]. Taken together, there might be some relationships between the phosphorylation of the residue T in the segment, if any, and the leukocytolytic function of LukS.

If leukocidin-specific activity is decided by the 5-residue segment IKRST of LukS, leukocidin activity might be endowed into Hlg2 by being inserted the 5-residue segment of LukS into between A²³⁸ and Y²³⁹ residues of Hlg2 (Fig. 1). Accordingly, we created the mutant plasmids pMHS-Z using the following primers, 5'-TACATATGCTATTTAAAAGATCAACGT ACGTGACAAGACATCGTTTACCCGTTGATAGAAAACATGATGC-3' and 5'-TTGTCACGTACGTTGATC-TTTTAATAGCATATGTAGCATCCATGTTTCTGCCGTAAGTGATTTCAAAC-TCGC-3' (double underlined nucleotide sequences in both primers are corresponding to the 5-residue, IKRST). The MHS-Z was expressed in *E. coli* DH5 α (pMHS-Z) and then purified from the sonicated extracts from the cells according to the methods described above. The MHS-Z showed leukocytolytic activity with LukF, although not full activity (Fig. 2, lane 8). Thus, we concluded that the 5-residue IKRST is the minimum segment essential for the LukS-specific function of staphylococcal leukocidin.

The findings described above suggest that LukS and the mutant proteins MLS-TS and MHS-Z possessing the 5-residue segment are phosphorylated by protein kinase A. Accordingly, we examined the possibility using a cell-free system [10]. LukS or mutant protein MLS-TS was incubated with protein kinase A from bovine heart in the presence of [γ -³²P]-ATP, MgCl₂, and EGTA, and the reaction mixture was applied to SDS-PAGE. The gel was scanned by the image scanner. The native LukS, MLS-TS, and MHS-Z were slightly phosphorylated. If they were boiled for 5 min and used as substrates, their phosphorylation was increased about 50-100 times

higher than that of native ones. The boiled MLS-TS was phosphorylated 4 times higher than the boiled LukS. The mutant proteins MLS-TA and MLS-TY, in which the T as a phosphorylated residue in the recognition site motif had been replaced by A and Y, respectively, were also slightly phosphorylated. However, the intensity of their phosphorylation did not increase regardless of their denature by boiling. It is known that LukS binds specifically to monosialoganglioside GM₁ (GM₁) on the leukocyte membrane [11]. We previously clarified that W²⁷⁵ residue of LukS is an essential amino acid for its binding to GM₁ and that this binding is accompanied by a conformational change of LukS [4, 7]. Accordingly, we examined whether or not the phosphorylation of the native LukS was activated by adding GM₁ in the reaction mixture. However, its phosphorylation was not activated by GM₁. These findings indicated that T²⁴⁶ residue in the segment of native LukS and MLS-TS is not exposed to the protein surface even upon binding to GM₁ in the reaction mixture. The data also strongly suggest that the phosphorylation of T²⁴⁶ residue in the 5-residue segment of LukS is important for the LukS function of leukocidin, and that the lack of leukocidin activity in Hlg2 is due to the absence of the 5-residue segment in itself.

Fink-Barbacon *et al.* reported that leukocidin forms pores inducing an increase in the free intracellular Ca²⁺ which may activate human polymorphonuclear leukocytes [12]. However, no direct evidence on the complex formation of the leukocidin components on the human leukocytes is available. Recently, we reported that LukF and Hlg2 of γ -hemolysin assemble into a ring-shaped 195 kDa complex in a molar ratio of 1:1, which form a membrane pore with a functional diameter of 2.1-2.4 nm [13]. Using our established systems, we examined complex formation of LukF with either LukS, MLS-TS, MLS-TA, MLS-TY, or MHS-Z on the human polymorphonuclear leukocytes. Our data showed that all of the mutants as well as LukS in combination with LukF assembled into above 100 kDa complex, which may form a pore on the surface of human polymorphonuclear leukocytes. LukS alone did not assemble on the cell surface (data not shown). From these findings, we could distinguish the leukocytolytic function from the complex formation which forms the membrane pore on the surface of human polymorphonuclear leukocytes.

We monitored the change of morphology of the cells of human polymorphonuclear leukocytes under a phase contrast microscope. Intact cells became swollen after the incubation at 37°C with LukF and Hlg2 for 10 min. However, any lysed cell was not observed by the incubation for more 20 min. The data indicated that LukF and Hlg2 cooperatively caused only swelling of the human polymorphonuclear leukocytes without lysis.

References

- [1] Tomita, T., and Kamio, Y. (1997) **Biosci. Biotech. Biochem.**, 61, 565-572.
- [2] Kamio, Y., Rahman, A., Nariya, H., Ozawa, T. and Izaki, K. (1993) **FEBS Lett.**, 321, 15-18.

- [3] Rahman, A., Izaki, K. and Kamio, Y. (1993) **Biosci. Biotech. Biochem.**, 57, 1234-1236.
- [4] Nariya, H., and Kamio, Y. (1995) **Biosci. Biotech. Biochem.**, 59, 1603-1604.
- [5] Choorit, W., Kaneko, J., Muramoto, K. and Kamio, Y. (1995) **FEBS Lett.**, 357, 260-264.
- [6] Nariya, H., Asami, I., Ozawa, T., Beppu, Y., Izaki, K. and Kamio, Y. (1993) **Biosci. Biotech. Biochem.**, 57, 2198-2199.
- [7] Nariya, H., Izaki, K. and Kamio, Y. (1993) **FEBS Lett.**, 329, 219-222.
- [8] Horton, R., Hunt, H., Ho, S., Pullen, J. and L. Pease, L. (1989) **Gene**, 77, 61-68.
- [9] Edelman, A. M., Blumenthal, D. K., and Krebs, E. G. (1987) **Annu. Rev. Biochem.**, 56, 567-613.
- [10] Kennelly, P. J., and Krebs, E. G. (1991) **J. Biol. Chem.**, 15555-15558.
- [11] Noda, M., Kato, I., Hirayama, T. and Matuda, F. (1980) **Infect. Immun.**, 29, 678-684.
- [12] Finck-Barbacon, V., Duportail, G., Meunier, O. and Colin, D. A. (1993) **Biochim. Biophys. Acta**, 1182, 275-282.
- [13] Sugawara, N., Tomita, T. and Kamio, Y. (1997) **FEBS Lett.**, 410, 333-337.

MOLECULAR MODELLING OF G PROTEIN-COUPLED RECEPTOR (GPCR)-BIOLIGAND INTERACTIONS

Jerzy Ciarkowski^{1,2}, Cezary Czaplewski¹, Rajmund Kaźmierkiewicz¹, Ewa Politowska^{1,3}

¹Faculty of Chemistry, University of Gdańsk, ul. Sobieskiego 18, 80-952 Gdańsk, Poland; ²e-mail jurek@sun1.chem.univ.gda.pl

³Academic Computer Centre in Gdańsk TASK, ul. Narutowicza 11/12, 80-952 Gdańsk, Poland

Abstract: G protein-coupled receptors (GPCRs), the largest superfamily of *integral* membrane proteins, feature the following common properties: (i) They all exhibit 7 hydrophobic α -helices of length of $\sim 38\text{\AA}$ (25 amino acids, 7 turns) along a single chain. The consecutive helices cross the lipid bilayer to and fro starting from the extracellular side, to form a heptahelical transmembrane domain (7TM¹). This arrangement implies 6 inter-helical loops, whereof the even ones plus the *N*-terminus form the receptor's extracellular while the odd ones plus the *C*-terminus - its intracellular domain. (ii) All GPCRs are stimulated by extracellular signals of miscellaneous character. (iii) The stimulated GPCRs pass allosterically the extracellular message to the cytosolic *peripheral* heterotrimeric G proteins, the transducers, which mediate a downstream signal passage to various cellular second messenger systems. A current status of structural studies on GPCRs, consisting of low 7.5 \AA resolution experimental structures and supplementary molecular modelling, is presented. Some results of authors' own work on modelling essential interactions between the vasopressin renal receptor and its agonists vasopressin (AVP) and both the peptide desGly²-[Mca¹,D-Ile²,Ile⁴]AVP and the nonpeptide antagonist OPC-31260 are discussed.

The function of G protein-coupled receptors (GPCRs)

GPCRs are *integral* membrane proteins that receive, process and forward extracellular signals to cytosolic transducers, the heterotrimeric G α G β γ *peripheral* membrane proteins (G proteins). Given the GPCR message, a G protein exchanges of GTP for GDP, then dissociates into its active forms, the G α .GTP and G β γ subunits, which stimulate various cellular effectors like ion channels or enzymes (e.g. adenylyl cyclase, AC, or phospholipase C type β , PLC β , or else), giving rise to the cytosolic second messengers like Ca²⁺ ions or small molecules, respectively (e.g. cAMP or DAG+IP₃, or else, respectively), regulating cellular performance [1,2], see Figure 1. Over 1000 GPCR sequences are currently described [3].

¹**Abbreviations:** Natural amino acids and nucleotides are given their standard one- or three-letter codes. 7TM heptahelical transmembrane domain; TM1-TM7 transmembrane helices 1 to 7; **G protein** GTP-binding protein; **GPCR** G protein-coupled receptor; **AC** adenylate cyclase; **PLC β** phospholipase C type β ; **DAG** diacylglycerol; **IP₃** inositol triphosphate; **AVP** [Arg⁸]Vasopressin; **OT** oxytocin; **V2R** vasopressin V2 receptor; **V1a(b)R** vasopressin V1a(b) receptor; **OTR** oxytocin receptor; **RD** rhodopsin; **BRD** bacteriorhodopsin; **Mca** β , β -cyclopenta-methylene- β -mercaptpropionyl; **OPC-31260** [5-dimethylamino-1-{4-(2-methyl-benzoyl-amino)- benzoyl}-2,3,4,5-tetrahydro-1H-benzazepine]; **E(I)D** extra(intra)cellular domain; **E(I)L1**, **E(I)L2**, etc. extra(intra)cellular loop 1, extra(intra) cellular loop 2, etc, resp.; **3D** three dimensional; **CSA** constrained simulated annealing.

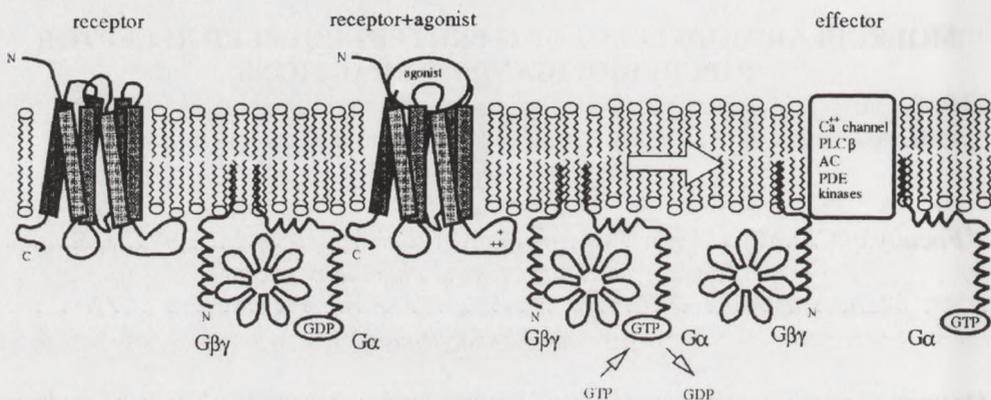


Figure 1. GPCR signal transduction system. The activated receptor catalyses the exchange of GTP for GDP in the $G_{\alpha}G_{\beta\gamma}$ heterotrimer. This triggers the dissociation of the heterotrimer into $G_{\alpha}\cdot\text{GTP}$ and $G_{\beta\gamma}$, which activate effectors: enzymes and/or ion channels. Native G_{β} never separates of helical G_{γ} and has a structure reminiscent of a 7-blade propeller [4] as reflected on the scheme.

Besides *functional*, all GPCRs also share a *structural* similarity, consisting of 7 hydrophobic α -helices of $\sim 38\text{\AA}$ long (25 amino acids, 7 turns). The consecutive helices traverse the lipid bilayer to and fro starting from the external side, to form a heptahelical transmembrane domain (7TM). This arrangement simultaneously implies 6 inter-helical loops, whereof the even ones plus the *N*-terminus form the receptor's extracellular domain (ED) while the odd ones plus the *C*-terminus make its intracellular domain (ID), see Figure 1. From here on, we will be dealing exclusively with the most abundant and best studied GPCR superfamily, known as the rhodopsin(RD)-like or classical one. Other major GPCR superfamilies include the secretin- and the metabotropic glutamate-type receptors, both groups considerably less abundant than the classical one and, apart of the 7TM, having no structural homologies with it.

GPCRs are individually tuned for very diverse *primary* extracellular signals, ranging from biogenic amines (adrenaline, noradrenaline, acetylcholine, dopamine, histamine, serotonin, etc.) through tastes and smells (received by the most abundant gustatory and olfactory GPCRs, respectively), peptide hormones (OT, AVP, angiotensin, bradykinin, tachykinins and many more), proteins (chemokins, glykoprotein gonadotropins, other), to the photons, which are the primary signals for the landmark GPCR rhodopsin (RD). Despite so diverse primary signals all GPCRs use a common transduction mechanism, embodied in coupling with a *peripheral* heterotrimeric $G_{\alpha\beta\gamma}$ (G) protein, being simultaneously a GTPase [1,2,4,5]. $G_{\alpha\beta\gamma}$ activated by the GPCR's intracellular domain (ID) exchanges of GTP for GDP. This induces considerable conformational changes in the three switch regions within the G_{α} subunit [6,7] and leads to the dissociation of $G_{\alpha}G_{\beta\gamma}$ into the active forms, the $G_{\alpha}\cdot\text{GTP}$ and $G_{\beta\gamma}$ subunits. Since now on, a diversity occurs again as there are about 20, 5 and 10 isoforms of G_{α} , G_{β} and G_{γ} , respectively, making some $20 \times 5 \times 10 = 1000$ theoretically possible combinations of $G_{\alpha}G_{\beta\gamma}$ and respective further actions. The

downstream effectors are being stimulated until GTP on G_{α} is not hydrolysed to GDP. Once done, this is a signal for G_{α} -GDP and $G_{\beta\gamma}$ to reassociate into the heterotrimer $G_{\alpha}G_{\beta\gamma}$ and the entire cycle repeats. Typically, a single receptor stimulation triggers a several hundred to several thousand G protein cycles [1].

GPCR structure

Contrary to the G protein structures, known at various steps of the cycle [6,7] and also implying a definite GTPase mechanism [8,9], the GPCR atomic-resolution structures still await their elucidations. Thus far, only a general shape of RD at a resolution of at best 7.5Å parallel and 16.5Å perpendicular to the membrane is known due to the electron criomicroscopy [10,11]. Hence, all GPCR structural considerations are currently limited to molecular modelling.

v2rMLMAS	TTSVAVPGHPS	LPSLPSNSSQ	ERPLDTRDPL	LARAELALLS	IVFVAVALSN	GLVLAALARR	65	
v1ar	.MRLSAGPDA	GPSGNSSPW	PLATGAGNTS	REAEALGEGN	GPPRDVRNEE	LAKLEIAVLA	VTFAVAVLGN	SSVLLALHRT	79
v1brMD	SGPLWDANPT	PRGTLSAPNA	TTPWLGRDEE	LAKVEIGVLA	TVLVLATGGN	LAVLLTLGQL	62	
otrMEGALAA	NWSAEANAS	AAPPGAEGNR	TAGPPRNEA	LARVEVAVIC	LILLLALSGN	ACVLLAL.	65	
v2r	GRRGHWAPEIH	VFIGHLCLAD	LAVALFQVLP	QLAWKATDRF	RGPDALCRAV	KYLQVMGYA	SSYMLLAMEL	DRHRAICRAM	145
v1ar	PR..KTSRMH	LFIRHLSLAD	LAVAFFQVLP	QMCWDI TYRF	RGPDWLCRVV	KHLQVGFMEFA	SAIYMLVMTA	DRYIIVACHP.	156
v1br	GR..KRSRMH	LFVLHLALTD	LAVALFQVLP	QLLWDI TYRF	QGPDLRCRAV	KYLQVLSMFA	STYMLLAMEL	DRYLAVACHP.	139
oxyr	TTRQKHSRLF	FFMKHLSIAD	LVVAVFQVLP	QLLWDI TFRF	YGPDLLCRLV	KYLQVVGMEFA	RTYLLMLLSL	DRCLAI CQPL	145
v2r	LAYRHGSGAH	WNRPVLVAVFA	FSLLLSLPQL	FIF..AQRNV	EGSGSVTDWCW	ACFAE PWGR	TYVTWIALMV	FVAFTLGIAA	223
v1ar	LKTQQPARR	SRLMIAAAWV	LSFVLSIPQY	FVFS..MIEV	NNVTKARDCW	ATFIQPWGSR	AYVTWMTGGI	FVAPVIVLGT	234
v1br	LRSLQQPGQS	TYLLIAAPWL	LAIFSLPQV	FIFSL..REV	IQSGSVLDCW	ADFGFPWGR	AYLVTWTLAI	FVLPVMTLTA	217
oxyr	RSLR...RRT	DRLAVLATWL	GCLVASAPQV	HIFSLRE...	.VADGVFDCW	AVFIQVWGPK	AYITWITLAV	YIVFVIVLAT	218
v2r	<u>QVLFIREIH</u>	ASLVPGP...SERPGR	RRGRRTGSPG	EGAHVSAAVA	KTVRMTLVIV	VVYVLCWAPF	287	
v1ar	<u>CYGFICYNIW</u>	CNVRGKTASR	QSK.....	GAEQAGVAFQ	KGFL LAPCVS	SVKSI SRAKI	RTVKMTFVIV	TAYIVCWAPF	307
v1br	<u>CYSLICHEIC</u>	KNLKVKTQAW	RVGGGWRTW	DRPSPSTLAA	TTRGLPSRVS	SINTI SRAKI	RTVKMTFVIV	LAYIACWAPF	297
oxyr	<u>CYGLISFKIW</u>	QNLRLKTA.AAAA	EAEPEGAAGD	GGRVALARVS	SVKLI SKAKI	RTVKMTFVIV	LAFIVCWTFE	291
v2r	<u>FLVQLWAARD</u>	...PEAPLEG	APFVLLMLLA	SLNSCCNPWI	YASFSSSVSS	ELRSLCCAR	GRTPPSLGPQ	DESCTTASSS	364
v1ar	<u>FIIQMWSVWD</u>	PMSVWTESEN	PTITITALLG	SLNSCCNPWI	YMF FSGHLLQ	DCVQSFPCQ	NMKEKFNED	TDMSRRRTQF	387
v1br	<u>FSVQMWSVWD</u>	KNAPDESDTN	VAFTISMLLG	NLNSCCNPWI	YMGFNSHLLP	RPLRHLCAGC	GPQPRMRRRL	SDGSLSSRHT	377
oxyr	<u>FEVQMWSVWD</u>	AN...APKEA	SAFTIIVMLLA	SLNSCCNPWI	YMLFTGHLFH	ELVQRFLCCS	ASYLKGRRIG	ETSASKKSNS	368
v2r	LAKDTSS...	371	
v1ar	YSNNRSP TNS	TGMWKDSPKS	SKSIKFI PVS	T.....	418	
v1br	TL LTRSSCPA	TL SLSLSLTL	SGRPRPEESP	RDLELADGEG	TAETIIF...	424	
oxyr	SSFVLSHRSS	SQRSCSQPST	A.....	389	

Figure 2. Multiple sequence alignment of human V2R, V1aR, V1bR and OTR. The transmembrane helices TM1-TM7 are underlined. The *subfamily-invariant* residues are in *italic* and the **universally invariant/conservative** (~19% of 7TM) residues are in **bold**.

All approaches toward this task start from a multiple sequence alignment [3] and/or a hydrophobicity profile for making a selection of the TM1-TM7 helices along the sequence, see Figure 2. Subsequently a 3D model of 7TM is built and (sometimes) ID and ED added. Regarding the 7TM model, we favour the recent approach by Herzyk and Hubbard [12] using: (i) a convincing multiple sequence alignment of all GPCRs [13]; (ii) maximal extents of minimal loops 1-6 over the whole superfamily; (iii) numerous distance, position and orientation restraints from biophysical experiments on rhodopsin and its mutants; and (iv) a low-resolution 9Å structure of bovine RD from electron criomicroscopy [14]. Given these four groups of data as an input to a constrained simulated annealing (CSA), they received an RD

7TM domain self-consistent to 1.6Å at the C^α trace level, which provided for a reasonable template to all classical GPCRs [15]. While sequence homologies in 7TM extend over the whole superfamily, the extra and intracellular domains have no homologies at all, unless among very close relatives, see Figure 2. Accordingly, there is no rational basis for modelling the loops and the *N*- and *C*-termini.

Contrary to the model of Herzyk and Hubbard [12,15], most former models have used a low-resolution structure of bacteriorhodopsin (BRD) [16], a bacterial 7TM proton pump, being not a GPCR and having no homologies with it. Current comparisons of low to intermediate resolution structures of RD [17] to BRD [18,19], respectively, reveal essential differences in their 3D shapes, thus turning BRD improper as a 3D template for 7TM.

Vasopressin V2 receptor (V2R)-bioligand interactions

The nonapeptide hormone vasopressin (CYFQNCPRG-NH₂, AVP) regulates the renal water absorption via the interaction with the V2 receptor (V2R). V2R and the other three structurally related, the vascular V1a and pituitary V1b AVP receptors (V1aR and V1bR, respectively), and the oxytocin receptor (OTR), form a discrete GPCR subfamily. All four receptors share a high degree of sequence identity [20,21], in particular (up to 87%) in their transmembrane helices TM2, TM3, TM6 and TM7, in the extracellular loop EL1 and the *C*-terminal part of EL2, see Figure 2, warranting an assumption as to a similar recognition and binding modes. Major differences occur in the intracellular loops (IL)s and correlate with receptors' linking to specific secondary signal systems: V1aR, V1bR and OTR to the G_{q/11} protein/PLCβ tract while V2R to the G_s protein/cAMP system [1].

The V2R 7TM domain was modelled using the procedures mentioned above [12,15] while the loops and the *N*-terminus were built using the Sybyl software [22]. Initial ligand docking was attained in several ways, always respecting a complementarity in the electrostatic potentials of the V2R cleft, see below, and the ligand. The systems were equilibrated using a CSA protocol, with all but the 7TM C^α atoms free to move. Optimal ligand docking modes were chosen using the ligand/receptor interaction energies and structure-activity data [23,24] as the selection criteria. Details of computing and analyses are described elsewhere [25].

From Figure 3 it is seen that any GPCR modelled to the RD template [12,15] has a ~21Å deep cavity on the extracellular side, surrounded by TM3-TM7, with a narrower extension towards TM2. The cavity ends up on a floor from the hydrophobic residues TM3:M123, TM4:L170, TM5:V213,F214 and TM6:W284, F287,F288 in V2R. The cleft is large enough to accommodate the pressin ring (CYFQNC) of AVP and even more so to fit the OPC-31260 antagonist. Most of the simulations, whether with a peptide ligand or not, converged to the docking modes typical of V2R/AVP [25]. However, OPC-31260 as much thinner than the AVP pressin ring, cannot fill the entire V2R cleft and it adheres to the front side of the TM3-TM7 cavity in its most preferred arrangements, see Figure 3C [26].

In Figure 3 all V2R interacting residues are marked so that the significant receptor-ligand interactions could be seen. Both the V2R/peptide complexes develop

References

- [1] T.P. Iismaa, T.J. Biden, J. Shine, *G Protein-Coupled Receptors*, Springer-Verlag, Heidelberg, Germany, 1995.
- [2] S. Watson, S. Arkinstall, *The G-protein linked receptor. Facts Book*, Academic Press Ltd., London 1994.
- [3] G. Vriend, GPCR DataBase © 1996 (<http://swift.embl-heidelberg.de/7tm/>).
- [4] S.R. Sprang, *Annu. Rev. Biochem.*, 1997, **66**, 639.
- [5] A. Wittinghofer, *Structure*, 1996, **4**, 357.
- [6] D.G. Lambright, J. Sondek, A. Bohm, N.P. Skiba, H.E. Hamm, P.B. Sigler, *Nature*, 1996, **379**, 311.
- [7] M.A. Wall, D.E. Coleman, E. Lee, J.-A. Iñiguez-Lluhi, B.A. Posner, A.G. Gilman, S.R. Sprang, *Cell*, 1996, **83**, 1047.
- [8] J.P. Noel, H.E. Hamm, P.B. Sigler, *Nature*, 1993, **366**, 654.
- [9] D.E. Coleman, A.M. Berghuis, E. Lee, M.E. Linder, A.G. Gilman, S.R. Sprang, *Science*, 1994, **265**, 1405.
- [10] G.F.X. Schertler, P.A. Hargrave, *Proc. Natl. Acad. Sci.*, 1995, **92**, 11578.
- [11] V.M. Unger, P.A. Hargrave, J.M. Baldwin, G.F.X. Schertler, *Nature*, 1997, **389**, 203.
- [12] P. Herzyk, R.E. Hubbard, *Biophys. J.*, 1995, **69**, 2419.
- [13] J.M. Baldwin, *EMBO J.*, 1993, **12**, 1693; *Curr. Opin. Cell Biol.*, 1994, **6**, 180.
- [14] G.F.X. Schertler, C. Villa, R. Henderson, *Nature*, 1993, **362**, 770.
- [15] M.C. Peitch, P. Herzyk, T.N.C. Wells, R.E. Hubbard, Swiss-Model, an automated knowledge-based protein modelling server at Glaxo Wellcome Research and Development S.A. in Geneva (<http://expsy.hcuge.ch/swissmod/SWISS-MODEL.html>).
- [16] R. Henderson, J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckmann, K.H. Downing, *J. Mol. Biol.* 1990, **213**, 899.
- [17] V.M. Unger, P.A. Hargrave, J.M. Baldwin, G.F.X. Schertler, *Nature*, 1997, **389**, 203.
- [18] Y. Kimura, D.G. Vassilyev, A. Miyazawa, A. Kidera, M. Matsushima, K. Mitsuaka, K. Murata, T. Hirai, Y. Fujioshi, *Nature*, 1997, **389**, 206.
- [19] E. Pebay-Peyroula, G. Rummel, J.P. Rosenbusch, E.M. Landau, *Science*, 1997, **277**, 1676.
- [20] T. Sugimoto, M. Saito, S. Mochizuki, Y. Watanabe, S. Hashimoto, H. Kawashima, *J. Biol. Chem.*, 1994, **269**, 27088.
- [21] Y. De Keyser, C. Auzan, F. Lenne, C. Beldjord, M. Thibonnier, X. Bertagna, E. Clauser, *FEBS Lett.*, 1994, **356**, 215.
- [22] SYBYL, v. 6.1, 1994, Tripos Inc., St. Louis, MO, U.S.A.
- [23] M. Manning, W.H. Sawyer, *J. Recept. Res.* 1993, **13** 195.
- [24] H. Ogawa, H. Yamashita, K. Kondo, Y. Yamamura, H. Miyamoto, K. Kan, K. Kitano, M. Tanaka, K. Nakaya, S. Nakamura, T. Mori, M. Tominaga, Y. Yabuuchi, *J. Med. Chem.*, 1996, **39**, 3547.
- [25] C. Czaplewski, R. Kaźmierkiewicz, J. Ciarkowski, *J. Comp.-Aided Molec. Design*, 1998, **12**, in press.
- [26] C. Czaplewski, R. Kaźmierkiewicz, J. Ciarkowski, *Lett. Peptide Sci*, 1997, **4**, in press.
- [27] B. Mouillac, B. Chini, M.-N. Balestre, J. Elands, S. Trumpp-Kallmeyer, J. Hoflack, M. Hibert, S. Jard, C. Barberis *J. Biol. Chem.*, 1995, **270**, 25771.

α -HELIX FORMATION AS AN EARLY STEP OF PROTEIN FOLDING: A CRITICAL EXAMINATION OF THE HELIX-COIL TRANSITION THEORIES

Andrzej Bierzyński

Institute of Biochemistry and Biophysics, Polish Academy of Sciences
02-106 Warszawa, ul. Pawińskiego 5a, POLAND
e-mail: ajb@ibbrain.ibb.waw.pl

Abstract: Elucidation of the α -helix formation process is very important for understanding of how numerous protein polypeptide chains fold in their native structures. Helix-coil transition has been studied intensively and now seems to be well understood. Nevertheless, we have found that a very short pre-nucleated helical structure formed by residues AlaAlaAlaGlnNH₂ attached to the C-terminus of a calcium-binding peptide loop is extremely stable in water solutions even at room temperature, much more stable than the existing theories would predict. We explain this phenomenon by the effects of hydration of the helix ends. Implications of our findings for kinetics and thermodynamics of nucleation and propagation of a nascent helix are discussed.

In a number of protein polypeptide chains short segments have been found which even after they are cleaved off and separated from the rest of the molecule show high propensity for conformations they adopt in their native states. The most striking examples of such segments are: the nine residue fragment of haemagglutinin forming a high populated β -turn in water solutions (1) and C-peptide of RNase A - the 1-13 fragment of the protein maintaining about 30% of the native α -helical conformation (2).

Such segments are called "autonomous folding units" (3) and play a very important, or even decisive, role in folding processes of many proteins. The most spectacular corollary of this statement is provided by the famous Kato and Anfinsen experiment proving that RNase S protein can fold in its native structure only in the presence of S-peptide (4).

Among candidates for the autonomous folding units α -helical segments of proteins are of special interest for the following reasons:

1) The α -helix formation is a fast, several orders of magnitude faster than the protein folding process (5).

2) It was shown that many isolated short helical segments of proteins are characterized by very high propensity for α -helix formation (6,7).

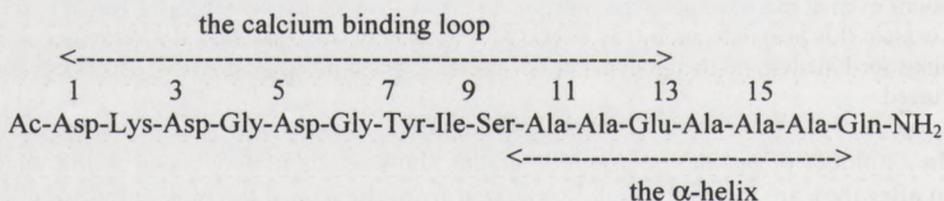
3) Such segments are localized precisely within sequences of folding polypeptide chains. They are initiated and terminated by the so called "stop signals" - one or a few residues that stabilize the helix ends but, consequently, prevent further propagation of the helix.

For these reasons the kinetics and thermodynamics of α -helix formation is of particular importance for our knowledge of the protein folding process.

Since the discovery of the α -helical structure the helix-coil transition in polypeptides has been studied intensively, particularly in the last 15 years and, by now, the process seems to be quite well understood (8). The theoretical basis for the helix-coil transition has been formulated by Zimm and Bragg (9) and Lifson and Roig (10). Using the modified theoretical descriptions and experimentally determined parameters, population of the α -helix conformation can be evaluated for any polypeptide chain of a given amino acid sequence (8).

Nevertheless, our studies of some short peptide systems with pre-nucleated α -helical structure (Siedlecka et al. in preparation) led to some findings incompatible with the commonly accepted views and implying that the formation of a nascent α -helix and its propagation depend on effects not taken into account so far.

The polypeptide studied by us consists of a calcium-binding loop analogous to the IIIrd loop of calmodulin and a short peptide sequence AlaAlaAlaGlnNH₂ with a high propensity for α -helix formation (11) attached to its C-terminus:



The loop coordinates La³⁺ ions (12,13) with a high binding constant (10⁵ M⁻¹) and adopts a very rigid structure (14) with three last residues Ala10, Ala11, and Glu12 fixed in the α -helical conformation as in the native protein (15).

Such a helix nucleation site should stabilize very efficiently the helical conformation of residues Ala13-Gln16. Indeed, NMR and CD measurements confirm the α -helical structure of this segment. Unexpectedly, though, the helix stability appeared to be amazingly high. The CD signal of the peptide at 222 nm, commonly accepted as a measure of α -helix content does not change with increasing trifluoroethanol concentration up to 60 %. This shows that the helix population is close to 100% - at least 95% taking into account measurement errors - not only at low (1^oC) but even at room (25^oC) temperature.

In our system the statistical weight of an n-residue helix starting from Ala13 is given, according to the modified Lifson-Roig theory, by the equation:

$$k_n = \prod_{i=13}^{12+n} w_i c_{(13+n)}$$

where w_i is the helix propagation parameter for residue number i and $c_{(13+n)}$ the C-cap parameter defined by the residue flanking the helical segment.

Using the highest values reported in the literature for w of alanine and glutamine (16,17) and c parameters determined for glutamine and the NH₂ group (18), the calculated helix content of Ala13-Gln16 segment should not exceed 77% at 1^oC and

drop with increasing temperature. Therefore, in our system the helix is stabilized by some additional contribution to its free energy that can be estimated to be no smaller than -3 kcal/mole.

What this contribution come from ? Our molecular dynamics calculations show that it most probably corresponds to a difference between hydration energy of three NH and CO groups at the helix ends separated in Ala10-Gln16 helix but merged with each other in the helix nucleus Ala10-Glu12. The hydration energy of the helix ends is almost twice as large as that of three non-hydrogen bonded peptide groups.

Our findings provide evidence that hydration of helix ends is one of the major factors determining its stability and lead to the following specific conclusions.

1) Helix nucleation is much more difficult than it has been suspected. It is linked not only, as recognized long ago (9,10), with a large conformational entropy drop of two residues, not compensated by intrahelical hydrogen bonding, but also with a dramatic increase of peptide group solvation energy.

2) Because of poor hydration very short helical structures ($n < 6$) are extremely unstable. The helix-coil transition theories are, therefore, applicable only to longer helices with well separated ends.

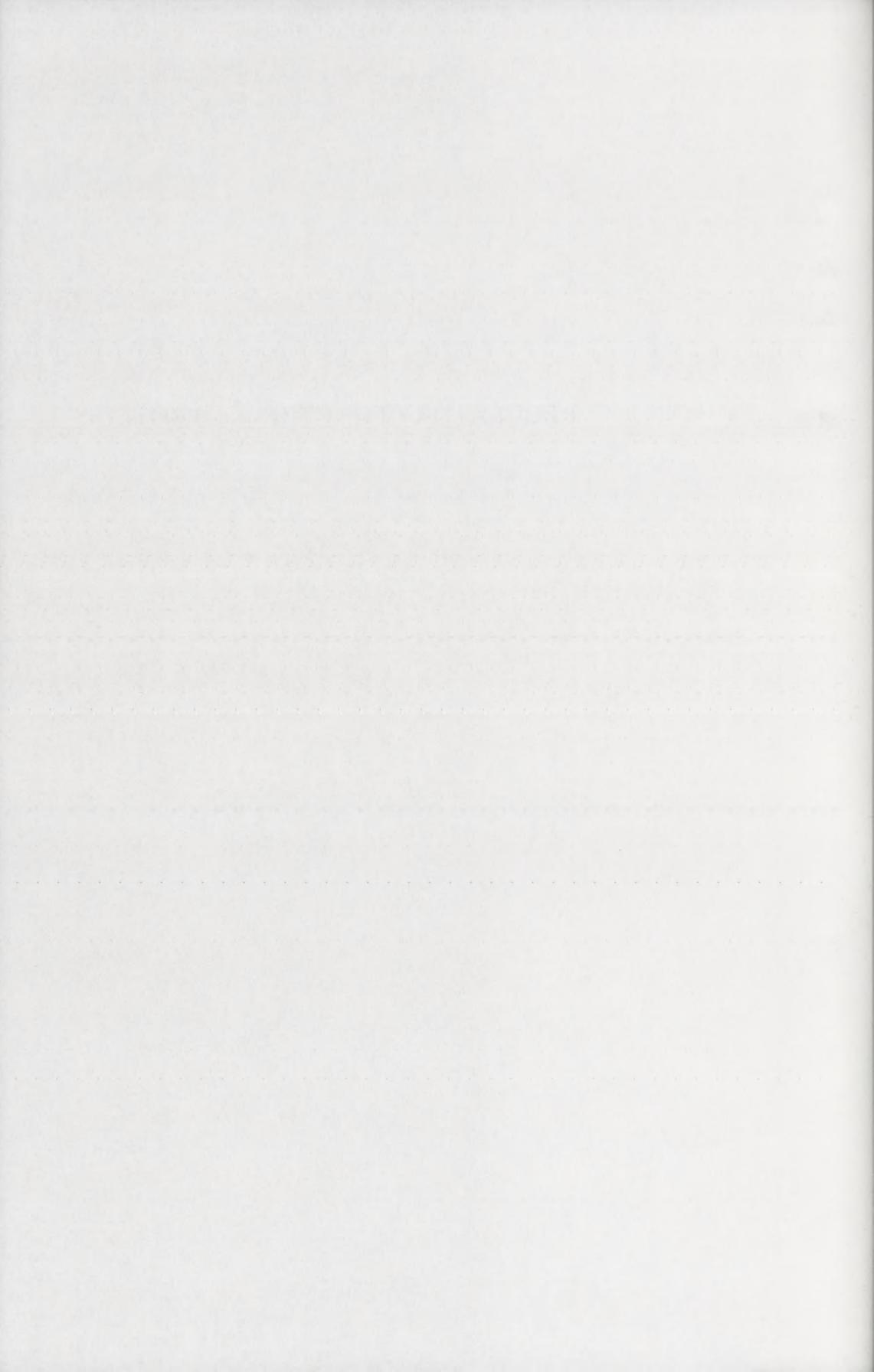
3) Hydration of the helix ends depends, of course, on their primary structure. Some sequences of amino acid residues, particularly those with large side chains, can be well accommodated into the interior of the helix when intolerable at its ends. Therefore, the helix propagation must be looked at, not as a basically continuous process with the rate determined by propagation parameters, but rather as series of jumps from one helical segment with well hydrated ends to another, across energy barriers imposed by end sequences that are poorly hydrated.

References

- 1) Dyson, H.J., Cross, K.J., Houghten, R.A., Wilson, I.A., Wright, P.E. and Lerner, R.A. (1985) *Nature* **318**, 480-483.
- 2) Bierzynski, A., Kim, P.S. and Baldwin, R.L. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2470-2474.
- 3) Shoemaker, K.R., Fairman, R., Kim, P.S., York, E.J., Stewart, J.M. and Baldwin, R.L. (1987) *Cold Spring Harb Symp Quant Biol* **52**, 391-398.
- 4) Kato, I. and Anfinsen, C.B. (1969) *J.Biol.Chem* **244**, 1004-1007.
- 5) Thompson, P.A., Eaton, W.A. and Hofrichter, J. (1997) *Biochemistry* **36**, 9200-9210.
- 6) Bierzynski, A. (1987) *Comments Mol. Cell. Biophys.* **4**, 189-214.
- 7) Munoz, V. and Serrano, L. (1995) *J. Mol. Biol.* **245**, 297-308.

- 8) Bierzynski, A. and Pawlowski, K. (1997) *Acta Biochim. Polon.* **44**, 423-432.
- 9) Zimm, B.H. and Bragg, J.K. (1959) *J. Chem. Phys.* **31**, 526-535.
- 10) Lifson, S. and Roig, A. (1961) *J. Chem. Phys.* **34**, 1963-1974.
- 11) Chakrabartty, A., Kortemme, T. and Baldwin, R.L. (1994) *Protein Sci.* **3**, 843-852.
- 12) Dadlez, M., Goral, J. and Bierzynski, A. (1991) *FEBS Lett* **282**, 143-146.
- 13) Wojcik, J., Goral, J., Pawlowski, K. and Bierzynski, A. (1997) *Biochemistry* **36**, 680-687.
- 14) Ejchart, A., Siedlecka, M., Sticht, H. and Bierzynski, A. (1997) *Bull. Acad. Polon. Sci., Sci. Chim.* In print.
- 15) Strynadka, N.C. and James, M.N. (1989) *Annu. Rev. Biochem.* **58**, 951-998.
- 16) Chakrabartty, A. and Baldwin, R.L. (1995) *Adv. Protein Chem.* **46**, 141-176.
- 17) Gans, P.J., Lyu, P.C., Manning, M.C., Woody, R.W. and Kallenbach, N.R. (1991) *Biopolymers* **31**, 1605-1614.
- 18) Doig, A.J. and Baldwin, R.L. (1995) *Protein Sci.* **4**, 1325-1336.

REGULATORY FUNCTION OF PROTEINS



STRUCTURE AND FUNCTION OF MOLECULAR CHAPERONE GROE: THE CHAPERONIN-FACILITATED PROTEIN FOLDING AND THE ROLE OF NUCLEOTIDE

Yasushi Kawata

Department of Biotechnology, Faculty of Engineering, Tottori University,
Koyama-Minami, Tottori 680, Japan
e-mail: kawata@bio.tottori-u.ac.jp

Abstract: Molecular chaperone, chaperonin GroE (GroEL and GroES) from *Escherichia coli* was found to facilitate folding reactions of various proteins *in vitro*. In order to function effectively, GroE requires ATP in the reaction and hydrolyzes it to ADP. We found that the chaperonin functions in the presence of not only ATP but also non-hydrolyzable nucleotides, such as AMP-PNP, ATP- γ -S, and ADP. This ability of chaperonin was proved by studying the functional characteristics of a GroEL T89W mutant, which can bind ATP but not hydrolyze it. The importance of nucleotide binding in chaperonin function is stressed by our results.

It has been revealed that molecular chaperones play important roles in numerous events which occur in the cell. One of the fundamental and critical issues is the folding of polypeptide that was newly synthesized or unfolded under stressful conditions *in vivo*. Recently, it has been shown that this folding process is mediated by proteins known as the chaperonins, which specifically facilitate the folding of various polypeptides *in vivo* and *in vitro*. The chaperonin GroE (GroEL and GroES) from *Escherichia coli* is one of the most extensively studied molecular chaperones (1, 2). The GroE protein not only facilitates protein folding but also prevents heat denaturation of proteins in the presence of ATP. This ATP-dependent chaperonin function is very important for living cells, especially under stressful conditions.

GroE is composed of two different subunits, GroEL and GroES. GroEL is a tetradecameric protein with a subunit molecular weight of 57,259 as determined by DNA sequence analysis. The subunits are arranged in a "double doughnut" configuration, with two heptameric rings sharing a common axis. From recent X-ray analyses (3-6), the subunit consists of three domains; equatorial, intermediate, and apical domains. There is a big hole (central cavity) surrounded by seven oligomers. Two heptameric rings interact each other back to back with the equatorial domains. GroEL has weak ATPase activity and ATP binding site is located in the equatorial domain. GroES is also heptameric protein of subunit molecular weight 10,368, and the subunit structure is consisted of mainly β -strands. In the presence of ATP or ADP, GroEL associates with GroES to form a huge complex of GroEL (14-mer)-GroES (7-mer).

The functional mechanism of the chaperonin-facilitated folding phenomenon involves interactions between the GroEL protein, the GroES protein, and the nucleotide ATP. This specific mechanism can be divided broadly into two parts as shown in Fig. 1; binding of protein folding intermediates prone to aggregation, and the controlled release of these intermediates to allow maximum refolding yields. In the former step, GroEL recognizes a folding intermediate and

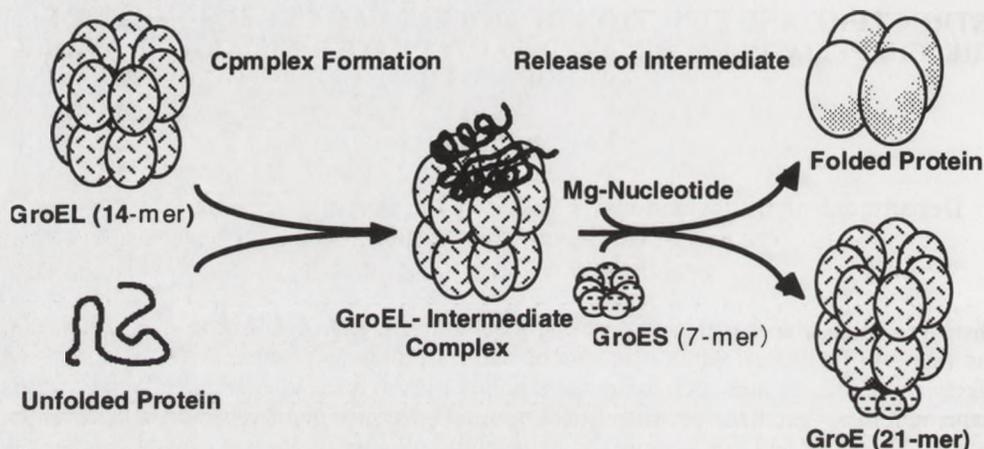


Fig. 1. A schematic model of GroE-facilitated protein folding.

binds it specifically by hydrophobic and electrostatic interactions (7). By forming this complex, the folding intermediate is stabilized and irreversible aggregation is suppressed. In the latter step, which usually occurs when ATP is added, an efficient release of the intermediate occurs and the native protein is formed. In this step, GroES protein plays an important role in assisting the function of GroEL. This intriguing mechanism is, however, very complicated and should be studied more extensively. In order to characterize the mechanism in more detail, we studied refolding reactions of various proteins with chaperonin GroE.

GroE-Facilitated Refolding of Various Proteins in the Presence of ATP and ADP

We have studied various protein folding in the presence of chaperonin GroE. The enzymes used were Taka-amylase A (TAA; monomer, Mw=54,000) from *Aspergillus oryzae*, malate dehydrogenase (MDH; dimer, subunit Mw=27,000) from *Thermus* species, enolase (dimer, subunit Mw=47,000) from *Saccharomyces cerevisiae*, lactate dehydrogenase (LDH; dimer, subunit Mw=32,000) from *Staphylococcus* species, glucose dehydrogenase (GLUCDH; tetramer, subunit Mw=26,300) from *Bacillus* species, and tryptophanase (TPase; tetramer, subunit Mw=52,000) from *Escherichia coli*. The quaternary structure of these six enzymes range from a monomeric state (Taka-amylase A) to a tetrameric state (GLUCDH, tryptophanase). Taka-amylase A is an extra-cellular enzyme and also distinguished from the other five enzymes by the presence of 4 disulfide bonds located in its tertiary structure. Guanidine hydrochloride was used to unfold the proteins completely and the refolding reaction was initiated by dilution into refolding buffers containing selected factors of the GroE complex.

Our experiments showed that GroE was capable of interacting with each of these all enzymes; in the absence of nucleotide, the refolding reactions of guanidine hydrochloride-unfolded proteins were arrested by GroEL. Addition of GroES and ATP to the reaction mixture initiated the dissociation of these proteins from GroEL and the enzyme activities were regained efficiently as shown in Fig.

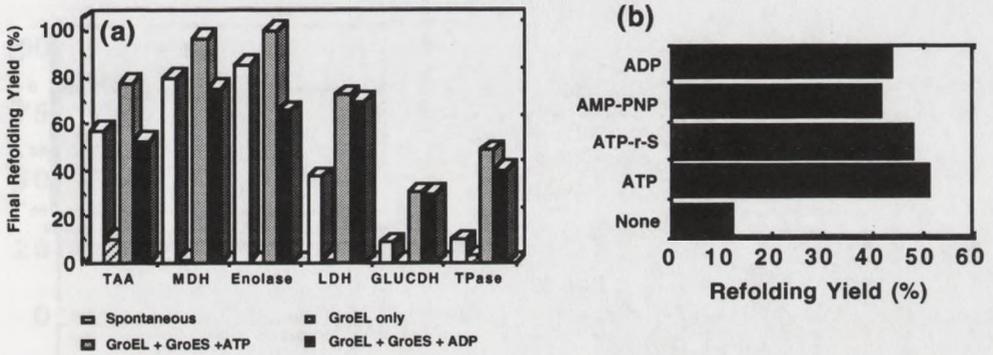


Fig. 2. GroE-facilitated protein folding characteristics of various enzymes in the presence of 2 mM ATP or ADP (a) and final refolding yield of tryptophanase in the presence of GroE and 2 mM various nucleotides (b). Five-fold molar excess GroE relative to enzyme's protomer was used and the activity was measured. Refolding temperature was 25°C.

2a. For more detailed, tryptophanase, which has been known to refold very poorly *in vitro*, showed a vast improvement in refolding yield (from 15% to 80%) in the presence of excess amounts of GroE and ATP (8). On the other hand, the dimeric enzyme enolase, which refolds almost quantitatively from the unfolded state, showed a negligible increase in refolding yield as a result of addition of GroE and ATP. However, the refolding of enolase was completely arrested in the presence of GroE when ATP was omitted from the refolding mixture, indicating that a specific interaction between GroE and enolase folding intermediates was taking place (9). In the case of Taka-amylase A, the addition of GroE to the refolding mixture resulted in a large increase in refolding yield, regardless of the presence or absence of disulfide bonds. The results of our studies support the proposition that GroE is capable of facilitating the folding of many proteins regardless of various differences in structural characteristics, cellular location, or origin. A systematic comparison of the refolding characteristics of these enzymes should be useful in elucidating a common mechanism for GroE-facilitated protein folding. Also, chaperonins may be utilized in a variety of processes related to protein engineering, in particular, in attaining a greater yield of overproduced protein from a given cellular system (10).

In the study of GroE-facilitated refolding of tryptophanase, we observed an interesting result that GroE is capable to refold the enzyme in the presence of not only ATP but also non-hydrolyzable ATP-analogs (ATP- γ -S and AMP-PNP) and ADP as shown in Fig. 2b. Further experiments showed that ADP was not hydrolyzed by GroEL during the refolding reaction, which indicated that the energy released by ATP hydrolysis is not required for an efficient release of tryptophanase from GroE.

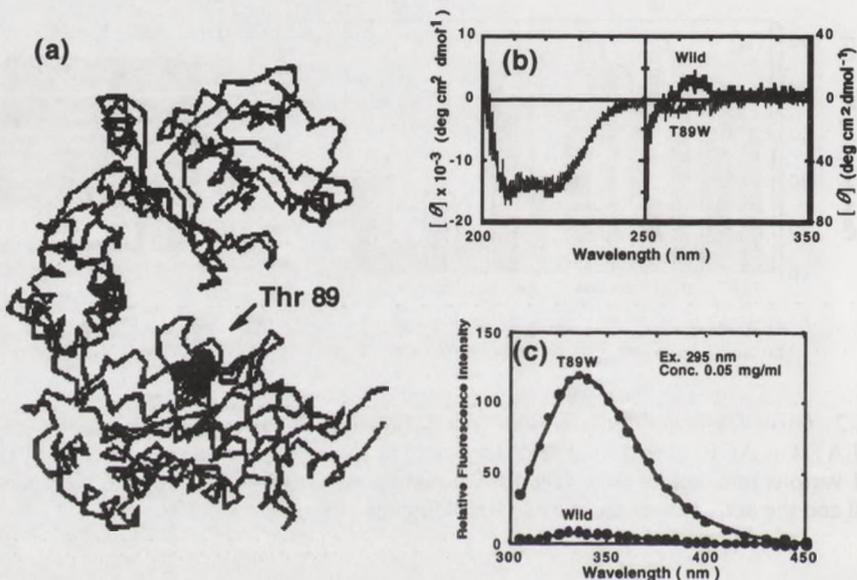


Fig. 3. Mutation site of Thr 89 located in equatorial domain of GroEL subunit (a), circular dichroism spectra (b) and tryptophyl fluorescence spectra (c) of wild-type and GroEL T89W mutant.

In order to examine whether GroE-facilitated protein folding in the presence of ADP occurs for various proteins in common, we have also performed refolding reactions of the above six enzymes in the presence of GroE and ADP. Very interestingly, as shown in Fig. 2a, it was found that refolding reactions of all enzymes were mediated in the chaperonin-dependent manner, as well as in the presence of ATP. We also found that the chaperonin GroE and ADP not only facilitated the refolding of all these enzymes, but also prevented the irreversible heat inactivation of LDH and GLUCDH (11). These findings suggest that nucleotide binding is an important event in the mechanism of GroE-facilitated protein folding.

The Role of Nucleotide

Studies regarding the effects of the chaperonins on the folding reactions of various proteins have yielded many important clues to understanding the molecular mechanism of chaperonin-facilitated protein folding. In many protein refolding reactions, addition of ADP was sufficient to release the bound intermediates, as well as ATP. This fact strongly suggested that the binding of nucleotide to GroEL is important and that the energy derived from ATP hydrolysis may not be directly involved in the mechanism. In order to understand the role of ATP hydrolysis during chaperonin function, we have produced some mutant GroEL proteins using site-directed mutagenesis. One of them, GroEL

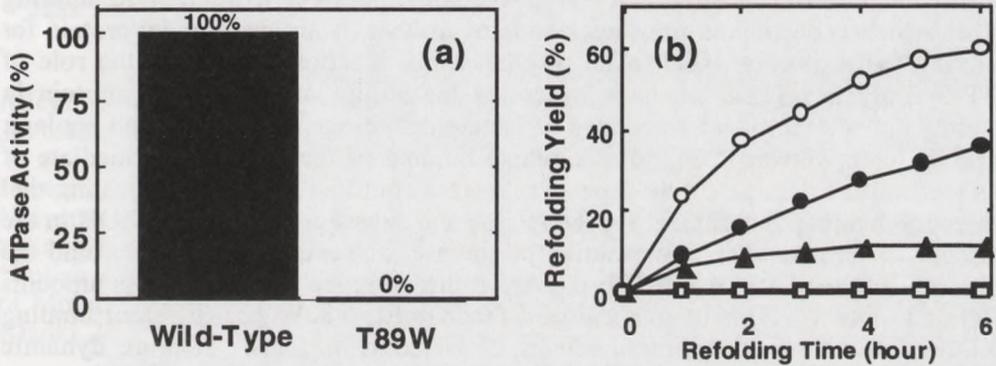


Fig. 4. ATPase activity of wild-type and T89W mutant GroEL at 37°C (a) and the chaperonin-facilitated refolding of LDH at 25°C (b). In the presence of wild-type GroEL only (open square); T89W GroEL only (closed square); wild-type GroEL, GroES and 2 mM ATP (open circle); T89W GroEL, GroES, and 2 mM ATP (closed circle). Five-fold molar excess chaperonins relative to LDH protomer were used. Closed triangle indicates spontaneous refolding of LDH.

T89W, showed the most interesting effects of residue substitution, and has been utilized in probing the specific reactions which take place during facilitation of folding by chaperonin.

GroEL T89W, in which the residue Thr 89 has been replaced to a tryptophan, is a mutant with a sole tryptophan residue introduced into the (tryptophan-lacking) wild type sequence. As shown in Fig. 3a, the mutated residue of Thr 89 is located in the vicinity of the nucleotide binding site at the equatorial domain of GroEL, and the γ OH group interacts with γ -phosphate group of ATP. We have studied this mutant protein in detail from the structural and functional characteristics point of view.

GroEL T89W mutant protein was highly expressed in *E. coli* in soluble form, indicating that this mutant was stable *in vivo*. Circular dichroism (CD) spectra (Fig. 3b) and tryptophyl fluorescence spectra (Fig. 3c) of the purified protein are shown in Fig. 3. Far-UV CD spectrum of T89W mutant was very similar to that of wild-type although near-UV CD spectrum was somewhat different. The difference of the aromatic region in CD spectrum was thought to be due to the mutation to tryptophan. The same reason was in the spectrum in tryptophyl fluorescence as shown in Fig. 3c. Since wild-type GroEL does not have any tryptophan residues, this tryptophyl fluorescence of T89W mutant was very useful in probing nucleotide binding as described below. Altogether with gel-filtration experiments it was concluded that GroEL T89W was similar to the wild type protein in aspects of quaternary and secondary structural characteristics.

Next, we have studied the functional properties of this chaperonin. Very interestingly, the mutant chaperonin lost completely its ability to hydrolyze ATP as shown in Fig. 4a. However, binding experiments performed by monitoring the changes in tryptophan fluorescence of the mutant protein indicated that the mutant was capable of binding nucleotides such as ATP and ADP with a reduced

affinity. Therefore, GroEL T89W, possessing in effect a nucleotide binding ability which is decoupled from nucleotide hydrolysis, is an ideal model protein for studying the molecular mechanism of chaperonin function, especially the role of ATP-hydrolysis. Then, we have examined the ability of this mutant protein in binding and release of intermediates of lactate dehydrogenase (LDH) and enolase. Expectedly, as shown in Fig. 4b, complete binding of the folding intermediate of LDH as well as the nucleotide-dependent release could be observed, indicating that nucleotide binding is sufficient for the release and subsequent folding of LDH in the presence of GroEL. For the refolding of enolase, however, the ability to bind the refolding intermediates was reduced even in the presence of large excess amounts of GroEL T89W. Thus it may be said that GroEL T89W has different binding abilities for the refolding intermediates of different proteins. A more dynamic mechanism may be necessary to explain the binding mechanism completely.

Our experimental system using GroEL T89W mutant allows a detailed analysis of the specific role of nucleotide in chaperonin function. The fact that the trapped refolding intermediates of LDH and enolase were released from GroEL T89W upon ATP addition and refolded effectively shows clearly the importance of nucleotide binding but not hydrolysis. A significance of ATP-hydrolysis consists probably in maintaining the stable GroEL-GroES complex kinetically and its regulation (12). Thus, the ADP-dependent chaperonin function ("ADP-cycle") is also suggested by our results and may play a critical role *in vivo*, especially under stressful conditions. Further studies using this system should clarify many puzzling details of this intriguing phenomenon.

References

1. Hartl, F.U. (1996) *Nature* **381**, 571-580.
2. Fenton, W.A. and Horwich, A.L. (1997) *Protein Science* **6**, 743-760.
3. Braig, K., Otwinowski, Z., Hegde, E., Boisvert, D.C., Joachimiak, A., Horwich, A.L. and Sigler, P.B. (1994) *Nature* **371**, 578-586.
4. Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L. and Deisenhofer, J. (1996) *Nature* **379**, 37-45.
5. Boisvert, D.C., Wang, J., Otwinowski, Z., Horwich, A.L. and Sigler, P.B. (1996) *Nature Struct. Biol.* **3**, 170-177.
6. Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) *Nature* **388**, 741-750.
7. Hoshino, M., Kawata, Y. and Goto, Y. (1996) *J. Mol. Biol.* **262**, 575-587.
8. Mizobata, T., Akiyama, Y., Ito, K., Yumoto, N. and Kawata, Y. (1992) *J. Biol. Chem.* **267**, 17773-17779.
9. Kubo T., Mizobata, T. and Kawata, Y. (1993) *J. Biol. Chem.* **268**, 19346-19351.
10. Ashiuchi, M., Yoshimura, T., Kitamura, T., Kawata, Y., Nagai, J., Gorlatov, S., Esaki, N. and Soda, K. (1995) *J. Biochem.* **117**, 495-498.
11. Kawata, Y., Nosaka, K., Hongo, K., Mizobata, T. and Nagai, J. (1994) *FEBS Letters* **345**, 229-232.
12. Kawata, Y., Hongo, K., Nosaka, K., Furutsu, Y., Mizobata, T., and Nagai, J. (1995) *FEBS Letters* **369**, 283-286.

HSP70 MOLECULAR CHAPERONE MACHINERY: STRUCTURE AND FUNCTION

Alicja Wawrzynów, Bogdan Banecki, Krzysztof Liberek, Adam Błaszczak*, Jarosław Marszałek, Igor Konieczny oraz Maciej Żylicz

Department of Molecular and Cellular Biology, Faculty of Biotechnology, University of Gdańsk, 80-822 Gdańsk, Kładki 24, Poland. E-mail: zylicz@biotech.univ.gda.pl

*Polish Academy of Science, Institute of Biochemistry and Biophysics, Laboratory of Molecular Biology affiliated to the University of Gdańsk, 80-822 Gdańsk, Kładki 24, Poland.

Determination of molecular mechanism of bacterial DnaK/DnaJ/GrpE chaperone machinery provides insights into the principles governing Hsp70 interaction with polypeptide chains. The proteins belonging to Hsp70 family (eukaryotic DnaK homologue) together with Hsp40 and Hsp24 (DnaJ and GrpE eukaryotic homologues respectively) are involved in protein folding, protein transport and regulation of gene expression. During stress conditions, such as heat shock, the Hsp70/Hsp40 chaperone machine protects and reactivates other proteins. It has been shown recently that Hsp70-substrate complex isolated from cancer cells may act as a vaccine suppressing a cancer formation. These findings combined with involvement of molecular chaperones in prion-dependent diseases make the problem of Hsp70 recycling from complex with protein substrates medically relevant. In this paper we discuss differences in few, existing in the literature and proposed by us, models of Hsp70 action.

The *Escherichia coli* DnaK heat shock protein was identified as a peptide-stimulated ATPase, directly involved in the initiation of λ DNA replication (Żylicz *et al.*, 1983; Żylicz and Georgopoulos, 1984). This 70 kDa protein in the presence of two other bacterial heat shock proteins, namely DnaJ (Żylicz *et al.*, 1985) and GrpE (Żylicz *et al.*, 1987) triggers the initiation of λ DNA replication by releasing inhibitor, the λ P replication protein from the λ preprimosomal complex (Liberek *et al.*, 1988; Żylicz *et al.*, 1989; Osipiuk *et al.*, 1993). The homologues of bacterial DnaK (Hsp70), DnaJ (Hsp40) and GrpE (Hsp24) heat shock proteins were found in all tested prokaryotic and eukaryotic organisms (Deloche *et al.*, 1997; for review see Georgopoulos and Welch, 1994). It was shown that Hsp70/Hsp40/Hsp24 machine performs many chaperone function. These proteins are involved in protein transport (Hendrick *et al.*, 1993), proteolysis (Sherman and Goldberg, 1992; Gottesman, 1996), protein folding (Frydman and Hartl, 1996; Hartl, 1996), uncoating of clathrin-coated vesicles (Barouch *et al.*, 1997), regeneration of steroid hormone receptors (Hutchison *et al.*, 1994). During stress conditions the Hsp70/Hsp40/Hsp24 machine is able to protect other protein from inactivation, dissolve already formed aggregates (Skowyra *et al.*, 1991; Schroder *et al.*, 1993; Ziemienowicz *et al.*, 1993; 1995) and regulate heat shock response (Straus *et al.*, 1990; Liberek *et al.*, 1992; Liberek and Georgopoulos, 1993; Błaszczak *et al.*, 1996).

The chaperone functions of Hsp70/Hsp40/Hsp24 machine are performed by transient binding of Hsp70 and Hsp40 to denatured or native proteins. Where binding of bacterial DnaK to protein substrates is not very specific, the DnaJ protein alone seems to recognize more specific peptides sequence and/or structure (Wawrzynów and Żylicz, 1995). It has been shown that DnaJ alone can perform, at least *in vitro*, some chaperone activities (Hendrick *et al.*, 1993; Cyr *et al.*, 1994).

Relevance of these findings needs to be confirmed *in vivo*. In bacteria and mitochondria a third protein, GrpE (Hsp24) is also required. Recently, it has been suggested that in the presence of ATP, the DnaJ protein changes the affinity of DnaK to different substrates (Wawrzynów *et al.*, 1995). These findings make studies on affinity of DnaK alone to different peptides, biologically irrelevant (Gragerov and Gottesman, 1994; Rudiger *et al.*, 1997).

The DnaK protein binds ATP very tightly (k_D 3-7 nM) and in the absence of DnaJ and GrpE hydrolyzes ATP in a rather slow reaction. One molecule of ATP is hydrolyzed by DnaK every 10-18 minutes. The presence of DnaJ and GrpE accelerates the steady state rate of ATP hydrolysis up 50 to 180 fold (Liberek *et al.*, 1991; Jordan and McMacken, 1995; McCarty *et al.*, 1994). The DnaJ protein alone as well as protein substrates can also stimulate the DnaK ATPase albeit to much lower extent (2-5 fold). The GrpE was shown to be a nucleotide exchange factor which accelerates the release of ADP from the DnaK-ADP complex (Liberek *et al.*, 1991).

The DnaK/DnaJ/GrpE chaperone machine action requires ATP hydrolysis. It was postulated that ATP hydrolysis is required for recycling of DnaK from the DnaK-substrate complex (Żylicz *et al.*, 1989; Liberek *et al.*, 1991b). Only recently the question of involvement of ATP in this cycle was addressed using the physiological situation when all three elements of chaperone machine, namely DnaK, DnaJ and GrpE were present in the reaction mixture. Previously, the DnaK protein was used alone and the results were extrapolated to the situation where all three components were present (Palleros *et al.*, 1993; Schmid *et al.*, 1994). Special effort was undertaken to strip the DnaK from nucleotide (Theysen *et al.*, 1996), the procedure which simultaneously induces an artificial DnaK oligomerization (Banecki, unpublished results).

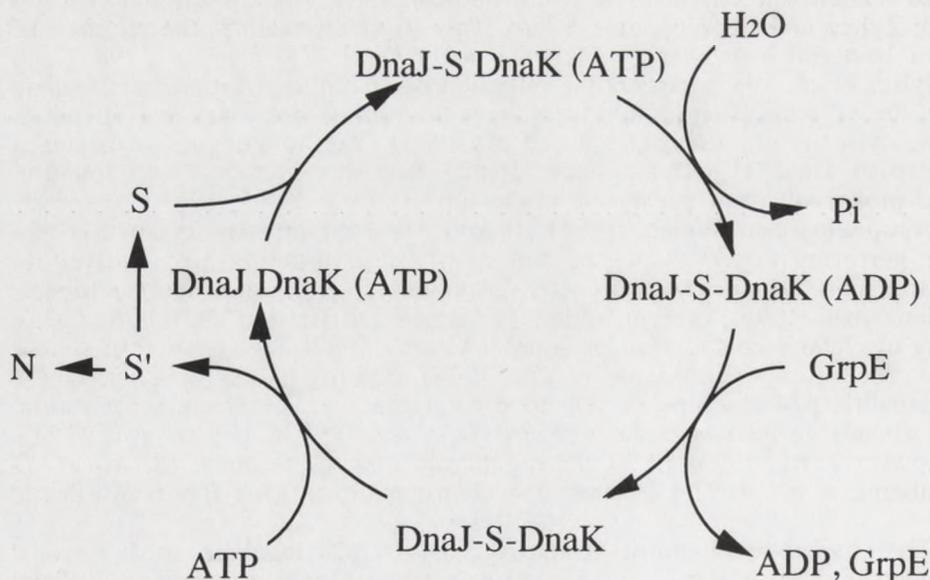


Fig. 1. The model of Hsp70 cycle proposed by Szabo *et al.*, 1994.
S - protein substrate (luciferase), N - native substrate

Szabo and co-workers have analyzed a model reaction in which DnaK, DnaJ and GrpE mediate the folding of denatured firefly luciferase, the substrate which also forms a stable complex with the DnaJ protein alone (Szabo *et al.*, 1994). The binding and release of substrate from the complex with DnaK was modulated by ATP hydrolysis. The following model of the DnaK/DnaJ/GrpE cycle was proposed: (i) DnaJ protein binds unfolded luciferase and delivers it to the DnaK(ATP) (ii) upon interaction of luciferase-DnaJ with DnaK, ATP is hydrolyzed and a stable DnaJ-luciferase-DnaK(ADP) complex is formed. (iii) GrpE releases ADP from DnaK and (iv) the binding of ATP to DnaK releases the protein substrate thus completing the reaction cycle. Several rounds of ATP-dependent interactions with DnaK and DnaJ are required for full efficient folding (Fig. 1).

Interestingly, the addition of ATP analogue (AMP-PNP) and GrpE to the ATP-dependent activated DnaJ-luciferase-DnaK(ADP) complex resulted in an intermediate level of luciferase reactivation (about 35%). This lead to the interpretation that AMP-PNP and GrpE allowed the release of luciferase from DnaK and DnaJ but not the reformation of the ternary complex. Based on these experiments authors conclude that ATP hydrolysis by DnaK is required only for rapid formation of the polypeptide-DnaK complex and not for its dissociation. The Szabo's model presented in Figure 1 assumes that DnaJ protein does not dissociate luciferase until the luciferase is folded correctly. Based on different protein substrate - σ^{32} , Gamer *et al.*, (1996) proposed a similar DnaK/DnaJ/GrpE cycle. The only difference was the suggestion that DnaJ could dissociate from the ternary or quaternary complex independently of DnaK, thereby allowing, in principle, to act in this cycle in substoichiometric amount. It had been shown before that substoichiometric amount of DnaJ was efficient to activate (in the presence of ATP) the binding of DnaK to σ^{32} (Liberek *et al.*, 1995).

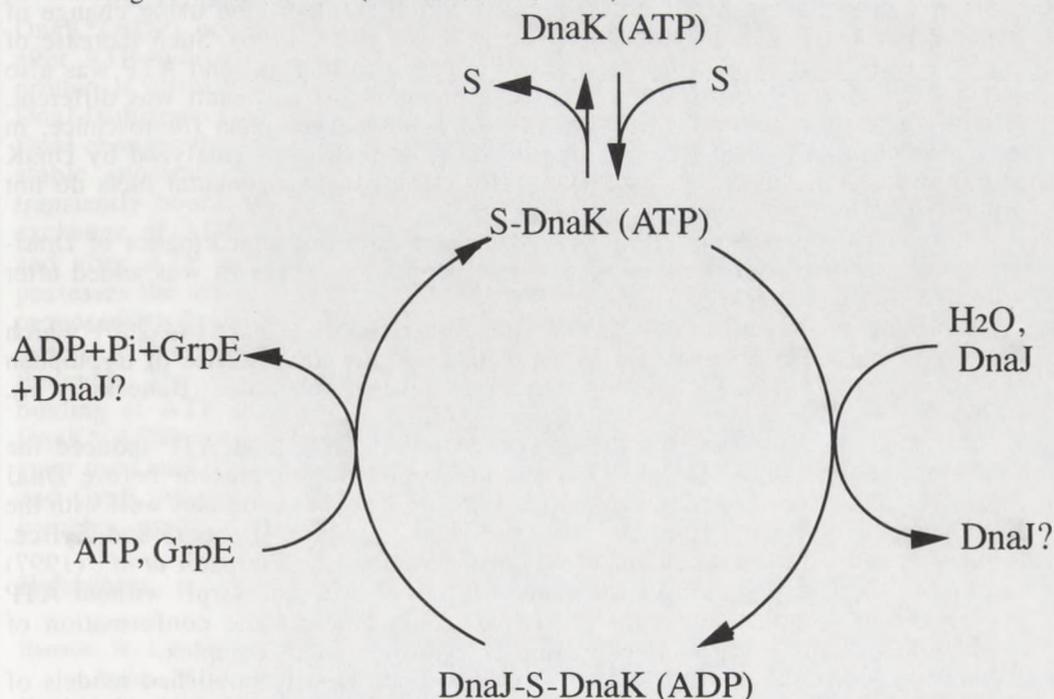


Fig. 2. The model of Hsp70 cycle proposed by Pierpaoli *et al.*, 1997.

Recently, another model of the DnaK/DnaJ/GrpE cycle has been proposed (Fig. 2; Pierpaoli *et al.*, 1997). Based on the real-time kinetic measurements of the intrinsic fluorescence of tryptophan in DnaK and fluorescence of dansyl-labeled peptide ligands the following steps of DnaK/DnaJ/GrpE cycle were proposed: (i) binding of target protein to fast-binding-and-releasing DnaK(ATP) stage which has low affinity to substrates; (ii) DnaJ-triggered conversion of substrate-DnaK(ATP) form to a high-affinity substrate-DnaK(ADP+Pi) complex; (iii) GrpE-facilitated ADP/ATP exchange which leads to the conversion of DnaK to the low affinity form and dissociation of the substrate-DnaK complex. The DnaK/DnaJ/GrpE cycle proposed by Pierpaoli and co-workers does not substantially differ from the initial suggested by Wall *et al.*, (1995). The experiments described by Pierpaoli and co-workers were performed using the peptide substrates which possessed the limited affinity to the DnaJ protein. This explains why the reaction of initial binding of DnaJ to the protein substrates, proposed by other authors, was omitted.

Some experimental data generated during last two years do not support the overall view of the DnaK/DnaJ/GrpE cycle model, mentioned above:

1. The DnaK and DnaJ proteins interact with each other only in the presence of ATP. The poorly hydrolysable ATP analogues inhibit such reaction. This suggests that DnaJ interacts stably not with the DnaK(ATP) but with DnaK(ADP) form (Wawrzynów and Żylicz, 1995).

2. The DnaJ-dependent activation of DnaK for binding to both native and denatured protein substrates requires ATP hydrolysis. The simultaneous presence of protein substrate and DnaJ is not required for this activation reaction, the substrate can be added after ATP hydrolysis (Wawrzynów *et al.*, 1995).

3. The activation reaction of DnaK for binding to protein substrate could be divided into two processes: i) hydrolysis of ATP ii) DnaJ-dependent conformational change of DnaK. This last reaction was detected using change of fluorescence of DnaK's tryptophan (Banecki and Żylicz, 1996). Such increase of the fluorescence of DnaK's tryptophan in the presence of DnaJ and ATP was also detected by Pierpaoli *et al.*, (1997) but interpretation of that result was different. According to their interpretation, increase of DnaK's tryptophan fluorescence, in the presence of ATP and DnaJ, is due to the ATP hydrolysis catalyzed by DnaK and stimulated in the presence of DnaJ. However, two experimental facts do not support such interpretation:

(i) 0.5 mM Pi which inhibits the ATP hydrolysis does not alter kinetics of DnaJ-dependent conformational changes, under the condition where Pi was added after ATP hydrolysis (Banecki and Żylicz, 1996);

(ii) DnaJ mutants with the deletion of zinc finger motifs (DnaJ Δ 144-200) which stimulate DnaK's ATPase activity failed to increase the fluorescence of tryptophan and activate the DnaK for binding to various protein substrates (Banecki *et al.*, 1996).

It has been shown that the presence of both GrpE and ATP induced the conformational change of DnaK back to its conformation present before DnaJ dependent activation (Banecki and Żylicz, 1996). This event correlates well with the dissociation of substrate from the substrate-DnaK complex (Banecki and Żylicz, 1996). Recently, similar phenomena was also described by Pierpaoli *et al.*, (1997) and again, the interpretation of the same effect was different: GrpE without ATP hydrolysis, by simple exchange of ADP to ATP, changes the conformation of DnaK, allowing the substrate release from the substrate-DnaK complex.

Taking into consideration all the above critics of previously published models of DnaK/DnaJ/GrpE cycle we suggest a new model which embraces and suits all the experimental data published before (Fig.3).

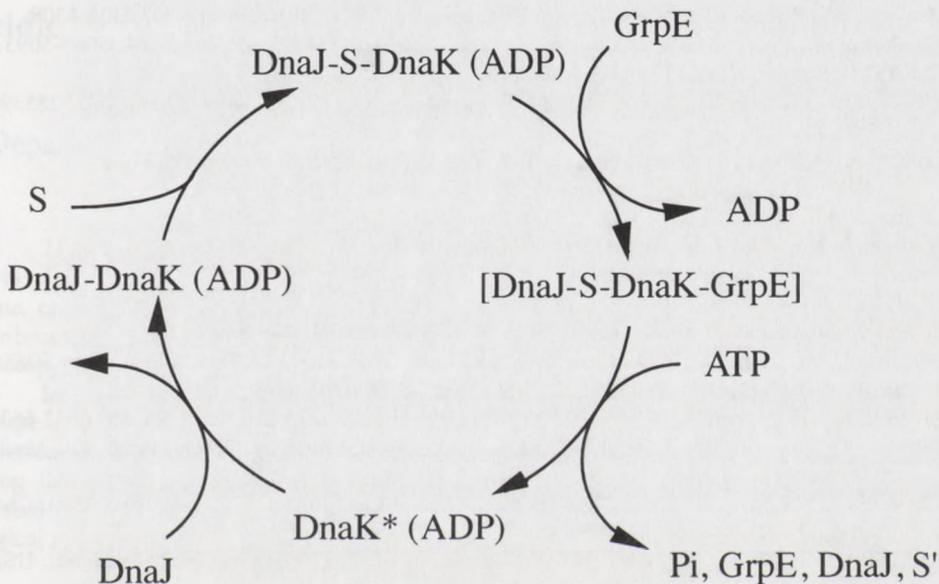


Fig.3. The model of Hsp70 cycle proposed in this paper (see also Banecki and Żylicz, 1996).

We propose that the DnaJ protein interacts with DnaK only when DnaK is in DnaK*(ADP) conformation. Such conformation can be achieved by DnaK only after ATP hydrolysis (not by simple incubation of DnaK with ADP). If the DnaJ protein is already in the complex with substrate the ternary complex is formed: DnaJ-(substrate-DnaK(ADP)). If the affinity of DnaJ to the substrate is low then DnaJ changes the DnaK*(ADP) conformation in such a way that the formation of a stable substrate-DnaK(ADP) complex is possible. The DnaJ protein is only transiently bound to such complex. In the presence of ATP, GrpE accelerates the exchange of ADP to ATP in DnaK (quaternary intermediate complex is formed) and after ATP hydrolysis the complex dissociates. If the protein substrate still possesses the affinity to DnaJ it will immediately move to another cycle (as it was proposed by Szabo *et al.*, 1994). If the DnaJ protein loses affinity to the substrate (after e.g. folding), it needs to activate the DnaK* form before this substrate enters the cycle. The model presented in Fig. 3 does not exclude the possibility that after binding of ATP analogues, the complex will also fall apart but in such case the DnaK*(ADP) conformation will not be reached and the substrate will never again enter the DnaK/DnaJ/GrpE cycle. This is probably the explanation why (AMP-PNP) and GrpE, when added to ATP-dependent activated DnaJ-luciferase-DnaK(ADP) complex, reactivated the luciferase only to a limited extent (Szabo *et al.*, 1994).

References

- Banecki, B., Liberek, K., Wall, D., Wawrzynów, A., Georgopoulos, C., Bertoli, E., Tanfani, F., and Żylicz, M. (1996) *J. Biol. Chem.* 271, 14840-14848.
 Banecki, B. and Żylicz, M. (1996) *J. Biol. Chem.* 271, 6137-6144.

- Barouch, W., Prasad, K., Greene, L., and Eisenberg, E. (1997) *Biochemistry* 36, 4303-4308.
- Błaszczak, A., Żylicz, M., Georgopoulos, C., and Liberek, K. (1995) *EMBO J.* 14, 5085-5093.
- Cyr, D.M., Langer, T., and Douglas, M.G. (1994) *TIBS* 19, 176-181.
- Deloche, O., Liberek, K., Żylicz, M., and Georgopoulos, C. (1997) *J. Biol. Chem.* 272, 28539-28544.
- Gamer, J., Multhaup, G., Tomoyasu, T., McCarty, J.S., Rüdiger, S., Schönfeld, H.J., Schirra, C., Bujard, H., and Bukau, B. (1996) *EMBO J.* 15, 607-617.
- Georgopoulos, C. and Welch, W.J. (1993) *Annu. Rev. Cell. Biol.* 9, 601-635.
- Gottesman, S. (1996) *Ann. Rev. Genet.* 30, 465-506.
- Gragerov, A. and Gottesman, M.E. (1994) *J. Mol. Biol.* 241, 133-135.
- Hartl, F.U. (1996) *Nature* 381, 571-580.
- Hendrick, J.P. and Hartl, F.-U. (1993) *Annu. Rev. Biochem.* 62, 349-384.
- Hutchison, K.A., Dittmar, K.D., Czar, M.J., and Pratt, W.B. (1994) *J. Biol. Chem.* 269, 5043-5049.
- Jordan, R. and McMacken, R. (1995) *J. Biol. Chem.* 270, 4563-4569.
- Liberek, K., Georgopoulos, C., and Żylicz, M. (1988) *Proc. Nat. Acad. Sci. USA* 85, 6632-6636.
- Liberek, K., Marszałek, J., Ang, D., Georgopoulos, C., and Żylicz, M. (1991a) *Proc. Nat. Acad. Sci. USA* 88, 2874-2878.
- Liberek, K., Skowrya, D., Żylicz, M., Johnson, C., and Georgopoulos, C. (1991b) *J. Biol. Chem.* 266, 14491-14496.
- Liberek, K., Galitski, T.P., Żylicz, M., and Georgopoulos, C. (1992) *Proc. Natl. Acad. Sci. USA* 89, 3516-3520.
- Liberek, K. and Georgopoulos, C. (1993) *Proc. Natl. Acad. Sci. USA* 90, 11019-11023.
- Liberek, K., Wall, D., and Georgopoulos, C. (1995) *Proc. Nat. Acad. Sci. USA* 92, 6224-6228.
- McCarty, J.S., Buchberger, A., Reinstein, J., and Bukau, B. (1995) *J. Mol. Biol.* 249, 126-137.
- Osiptuk, J., Georgopoulos, C., and Żylicz, M. (1993) *J. Biol. Chem.* 268, 4821-4827.
- Palleros, D.R., Reid, K.L., Shi, L., Welch, W.J., and Fink, A.L. (1993) *Nature* 365, 664-666.
- Pierpaoli, E.V., Sandmeier, E., Baici, A., Schonfeld, H.J., Gisler, S., and Christen, P. (1997) *J. Mol. Biol.* 269, 757-768.
- Rüdiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997) *EMBO J.* 16, 1501-1507.
- Schmid, D., Baici, A., Gehring, H., and Christen, P. (1994) *Science*, 263, 971-973.
- Schröder, H., Langer, T., Hartl, F.-U., and Bukau, B. (1993) *EMBO J.* 12, 4137-4144.
- Sherman, M. and Goldberg, A.L. (1992) *EMBO J.* 11, 71-77.
- Skowrya, D., Georgopoulos, C., and Żylicz, M. (1990) *Cell* 62, 939-944.
- Straus, D., Walter, W., and Gross, C.A. (1990) *Genes and Development* 4, 2202-2209.
- Szabo, A., Langer, T., Schröder, H., Flanagan, J., Bukau, B., and Hartl, F.U. (1994) *Proc. Natl. Acad. Sci. USA* 91, 10345-10349.
- Theysen, H., Schuster, H.P., Bukau, B., and Reinstein, J. (1996) *J. Mol. Biol.* 263, 657-670.
- Wall, D., Żylicz, M., and Georgopoulos, C. (1995) *J. Biol. Chem.* 270, 2139-2144.
- Wawrzynów, A., Banecki, B., Wall, D., Liberek, K., Georgopoulos, C., and Żylicz, M. (1995) *J. Biol. Chem.* 270, 19307-19311.
- Wawrzynów, A. and Żylicz, M. (1995) *J. Biol. Chem.* 270, 19300-19306.
- Ziemenowicz, A., Skowrya, D., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Żylicz, M. (1993) *J. Biol. Chem.* 268, 25425-25431.
- Ziemenowicz, A., Żylicz, M., Floth, C., and Hubscher, U. (1995) *J. Biol. Chem.* 270, 15477-15479.
- Żylicz, M., LeBowitz, J.H., McMacken, R., Georgopoulos, C. (1983) *Proc. Natl. Acad. Sci. USA* 80, 6431-6435.
- Żylicz, M., Georgopoulos, C. (1984) *J. Biol. Chem.* 259, 8820-8825
- Żylicz, M., Yamamoto, T., McKittrick, N., Sell, S., and Georgopoulos, C. (1985) *J. Biol. Chem.* 260, 7591-7598.
- Żylicz, M., Ang, D., and Georgopoulos, C. (1987) *J. Biol. Chem.* 262, 17437-17442.
- Żylicz, M., Ang, D., Liberek, K., and Georgopoulos, C. (1989) *EMBO J.* 8, 1601-1608.
- Żylicz, M. (1993) *Phil. Trans. R. Soc. Lond.* 339, 271-278.

Acidic Ribosomal Proteins and their phosphorylation in yeast cells

Pilecki Marek, Dukowski Piotr and Grankowski Nikodem

Department of Molecular Biology, Maria Curie-Skłodowska University,
Akademicka 19, 20-033 Lublin, Poland

The first part of this report concerns the unknown so far localization of one of the protein kinases phosphorylating acidic ribosomal proteins in yeast cells. The best substrates for this enzyme, called PK60S are the two endogenous acidic ribosomal proteins YP1 β and YP2 α from 60S ribosomal subunit. Using the method of immunogold labeling of enzyme and electron microscopy, we showed the localization of PK60S around the mitochondria.

In the second part we present our observation of the ribosome phosphorylation-dephosphorylation processes and its effect on the ribosome structure. The complete dephosphorylation of ribosome, in contrast to the native and phosphorylated ribosomes, leads to creating the different forms of 60S subunit, containing surprisingly the whole set of acidic ribosomal dephosphorylated proteins. The presented results are part of our studies on the biological function of ribosomes phosphorylation.

INTRODUCTION

It is known that large ribosomal subunit of all organisms contains a set of strongly acidic proteins (A-proteins) having molecular masses of about 12-13 kDa. This kind of proteins is present in a few copies per ribosome and play an important role in translation [1]. Bacterial (*E.coli*) ribosomes have two A-proteins called L7 and L12. Both proteins are encoded by the same gene (rplL) and therefore the primary structure of L7/L12 polypeptides are identical. The difference between them concerns the N-terminus only, which is acetylated in L7 [2]. The number of acidic proteins in eukaryotic ribosomes varies from two in mammals to up to eight in protozoa [3]. In contrast with bacterial A-proteins, the eukaryotic acidic ribosomal proteins are phosphorylated and for that reason they are called P-proteins. These proteins can be classified in two subgroups according to their sequence homology to mammals proteins P1 and P2 [4].

The 60S ribosomal subunit from *Saccharomyces cerevisiae* cells has four P-proteins. They are encoded by independent genes [5] and are called YP1 α , Yp1 β (P1 subgroup), YP2 α and YP2 β (P2 subgroup) [4]. The P-proteins located on the characteristic lateral protuberance (*stalk*) of large ribosomal subunit can be extracted from ribosomes by washing them with NH₄Cl/ethanol solution, giving split proteins fraction and core particle [6]. P1/P2 proteins interact with non-acidic core protein - PO which binds directly to the r-RNA in the area of ribosomal GTP-ase center [7].

The P-proteins from yeast cells are phosphorylated as well *in vivo* as *in vitro* by multifunctional protein kinase CK-2 [8,9], high specific protein kinase - PK-60S [10] and ribosomal protein kinase - RAP [11]. The three enzymes phosphorylate exclusively the last serine residues located in a highly conserved carboxyl end of the polypeptide chains [12]. The role of this modification is not well recognized yet. It is

known that phosphorylated forms of P-proteins are found on the ribosomal particle. Non-phosphorylated forms are free in the cytoplasm [13-15]. Moreover, phosphorylation of P-proteins are required for the translational activity of ribosomes. Dephosphorylated P-proteins are unable to reconstitute elongation factor dependent activity of ribosomes from core particles and acidic proteins extracted from the particles with the solution of NH_4Cl /ethanol [16].

In this report we present some new data on the localization of the specific protein kinase - PK60S in yeast cells and the effect of acidic proteins phosphorylation on ribosomal association/dissociation processes.

RESULTS AND DISCUSSION

The assembly of a full ribosomes from ribosomal RNA (r-RNA) and ribosomal proteins (r-proteins) involve a number of steps. Ribosomal proteins, like other proteins, are synthesized in the cytoplasm. To be assembled onto r-RNA in the formation of a ribosomes, they must enter the nucleolus via the nucleus [17]. One of the latter r-proteins to be joined to a ribosomal core, is the low molecular mass acidic proteins located on the surface of the ribosomes. As mentioned earlier, the phosphorylated forms of acidic r-proteins are found on the ribosomal particle. Hence, the question arises whether protein kinases are localized in or out of the nucleus? To resolve this problem, we have carried out immunogold labeling [18] of PK60S as one of the three protein kinases phosphorylating the acidic r-proteins in the yeast cells (Fig. 1).

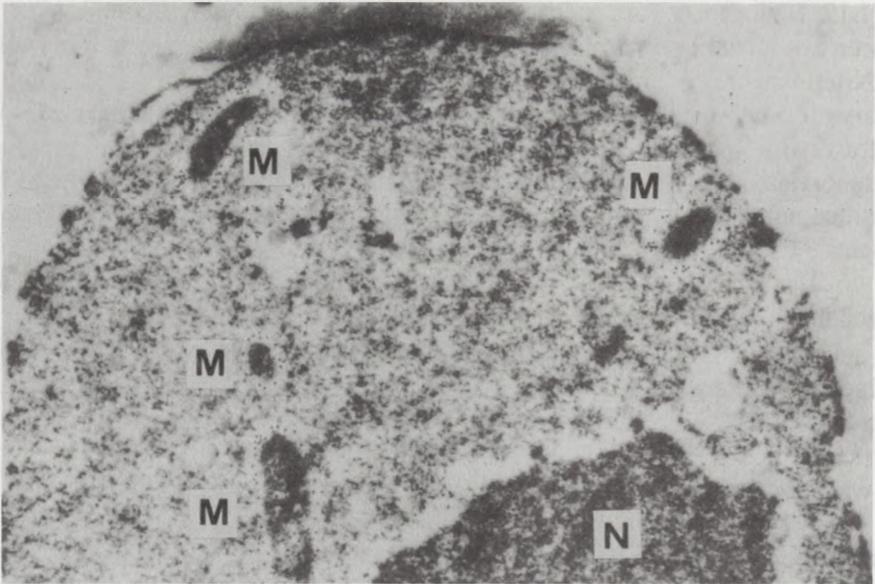


Fig. 1: Electron micrograph of immunogold labeling of ribosomal protein kinase - PK60S in *Saccharomyces cerevisiae* cells. Magnification approximately x 22.000, Nucleus (N), Mitochondrias (M).

The PK60S is present neither inside nor outside of the nucleus. Surprisingly, the mitochondria are surrounded by dark little dots which are the immunogold labeled particles of PK60S. It may suggest that PK60S has something to do in mitochondria. Isoelectrofocusing separation of P-proteins from mitochondrial ribosomes is shown in Fig. 2. They have only three proteins: YP1 α , YP1 β and YP2 β (Fig. 2B). The obtained result is quite curious, because the best two protein substrates for PK60S are YP1 β and the one missing in mitochondrial ribosomes YP2 α [19]. At the present time we have no explanation for that. In contrast to mitochondrial ribosomes, the cytoplasmic ones contain a set of four cardinal acidic proteins and one truncated form of YP1 β called YP1 β' (Fig. 2A) which is deprived of eight amino acids at the N-terminus [20].

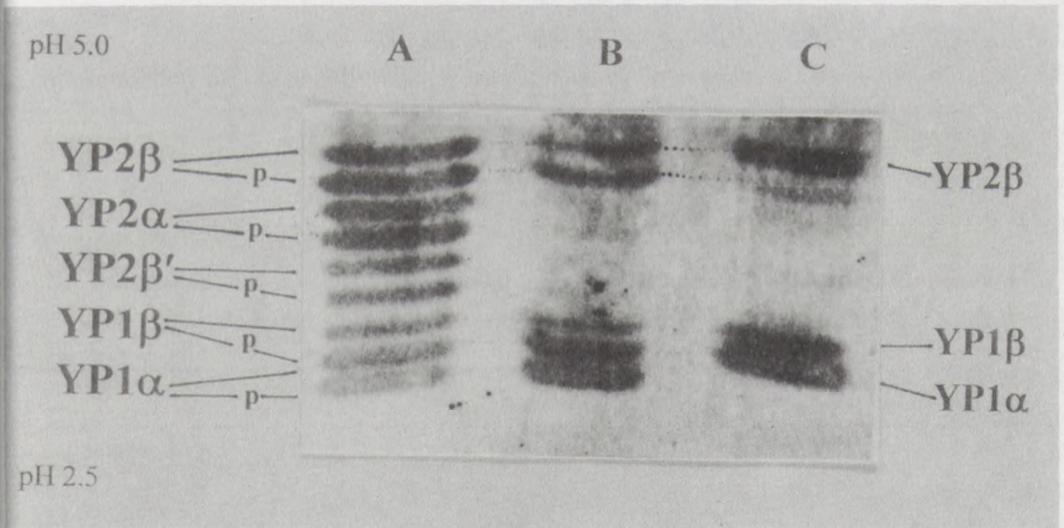


Fig. 2: Identification of acidic proteins from mitochondrial ribosomes. Separated proteins descending from cytoplasmic (A) and mitochondrial (B) native ribosomes were silver stained. The mitochondrial ribosomes were also dephosphorylated with alkaline phosphatase and acidic proteins were separated and stained (C). The small *p* letters mean phosphorylated form of acidic proteins.

So far, no data indicating whether phosphorylation of acidic r-proteins has whatever effect on the structure of ribosome. The native 80S ribosomes dissociate into the large-60S and small-40S subunits, which easily undergo separation by centrifugation (Fig. 3B). The isoelectrofocusing analysis of acidic r-proteins derived from the native yeast ribosomes shows the presence of the phosphorylated as well as the dephosphorylated forms of acidic r-proteins in the particle (Fig. 3b). Fully phosphorylated *in vitro* 80S ribosomes dissociate into subunits likewise to the native ribosomes (Fig. 3A and 3a). The graphic profile of subunit separation after

dephosphorylation of 80S ribosomes (Fig. 3C) differ from the native as from the phosphorylated ribosomes. It concerns the 60S subunits. In the gradient fractions where the single larger peak of 60S subunit should occur, there are two smaller peaks which have a complete set of acidic r-proteins (Fig. 3c). Dephosphorylation of ribosomes probably caused these perturbations in the structure of 60S ribosomes without splitting the acidic r-proteins leading to the formation of heavier forms of 60S subunits. These results suggest a protective role of acidic r-proteins in maintaining the native structure of the particle. This is important for the translational stability and readiness of a free 60S subunits appearing in the cytoplasm as a „run off” ribosomes ending the translation cycle and before the next one.

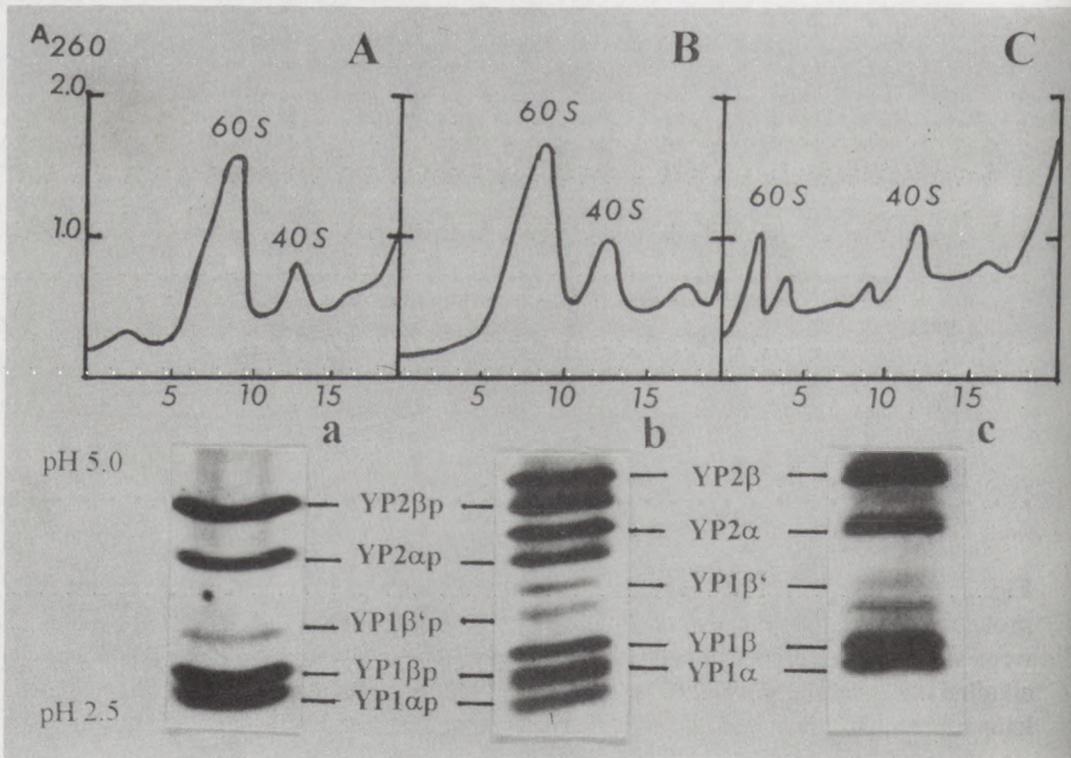
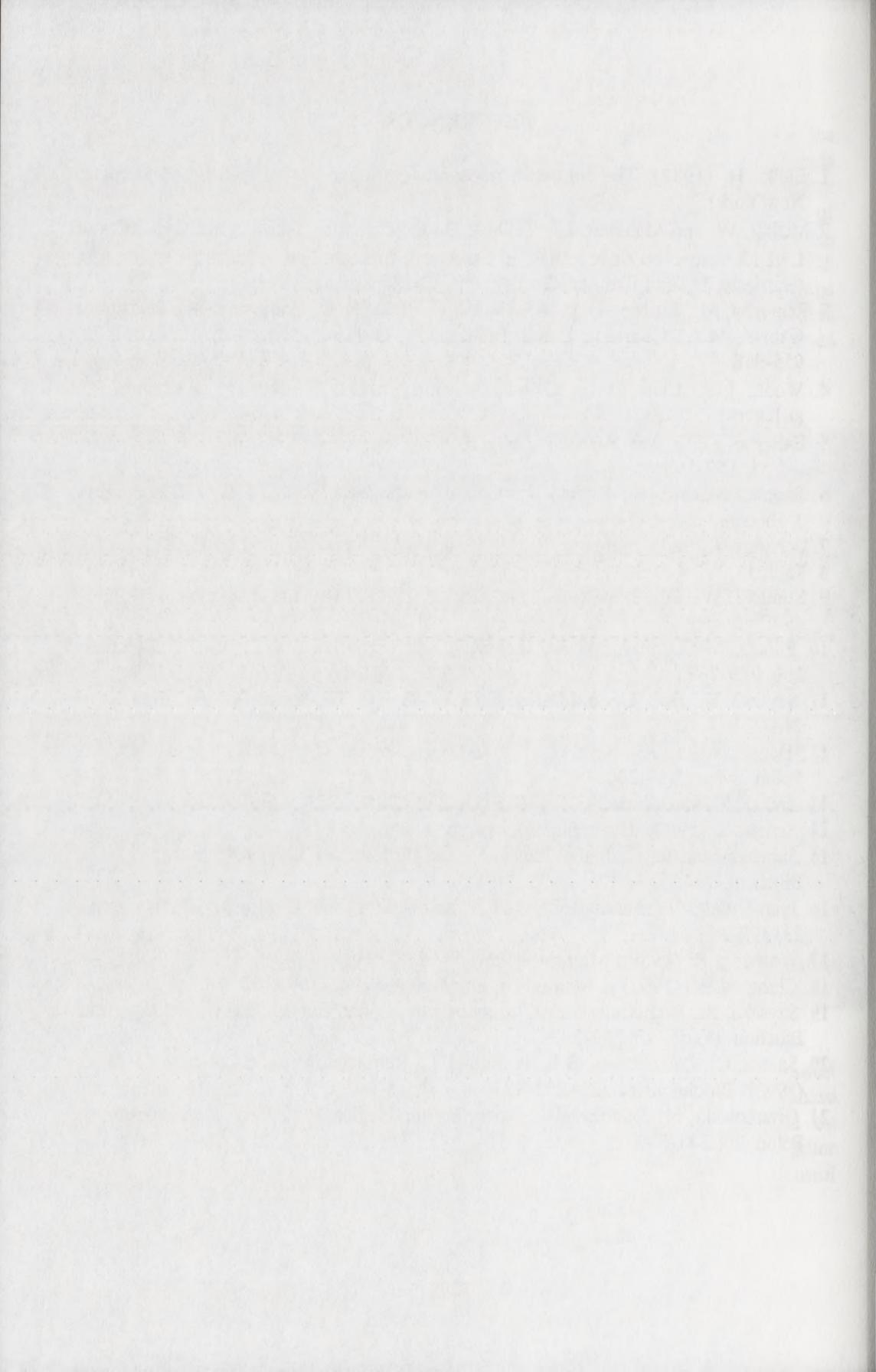


Fig. 3. Effect of phosphorylation - dephosphorylation processes on ribosomal structure. 80S ribosomes (1.5 mg) were phosphorylated *in vitro* with protein kinase (A) or dephosphorylated with alkaline phosphatase (C) and then separated on ribosomal subunits by centrifugation in a linear sucrose gradient as described earlier (21). The acidic r-proteins were analysed by isoelectrofocusing (a, c). The control batch was the native 80S ribosomes (B and b).

REFERENCES:

1. Bielka H. (1982): The eukaryotic ribosome. Springer-Verlag (Berlin, Heidelberg, New York).
2. Möller, W. and Maassen, J.A. (1986): On the structure, function and dynamics of L7/L12 from *E.coli* ribosomes. In Structure, function and genetics of ribosomes. (Hardesty, B., and Kramer, G. eds) New-York: Springer-Verlag.
3. Remacha, M., Jimenez-Diaz, A., Santos, C., Briones, E., Zambrano, R., Rodriguez Gabriel, M.A., Guarinos, E. and Ballesta, J.P.G. (1995): *Biochem. Cell Biol.* 73, 955-968.
4. Wool, I.G., Chan, Y.L., Glück, A., and Suzuki, K. (1991): *Biochemie* 73, 861-870.
5. Ballesta, J.P.G. and Remacha, M. (1996): *Progres Nucleic Acid Res. and Mol. Biol.* 55, 157-193.
6. Sanchez-Madrid, F., Reyes, R., Conde, P. and Ballesta, J.P.G. (1979) : *Eur. J. Biochem.* 98, 409-416.
7. Uchiumi, T., and Kominami, R. (1992): *J. Biol. Chem.* 267, 19179-19185.
8. Kudlicki, W., Grankowski, N., and Gąsior, E. (1976): *Mol. Biol. Rep.* 3, 121-129.
9. Kudlicki, W., Grankowski, N., and Gąsior, E. (1978): *Eur. J. Biochem.* 84, 493-498.
10. Pilecki, M., Grankowski N., Jacobs, J. and Gąsior E. (1992): *Eur. J. Biochem.* 206, 259-267.
11. Szyszka, R., Bou, G., and Ballesta, J.P.G., (1996): *Biochim. Biophys. Acta* 1293, 213-221.
12. Boguszewska, A., Szyszka, R., and Grankowski, N. (1997): *Acta Biochim. Polon.*, 44, 191 - 200.
13. Zinker, S., and Warner, J.T. (1976): *J. Biol. Chem.* 251, 1799-1807.
14. Zinker, S. (1980): *Biochim. Biophys. Acta*, 606, 76-82.
15. Sanchez-Madrid, F., Juan-Vidales, F., and Ballesta, J.P.G. (1981): *Eur. J. Biochem.* 114, 609-613.
16. Juan-Vidales, F., Saenz-Robles, M.T., and Ballesta, J.P.G. (1984): *Biochemistry* 23, 390-396.
17. Warner, J.R. (1989): *Microbiol. Rev.* 53, 256-271.
18. Clark, M.W. (1991) in *Methods in Enzymology* 194, 608-626.
19. Szyszka, R., Boguszewska, A., Grankowski, N., and Balesta, J.P.G. (1995): *Acta Biochim. Polon.* 42, 357-362.
20. Santos, C., Ortiz-Reyes, B.L., Naranda, T., Remacha, M., and Ballesta, J.P.G. (1993): *Biochemistry* 32, 4231-4236.
21. Grankowski, N., Kudlicki, W., Paleń, E., and Gąsior, E. (1976): *Acta Biochim. Polon.* 23, 341-352.



USE OF THE MATINSPECTOR PROGRAM AND GEL MOBILITY SHIFT ASSAYS FOR ANALYSIS OF REGULATORY SEQUENCES IN THE S100A6 (CALCYCLIN) GENE PROMOTER

Wiesława Leśniak, Gertruda Pokrywińska, Jacek Kuźnicki

Nencki Institute of Experimental Biology, 3 Pasteur 02-093 Warsaw, Poland
E-mail: jacek@nencki.gov.pl

Ca²⁺ signaling in signal transduction pathways is based on Ca²⁺ binding proteins. The most important group is the superfamily of EF-hand Ca²⁺ binding proteins representing about 250 highly homologous proteins. All of them share a common structural motif, which consists of a Ca²⁺ binding loop flanked by two α -helices. Calmodulin and a few other EF-hand proteins are present in all eukaryotic cells. Others, like calretinin and calyculin, are cell-specific. We study the structure, function, distribution, and regulation of the genes of these Ca²⁺ binding proteins. Specifically, the following projects are under investigation in our laboratory: cloning and expression of a novel calyculin target protein, the NMR structure of recombinant calretinin expressed in *Pichia pastoris*, the effect of overexpressed calretinin on Ca²⁺-homeostasis in cultured cells, the mechanisms of cell specific gene expression of calretinin, which is a neuronal protein, and of calyculin, which is expressed in epithelial cells and fibroblasts. The present article describes the recent developments in the latter project in which the calyculin human gene promoter is being studied.

Abstract

We describe the use of the MatInspector program for examining the putative transcription factor binding sites in the 1371 bp long calyculin gene promoter and an experimental evaluation of the data obtained from such a theoretical analysis. When the threshold of matrix probability was set at 0.92 the program indicated 23 transcription factors which could potentially recognize 59 different sequences. The binding of nuclear proteins from Ehrlich ascites tumor cells to the calyculin promoter fragments was examined. The protein-DNA interactions detected by means of gel mobility shift assays were fewer than these that can theoretically occur. However, we also observed a protein binding to the oligonucleotide corresponding to the putative enhancer sequence of the calyculin promoter and to the fragment representing -365/-256 bp that were not indicated by the computer search. This suggests a binding of transcription factors that were not included in the data base.

Introduction

Calcyclin, a member of the S100 family of calcium-binding proteins (for reviews see [1-3]), is expressed in a cell/tissue-specific manner [4-8]. Additionally, calyculin expression may be modulated in response to various extracellular factors [7-9] and an aberrantly high level of this protein is often observed in various pathological states, including many cancers [10, 11]. Cell specific expression and upregulation of the

protein under variety of conditions makes the calcyclin gene a suitable model to study the regulatory mechanisms directing the expression of the genomic information. The sequence of the calcyclin gene together with the putative 5' regulatory sequences have been established [12]. Functional studies using cells transfected with calcyclin gene promoter-reporter gene constructs are now also available [13, 14]. These studies helped to establish the so called minimal promoter, 5' proximal to the TATA box, which guarantees a base level of reporter gene transcription, and mapped the stimulatory or inhibitory properties of the upstream promoter regions in the investigated cell lines. However, no information is available concerning the binding of nuclear proteins to the calcyclin promoter DNA. Hence, no transcription factors that could be responsible for the observed stimulatory or inhibitory effects on transcription have been identified.

Recently, new methods have been published that help to evaluate the probability of a transcription factor binding to the examined sequences [15-17]. We believe the use of such an analysis may be helpful as a preliminary step in designing an experimental approach for studies on transcription factor binding. The combinatory use of this theoretical and experimental analyses might lead to the identification of novel transcription factors.

Materials and Methods

Theoretical analysis of the calcyclin gene

A human calcyclin gene clone isolated by Baserga and co-workers [13] was analyzed using the MatInspector program [15]. This program comes with a matrix library of about 200 transcription factors' matrices selected from the TRANSFAC database on transcription factors and their DNA binding sites [16, 17]. Parameter for core similarity (the four best-conserved consecutive nucleotides of the matrix) was chosen to be 1 (100%, full complementarity) and the threshold of minimal matrix similarity was arbitrarily set at 0.92 (92% complementarity). The calcyclin gene clone sequence, available under accession number JO2763 from the GenBank, was inserted to the MatInspector program and potential transcriptional factors binding to the promoter sequence (-1371/-1) and the respective DNA sequences were recovered. Overlapping binding sites for one transcription factor were considered as a single site, while overlapping sequences for two or more transcription factors were considered valid for each factor. Symmetrical sequences on the (+) and (-) DNA strands were considered as a single binding site.

Preparation of calcyclin gene promoter fragments

The 5' flanking region of the calcyclin gene, from position -1371 to +134 relative to the putative transcriptional start site was digested with pairs of restriction nucleases: *ApaI/NcoI*, *NcoI/BsmI*, *BsmI/EarI*, *EarI/EcoRV*, *EarI/SfiI*, *SfiI/BamHI*. This digestion yielded DNA fragments corresponding to the following regions of the

promoter: -1371/-1194, -1194/-1033, -1033/-748, -748/-476, -476/-277, -277/-52, -52/+134 (See Fig. 1). The resulting fragments were purified from a 1% agarose gel using Qiagen kit; DNA fragments were dephosphorylated, and subsequently labeled with [γ ³²P]ATP (>3000Ci/mmol) using T4 polynucleotide kinase (Promega). Free radiolabeled nucleotide was removed using Microspin G-50 columns (Pharmacia Biotech) according to the manufacturer instructions.

Gel mobility shift assays

Nuclear extracts (5-20 μ g of protein) were preincubated for 10 min with 1 μ g of poly(dIdC) and 1 μ g of herring sperm DNA in a buffer containing 20 mM Hepes pH 7.6, 1 mM EDTA, 10 mM (NH₄)₂SO₄, 1mM DTT and 0.1% Tween-20 in a final volume of 20 μ l. The ³²P labeled probe was added and samples were incubated for 20 min at room temperature. The incubation mixtures were resolved by electrophoresis on a 5% nondenaturing polyacrylamide gel in 1x TBE buffer. The gels were dried and analyzed by autoradiography. The oligonucleotides used were synthesized by TIB MOLBIOL, Poznań, and had the following sequences: Oligo 1 (USF core sequence underlined; NMYC, bold italic), **GTGCATGCCACGTGGGGAGC** (-598/-581); Oligo 2 and Oligo 2A (MZF1, underlined), TGAAGGGGAAATGGTGAA (-752/-735); TAGGGGGAGTAGCCAGTG (-463/-446); Oligo 3, (SV40 enhancer underlined) CAGGAGGCGTGGAAAGTC (-161/-144); Oligo 4 (SPE, underlined), TTGGCCGAGCTGGCCTC (-60/-43). For competition experiments the Competition Oligos kit (Stratagene) was used.

Nuclear extract preparations

Ehrlich ascites tumor cells were washed twice with 0.9% NaCl and pelleted at 1000 x g for 10 minutes. The cells were resuspended in a hypotonic lysis buffer containing 10 mM Tris-HCl, pH 7.4, 1 mM DTT, 0.25 mM PMSF, 0.05% sodium azide, 20 μ g/ml aprotinin, 20 μ g/ml pepstatin and 20 μ g/ml soybean trypsin inhibitor. After 10 min incubation on ice the cells were homogenized in a teflon-glass homogenizer. The nuclei were packed by a 10 min centrifugation at 1000 x g and resuspended in a buffer containing 10 mM Tris-HCl, pH 7.4, 0.25 mM sucrose, 0.25 mM PMSF and 0.05% sodium azide and subsequently pelleted at 3000 x g. The nuclei were resuspended in the same buffer containing 0.5% TRITON X-100, gently homogenized, pelleted as before and washed twice with the same buffer without detergent. The pellet was resuspended in the same buffer and extracted with 0.3 M KCl in substantia added gradually while the chromatin suspension was gently stirred on ice. The mixture was stirred for another 30 min, centrifuged as before, and the supernatant was dialysed 2 x 2 h against a buffer containing 10 mM Tris-HCl, pH 7.4, 40 mM KCl, 0.25 mM PMSF, 0.05% sodium azide. The dialysate was centrifuged at 3000 x g for 10 min to remove any precipitate; the supernatant was frozen and kept at -70°C. Protein content of the nuclear extract was measured by the method of Bradford.

Results

Table 1 shows the list of transcription factors whose putative binding sites were found throughout the 1371 bp long calcyclin gene promoter using the MatInspector program. When the threshold of matrix probability was set at 0.92, a limited set of

Table 1

Consensus sites for transcription factors in calcyclin promoter region (-1371/-1) with core = 1, and matrix values above 0.92

Transcription factors	No of sites	Core sequences	Positions
AP1/APIFJ	3	TGAC	-435(-), -321(-), -288(-)
AP4	1	CAGC	-252(-)
ARNT	2	CGTG	-591(+/-), -281(+/-)
ATF	2	TGAC	-1011(-), -231(-)
CAAT	1	CCAA	-1001(-)
CETS1P54	2	CGGA	-817(-), -49(-)
CREB	2	TGAC	-1007(-), -227(-)
CREL	1	TTCC	-169(-)
DELTAEF1	5	ACCT	-1165(-), -809(-), -798(-), -568(-), -522(-),
E47	2	CAGG	-1167(-), 799(-)
GATA	2	GATA	-853(-), 454(-)
IK2	6	GGGA	-1210(-), 1029(-), 946(-), 875(-), 596(-) -415(-)
LYF1	1	GGGA	-1209(-)
MAX/MYCMA	2	CACG	-593(-), -283(+/-)
MZF1	13	GGGG	-120(-), -1113(-), -963(-), -739(-), -710(-), -659(-), -631(-), -536(-), -495(-), -452(-), -294(-), -237(-), -193(-)
MYOD	2	CAGG	-1169(+/-), -799(+/-)
NF1	3	TGGC	-1231(-), -232(-), -805(-)
NFY	1	CCAA	-1004(-)
NKX25	1	AAGT	-1308(-)
NMYC	2	CGTG	-591(-), -281(+/-)
SP1	2	GGCG	-88(-), -73(-)
TH1E47	1	CTGG	-1153(-)
USF	2	CACG	-593(+/-), -283(+/-)

transcription factors was shown to have the potential to interact with the calcyclin gene promoter sequence. The analysis indicated 23 transcription factors which could potentially recognize 59 different sequences. The outcome of the search did not change significantly when the core similarity parameter was omitted (set at 0). Under such conditions the computer search indicated 6 additional sequences, only one of which was recognized by a factor not indicated previously. The sequences in the calcyclin promoter region could be recognized by both ubiquitous factors that include AP1, AP4, DELTA EF1, NF1, USF, as well as by cell specific factors such as MZF1, IK2 and DELTAEF1. Some factors had multiple potential binding sites (for instance MZF1 had 13 sites), but the majority had only one, or two binding sites. Some of these sequences are overlapping, indicating that the actual number of potential interactions is probably much smaller. Such overlapping sequences include two E-boxes (CANNTG) on both (+) and (-) DNA strands at regions -1167/-1161 and -798/-792 which can be recognized by MyoD, E47 and DELTAEF1. Another two overlapping binding sites at positions -596/-583 and -288/-273 included sequences recognized by USF, NMYC and ARNT.

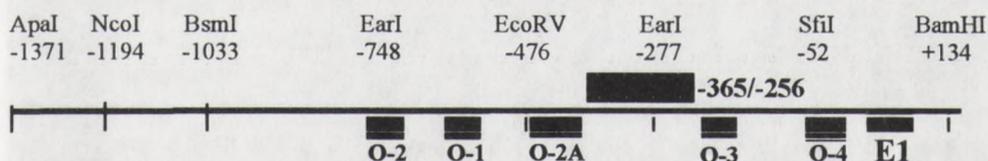


Fig. 1. Scheme of the calcyclin gene promoter region whose sequence is available [12]. Exon 1, fragments and oligonucleotides analyzed, as well as the restriction enzyme sites used to cut the promoter region are indicated.

To check some of the potential protein-DNA interactions indicated by the above described analyses we have performed binding studies using calcyclin promoter fragments or synthesized oligonucleotides and nuclear extracts from Ehrlich ascites tumor cells which express calcyclin at a high level. Promoter fragments corresponding to positions from -1033 to +134, *i.e.* extending 134 bp into the first noncoding exon, were used. Gel mobility shift assays (GMSA) performed under the conditions described in Materials and Methods show only a limited amount of DNA-protein interactions - a total of 10-12 bands were observed (Fig. 2A). According to the MatInspector search this part of the promoter (-1033/+134) should theoretically harbour 48 binding sites. When partially purified protein fractions were used several additional, weaker interactions could be observed (not shown). These results indicate that under the experimental conditions used only some of the theoretically possible DNA-protein interactions can be detected in nuclear extracts from cells expressing calcyclin. We have further compared the theoretical and experimental data by checking protein binding to short oligonucleotides (Fig. 1) harbouring binding sites for some of the transcription factors indicated by the computer search.

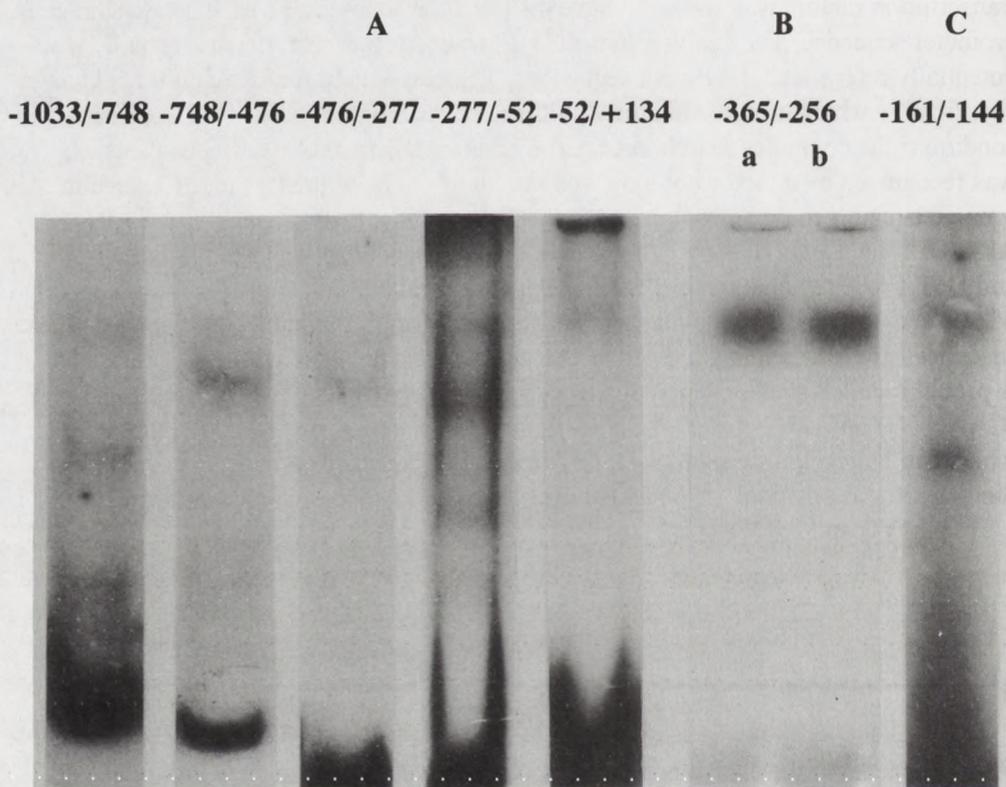


Fig. 2. Gel mobility shift assays using calcyclin gene promoter fragments and Ehrlich ascites tumor cells extracts. The fragments were labeled using T4 kinase and [$\gamma^{32}\text{P}$]-ATP and autoradiography of the dried polyacrylamide gel was performed overnight.

We have examined an 18 bp Oligo 1 corresponding to position (-598/-581) on the (-) DNA strand which contains overlapping binding sites for USF, NMYC, ARNT and IK2 transcription factors with the matrix similarities of 0.978, 0.956, 0.945 and 0.963 respectively. No protein binding to this oligonucleotide could be observed when the nuclear extract obtained from EAT cells was used (not shown). On the other hand, we observed a protein binding signal when we used a 110 bp long DNA fragment corresponding to position -365/-256 in which another overlapping putative binding site (-288/-273) for some of these factors (USF, NMYC, ARNT) was contained (Fig. 2B). To check whether this signal might be due to one of these factors binding to the -288/-273 bp region we have performed competition studies using Oligo 1, bearing an 8 bp long sequence identical to that contained in the -288/-273 region, as a competitor. Even at 50-fold excess, Oligo1 did not compete for the protein binding to the 110 bp long DNA fragment (Fig. 2B, line b). This suggests that none of the transcription factors indicated by the theoretical analysis (USF, NMYC, ARNT) was involved in this binding and indicates a possible binding of a protein not included in the data base.

This protein might represent a novel transcription factor. Other two examined oligonucleotides, Oligo 2 (-752/-735) and Oligo 2A (-463/-446) contained the MZF1 binding site at the (-) DNA strand (Fig. 1). No protein binding to these oligonucleotides was found under GMSA conditions (not shown) suggesting that active MZF1 protein is not present in the nuclear extract of EAT cells. On the other hand, we have observed a protein interaction with Oligo 3 (-161/-144) corresponding to the putative enhancer sequence of the calcyclin promoter [13] (Fig. 2C), although the computer search (matrix 0.92) did not indicate a transcription factor binding site in this sequence. The closest candidates were BARBIE and AHRARNT transcription factors with matrix similarity values of 0.874 and 0.858, respectively. Protein binding to the calcyclin gene enhancer sequence has been previously reported [13]. We have also tested a possible binding to the so called S100 protein element (SPE) - the only sequence that seems to be common to the promoters of almost all examined S100 proteins [20]. It is located at a similar distance from the TATA box in all S100 protein genes and was suggested to have a functional role [20]. The computer search of the SPE sequence GAGCTGGCCTC (-55/-45) did not indicate any transcription factor that could recognize it with a probability higher than 0.65 (core similarity 0, matrix similarity 0.65). In accordance with this prediction we could observe no binding to Oligo 4 corresponding to position -60/-43 of the promoter sequence (not shown). The computer search indicated two SP1 binding sites at positions -92/-79 and -76/-64 as well as three AP1 binding sites at positions -437/-427, -323/-314 and -290/-280. The AP1 binding sites are contained in the -476/-277 DNA fragment and the Sp1 binding sites in the -277/-52 DNA fragment, respectively. We have checked whether protein binding to these long DNA fragments can be competed out by commercially available oligonucleotides containing the consensus binding sites for these transcription factors. None of the competitive nucleotides was effective at 25-fold excess (not shown) suggesting that these transcription factors were not involved in the protein-DNA interactions observed for EAT cell nuclear extracts.

Discussion

The computer programs developed to analyze nucleotide patterns have been improved to utilize most of the available sequence information. The MatInd and MatInspector programs create a large library of sequence patterns and use them to locate matches in other sequences, employing a new search algorithm producing results superior to IUPAC searches [16]. These new programs establish a core sequence of the 4 most conserved nucleotides and start a preselection in which only matches to the core region, with the preset value, are considered. We have set the core value to 1, meaning that only sequences preserving the 4 consecutive most conservative nucleotides were considered. The matrix similarity threshold was set at 0.92. Although this choice was arbitrary, it was based on a test search performed for the ABF1 yeast transcription factor [15]. Out of 11 functional ABF1 binding sites, 10

had a matrix similarity ≥ 0.92 . Using these criteria (core = 1; matrix ≥ 0.92) we believe it was possible to select transcription factors which have the potential to interact with calcyclin promoter. To examine real interactions that may occur along the calcyclin promoter we have chosen EAT cells as the source of nuclear proteins. EAT cells are abundant in calcyclin [21], indicating a high level of calcyclin gene expression. We have experimentally checked protein binding to some sequences in the calcyclin promoter that were indicated as potential transcription factor binding sites. On the basis of these results several observations were made. The protein-DNA interactions detected by means of GMSA are fewer than those that can theoretically occur. This may be due to the fact that not all transcription factors, especially cell specific ones, may be present in a given cell type, *i.e.* the EAT cell extract. Also, some putative transcription factors might be inactive under the binding conditions used. Secondly, some weaker interactions may not be detected. In fact, additional weak signals could be seen in gels, when partially purified EAT cell nuclear extract fractions, instead of cellular extracts, were used. The observation that there is a fewer number of protein-DNA interactions than would be expected from the computer search might partly explain why we did not observe protein binding to oligonucleotides containing the consensus sequences. While MZF1 and IK2 are cell specific factors probably confined to the hematopoietic and lymphocyte lineages, with an unlikely involvement in calcyclin gene expression, USF and NMYC are ubiquitous. On the other hand, we have detected protein binding to the -365/-256 fragment that was not due to the binding of transcription factors (USF, NMYC, ARNT) found to recognize a sequence (-288/-273) present in this region. We have also observed protein binding to the putative enhancer sequence, to which no protein factors were ascribed, suggesting that novel transcription factors might be involved.

Our results show that the computer analysis performed using the MatInd and MatInspector programs can give a general scope of possible protein-DNA interactions. These interactions, however, must be experimentally verified for any given tissue or cell type. When a protein binding signal is detected the comparison with a theoretical search helps to establish which transcription factor may be involved or if a novel type of transcription factor should be considered. As the actual binding of a transcription factor depends on a cell type and intracellular processes, such as cell activation, phosphorylation, *etc.*, and on the experimental conditions, a negative experimental result does not automatically exclude the validity of the theoretical search. Hence, such interactions.

Acknowledgements

This work was supported by grant no 6 P04 01211 from the State Committee to Scientific Research.

References

1. Hilt, D. C. and Kligman, D. (1991) in *Novel Calcium-Binding Proteins* (Heizmann, C. W., ed.), pp. 65-103, Springer-Verlag, Berlin
2. Fano, G., Biocca, S., Fulle, S., Mariggio, M. A., Belia, S. and Calissano, P. (1995) *Progress in Neurobiology* **46**, 71-82.
3. Schafer, B. W. and Heizmann, C. W. (1996) *TIBS* **21**, 134-140
4. Kuźnicki, J., Filipek, A., Heimann, P., Kaczmarek, L. and Kamińska, B. (1989) *FEBS Lett* **254**, 141-144
5. Kuźnicki, J., Kordowska, J., Puzianowska, M. and Wozniewicz, B. M. (1992) *Exp Cell Res* **200**, 425-300
6. Timmons, P. M., Chan, C. T., Rigby, P. W. and Poirier, F. (1993) *J Cell Sci* **104**, 187-196
7. Leonard, D.G.B., Ziff, E.B., Greene, L.A. (1987) *Molec Cell Biol* **7**, 3156-3167
8. Gong, Y., Alkhalaf, B., Murphy, L. J. and Murphy, L. C. (1991) *Cancer Res* **51**, 1733-1737
9. Tonini, G. P., Casalaro, A., Cara, A. and Di Martino, D. (1991) *Eur J Biochem* **195**, 795-800
10. Weterman, M. A., Stoopen, G. M., van Muijen, G. N., Kuznicki, J., Ruiters, D. J. and Bloemers, H. P. (1992) *Cancer Res* **52**, 1291-1296
11. Berta G.N., Ghezzi, F., D'Avolio, A., Zulian, P., Carbone, V., Racca, S., Vercellino, V., Di Carlo, F. (1997) *J Oral Pathol Med* **26**, 206-210
12. Ferrari, S., Calabretta, B., deRiel, J. K., Battini, R., Ghezzi, F., Lauret, E., Griffin, C., Emanuel, B. S., Gurrieri, F. and Baserga, R. (1987) *J Biol Chem* **262**, 8325-8332
13. Ghezzi, F., Lauret, E., Ferrari, S., Baserga, R. (1988) *J Biol Chem* **263**, 4758-63
14. van Groningen, J. J., Weterman, M. A., Swart, G. W. and Bloemers, H. P. (1995) *Biochem Biophys Res Commun* **213**, 1122-1131
15. Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T. (1995) *Nucleic Acid Res* **23**, 4878-4884
16. Wingender, E., Dietze, P., Karas, H. and Knuppel, R. (1996) *Nucleic Acid Res* **24**, 238-241
17. Wingender, E. (1997) *Nucleic Acid Res* **25**, 265-268
18. Bottini F., Mazzocco, K., Abbondi, T., Tonini, G. P. (1994) *Neurosci Lett* **181**, 35-38
19. Allore, R.J., Friend, W.C., O'Hanlon, D., Neilson, K.M., Baumal, R., Dunn, R.J., Marks, A. (1990) *J Biol Chem* **265**, 15537-15543
20. Zimmer, D.B., Chessher, J., Song, W. (1996) *Biochim Biophys Acta* **1313**, 229-238
21. Filipek, A. and Kuźnicki, J. (1987) *Biochem J* **247**, 663-667

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14.
15.
16.
17.
18.
19.
20.
21.

THE STRUCTURE AND FUNCTION OF SYSTEMIN

Małgorzata Giel-Pietraszuk¹, Maciej Szymański¹, Genowefa Ślósarek², Thomas Specht³, Volker A. Erdmann³, Piotr Mucha⁴, Piotr Rekowski⁴, Gotfryd Kupryszewski⁴ and Jan Barciszewski¹

¹ *Institute of Bioorganic Chemistry, Polish Academy of Science, Poznan, Poland*

² *Institute of Physics, A. Mickiewicz University, Poznań, Poland,*

³ *Institute of Biochemistry, Free University, Berlin, Germany*

⁴ *Faculty of Chemistry, University of Gdansk, Poland*

Abstract: A synthetic palindromic DNA fragments of the promoter region of the tomato proteinase inhibitor I forms a complex with systemin. It was proved by capillary gel electrophoresis, footprinting with deoxynuclease I and circular dichroism (CD) measurements. We identified nucleotides protected by systemin and on this basis we performed computer modelling of the systemin-DNA complex.

Introduction

During the course of evolution, plants and pathogens have evolved a specific relationship, resulting from an exchange of a molecular information. In response to wounding and pathogens attack, plants developed a wide range of defence mechanisms. Recently much effort is devoted to the understanding of the expression of proteins that are synthesised in plants after pests or pathogen attacks. Wounding causes a release of local and systemic signals that induce expression of defence protein genes. Signalling molecules, regulating expression of those genes are suppose to meet several criteria. They should be synthesised by the plant, increase systematically after the attack by pathogen, move throughout the plant, induce defence-related substances and finally enhance resistance to wounding. Among regulatory molecules identified so far, there are low molecular weight compounds and short polypeptides (Pearce et al, 1991; Enyedi et al, 1992; Bernhamou, 1996; Becraft et al, 1996; van de Sande et al, 1996;). Recent reports from various laboratories, suggest that plants make wide use of a peptide signalling. A wound hormone, an 18-amino acid polypeptide called systemin, activates several defence genes (Pearce et al, 1991). Another peptide – ENOD40 – has been found to play a role as a regulator of the formation of nitrogen root nodules in legumes (van de Sande et al, 1996). The extracellular domain of the cr4 protein containing a cysteine-rich region similar to the ligand binding domain in mammalian tumour necrosis factor receptor (TNFR) and seven copies of a previously unknown 39-amino acid repeat function as a differentiation signal (Becraft et al, 1996). These peptides may be synthesised as inactive precursors e.g. prosystemin (McGurl et al, 1992), or as fully active molecules e.g. ENDO40 (van de Sande et al, 1996).

Systemin is transported in a phloem of tomato plants when placed on leaf wounds. A signalling pathway for defence gene activation has been proposed, in which

oligouronides and systemin activate the lipid-derived pathway (the octadecanoid pathway) where linolenic acid is released from membranes, resulting in the synthesis of jasmonic acid – a well known low molecular weight plant hormone (Doares et al, 1995). Systemin has been found to be processed *in vitro* at the putative furin cleavage site (Arg10-Asp11) by a plasma membrane associated protease and the systemin binding protein (SBP50) at the same time. The function of SPB 50 as a hormone receptor is not clear but it may play a role in the degradation of systemin (Schaller and Ryan, 1994). Low levels of systemin (femtomoles per plant) are active in inducing proteinase inhibitor synthesis. Systemin is synthesised in tomato as a 200 amino acids precursor protein, prosystemin. Its synthesis is essential for wound-induced proteinase inhibitors accumulation. Several other proteins including polyphenol oxidase, a sulfhydryl proteinase inhibitor, a cathepsin D inhibitor, carboxypeptidase, leucine aminopeptidase, aspartic proteinase and threonine deaminase are also systemically regulated (Schaller and Ryan, 1996). Systemin is translocated throughout phloem of tomato plants. Its movement is similar to that of sucrose and supports a role of this peptide as a systemic wound signal in tomato plants (Narvaez-Vasquez et al, 1995).

Although there are lot of data concerning biological function of systemin, very little is known about its structure and properties (Russell et al, 1992; Toumadje and Johnson, 1995). A molecular mechanism of action of the peptide is also not clear as well.

This work was undertaken to understand the structure and properties of systemin in solution and propose a model for its interaction with DNA.

Materials and methods

Systemin was synthesised manually using a solid phase procedure. The peptide was purified on HPLC RP Vydac C-18 column (Ślósarek et al, 1995).

Chemically synthesised DNA was labelled at the 5'end with [γ - 32 P]-ATP using T4 polynucleotide kinase (Sambrook et al., 1989). Radioactive DNA was purified on polyacrylamide gel and hybridised overnight in buffer containing 50 mM Tris-HCl pH 7.5, 2 mM MgCl₂. Guanine-ladder was prepared by chemical methods (Sambrook et al., 1989).

Ribosomal 5S RNA was isolated from yellow lupin seeds (*Lupinus luteus*) by phenol extraction, purified on Sephadex G-75 and repurified on 15 % polyacrylamide gel containing 7M urea, 50 mM Tris/borate buffer, pH 8.3 and 1mM EDTA (TBE). tRNA^{Phe} extracted from yeast was additionally purified on 15% polyacrylamide gel with 7M urea, 50 mM TBE buffer. 5S rRNA and tRNA were 3'-end labelled with [32 P]pCp and T4 RNA ligase. [32 P]-labelled RNAs were purified by 10 % polyacrylamide gel electrophoresis (PAGE) with 7M urea, eluted from the gel and renatured (Barciszewska et al., 1986).

The complex formation of DNA with systemin was carried out using 1nM of labelled DNA (40 000 cpm) and appropriate amount of systemin as specified in the legend to the figures, in Buffers: B1) 20mM Tris-HCl pH 7.5, 50 mM KCl, 5mM MgCl₂, 0.5 %

Nonidet P-40; B2) 20mM Tris-HCl pH 8.2, 50 mM KCl, 5mM MgCl₂, 0.5 % Nonidet P-40 and B3) 20mM Tris-HCl pH 7.5, 70 mM KCl, 2 mM MgCl₂, 0.1% Nonidet P-40, 0.5 mM DTT, 50μM ZnCl₂, 100μM BSA, 6% glycerol (10 μl of total volume). The reaction mixture was incubated 40 min. at 22°C. The complexes formation of systemin with [3'-³²P] labelled crude tRNA from yeast, RNA^{Phe} from yeast and 5S rRNA (*L. luteus*) were analysed similarly. Appropriate amount of systemin was incubated 40 min. at 22°C with RNA in buffer containing 20 mM Tris-HCl, 100 mM KCl and 1mM MgCl₂.

DNase I footprint analysis was performed as follows: DNA-systemin complex formation was carried out using 1nM of labelled DNA (40 000 cpm) and 0.8 mM systemin in buffer containing 20mM Tris-HCl pH 7.5, 70 mM KCl, 2 mM MgCl₂, 50 μM ZnCl₂, 0.5 mM DTT, 0.1% Nonidet P-40, 6% glycerol (10 μl of total volume). Reaction mixture was incubated for 40 min. at 22°C. After that 0.2 U and 0.02 U of DNase I was added and samples were incubated 1min at 22°C. Reactions were terminated by addition of 1μl of 100 mM EDTA solution. Digestion products were analysed by electrophoresis on 20% polyacrylamide gels with 7M urea (Sambrook et al., 1989).

Capillary electrophoresis was done on Beckmann apparatus and circular dichroism (CD) spectra were recorded on AVIV (New York) spectrometer.

Computer modelling was performed on Silicon Graphics Indigo2 workstation with the Insight and Discover software package (version 95.0) from Biosym Technologies, Inc. (San Diego, USA) using Amber forcefield.

Results

Recently we analysed the tertiary structure of systemin (¹Ala-Val-Gln-Ser-Lys-Pro-Pro-Ser-Lys-Arg-Asp-Pro-Pro-Lys-Met-Gln-Thr-Asp¹⁸) with two-dimensional NMR spectroscopy. By measuring the spectra at conditions different from the other authors (Russell et al, 1992), we have been able to identify the *cis* and *trans* isomers and to conclude that in solution, systemin can adopt a *Z*-like-β-sheet structure (Ślósarek et al, 1995). This finding was in some contrast to conclusion reached by Russell et al, who have not been able to find any persistent secondary or tertiary structure by NMR. However, they have noted two distinct weak molecular conformations at the C terminus of the peptide in solution near neutral pH (Russell et al, 1992). An analysis of their data suggests that those conformers could be similar to these observed by us. Furthermore, NMR analysis has not confirmed a possibility of a helical structure for systemin, which has been proposed by analogy to poly(L-proline) II (Toumadje and Johnson, 1995). Systemin is unusual peptide. It contains two pairs of proline residues separated by tetrapeptide Ser-Lys-Arg-Asp and can be hardly compared with a typical polyproline II helical motif as e.g. recently studied peptide: Pro-Arg-Pro-Pro-Arg-Pro-Pro-Arg-Pro-Pro-Asp-Pro-Pro (Gresh, 1996). A conformational stability of short peptides is very interesting problem. Several recent papers however exemplify, that they can show up bioactive structure (Perez et al, 1996).

It is very well known that the β -sheet domain has been found in various DNA binding proteins (Efimov, 1994). If so, one can suggest that systemin showing up such motif should bind to DNA. This was the case and we showed that random DNA immobilised on cellulose retards systemin, (but not a random peptide) which can be further eluted from the affinity column with 0.2 M sodium chloride (Ślósarek et al, 1995). These observations prompted us to study in more details specificity of DNA - systemin interactions. It is already known that among others, systemin activates synthesis of proteinase inhibitors I and II in tomato (Schaller and Ryan, 1996). Such effect could be explained through direct interaction with the promoter region of these genes. Two synthetic oligodeoxy dodecamers form double-stranded palindromic fragment of the tomato proteinase inhibitor I promoter and form the complex with systemin as was shown in capillary electrophoresis (Fig.1). The data presented on Fig. 1 indicated that the peak corresponding to the dsDNA with migration time (M_t) 20 min. after addition of peptide is shifted. A new peak corresponding to the DNA-systemine complex appears with M_t 23 min. These result was confirmed by gel retardation analysis of DNA and RNA systemine complex formation (Fig. 2). Results show that systemin form a complex with DNA (Fig. 2A), but not with crude tRNA or 5S rRNA. To learn more about the specificity of interaction we synthesised a 30 nucleotides long double stranded DNA, a fragment of -96 to -65 the tomato serine proteinase inhibitor I promoter containing palindromic dodecamers (Keil et al, 1986). 30-mer of DNA in complex with systemin was footprinted with deoxynuclease I. Two concentration of DNase I were used in these experiments. Lower concentration (Fig. 3. lanes 3, 5) visualised structure of 3'-end of DNA molecule in the complex. As one can see, the nucleotides 11-13, 17-21 on top strand and 11'-16', 20'-23' on lower strand are clearly protected from DNase I. Additionally weaker protection of nucleotides 8, 9 and 9', 10' are visible. Strong hydrolysis of position 27 in complex can suggest conformational changes of DNA. That was additionally confirmed by the CD-spectra measurement where maximum of the Cotton effect of DNA at 265nm decreases in the presence of systemin (Fig. 4). Shift of the Cotton effect maximum to higher waves suggest changes in DNA conformations. One can interpret them that B-DNA after addition of peptide adopt A form.

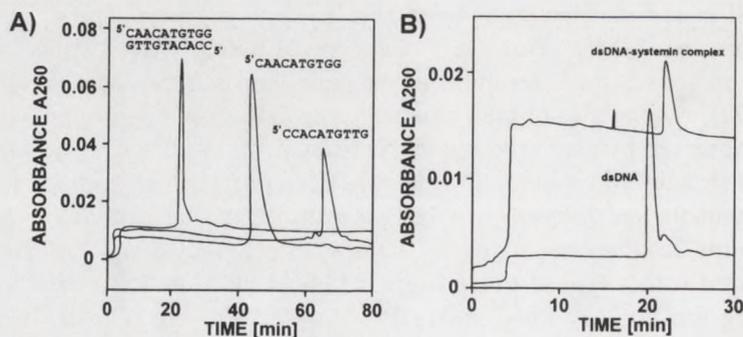


Fig. 1 Capillary electrophoresis analysis of double stranded DNA formation (A) and DNA-systemin complex migration (B).

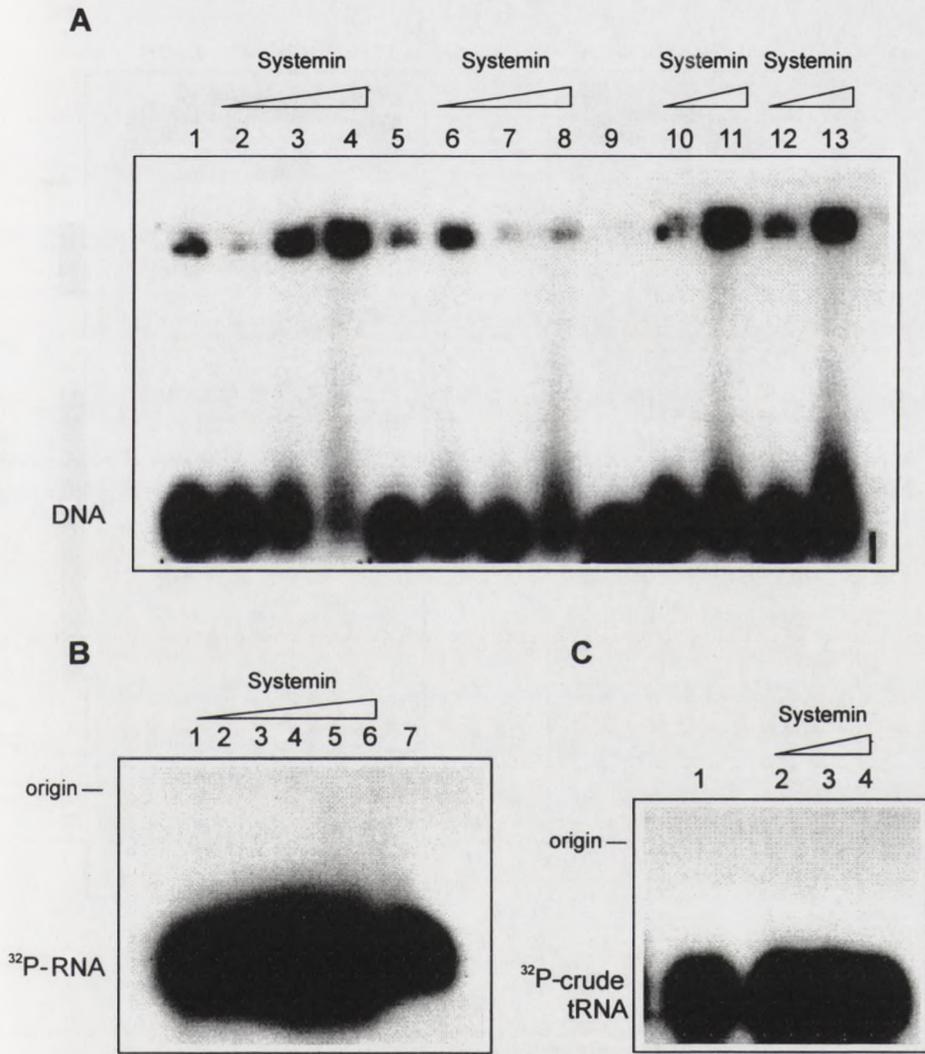


Fig. 2. An agarose gel-shift analysis of the systemin binding to A: 1nM [5'-³²P] DNA lanes 1,5,9 control, 2,6,10 DNA with 0.8 mM systemin, 3,7,11 DNA with 2.4 mM systemin, 4,8 DNA with 4.0 mM systemin, B: [3' - ³²P] labelled 5S rRNA (lanes 1-3) with 0.8, 1.6, 2.4 mM of systemin and [3' - ³²P]tRNA^{Phe} (lanes 4-6) with 0.8, 1.6, 2.4 mM of systemin and free tRNA as a control (lane 7) C: [3' - ³²P] labelled crude tRNA, 1-control, 2-4 crude tRNA with 0.8, 1.6, 2.4 mM of systemin respectively. Reaction were carried out in buffers B1 lanes 1-4, buffer B2 lanes 5-8, buffer B3 lanes 9-11. Complex formation was analysed on 0.7% agarose gel.

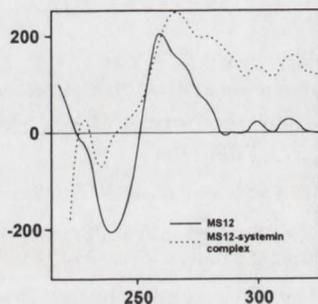


Fig. 4 CD spectra of synthetic oligodeoxynucleotide duplex (MS12) and its complex with systemin

To better understanding the binding properties of systemin to the promoter of proteinase inhibitor I gene, we carried out computer studies of the complex (Fig.5). DNA was built according to standard B DNA geometry. The systemin structure was taken from NMR data analysis (Ślósarek et al, 1995). As one can see, the peptide having the Z-like β -sheet structure can easily be docked into the major groove of DNA. We identified several interactions: amino group of arginine 10 interacts with oxygen 4 (O4) of thymine residue (T6), lysine 5 binds to phosphate oxygen of A5, the oxygen backbone proline 7 is in close contact with N7 of adenine 3. Also serine 8 and glutamine 16 interact with phosphate oxygen of A3 and T6 respectively. Furthermore lysine 14 interacts with phosphate oxygen. This model is in agreement with the experimental data shown in Fig.1.



Fig. 5 Hypothetical model of systemin - DNA interaction. It was obtained by docking of the systemin NMR structure (Ślósarek et al, 1995) to the promoter region of the tomato proteinase inhibitor I gene (-96 to -65). One turn of space filling DNA double helix model is shown. Systemin (stick model) is located in the major groove of B DNA.

Conclusion

We showed that systemin having Z-like β -sheet structure forms a complex with DNA, but not with RNA. Several specific interactions have been identified by DNase I footprint experiments of the complex. Docking of the peptide on B-DNA confirms several contacts of bases and backbone with amino acids.

References

- Barciszewska, M., Mashkova, T.D., Zwierzynski, T., Kisselev, L.L. & Barciszewski, J. 1986. The primary structure of yellow lupin seeds 5S ribosomal RNA. *Bull. Acad. Polon. Sci.* 34, 369-373.
- Becraft, P. W., Stinard, P. S., McCarty, D. R., 1996. CRINKLY4: A TNRR receptor kinase involved in maize epidermal differentiation. *Science*, 273, 1406-1409.
- Bernhamou, N., 1996. Elicitor-induced plant defence pathways. *Trends in Plant Science*, 1, 233-240.
- Doares, S. H., Syrovets, T., Weiler, E.W., Ryan, C. A., 1995. Oligogalacturonides and chitosan activate plant defensive genes through the octadecanoid pathway. *Proc. Natl Acad. Sci. USA*, 92, 4095-4098.
- Efimov, A. V., 1994. Favoured structural motifs in globular proteins. *Structure*, 2, 999-1002.
- Enyedi, A. J., Yalpani, N., Silverman, P., Raskin, I., 1992. Signal molecules in systemic plant resistance to pathogens and pests. *Cell*, 70, 879-886.
- Gresh, N., 1996. Can polyproline II helical motif be used in the context of sequence-selective major groove recognition of B-DNA? A molecular modelling investigation. *J. Biomol. Struct. Dynam.*, 14, 255-273.
- Keil, M., Sanchez-Serrano, M., Schell, J., Willmitzer, L., (1986). Primary structure of a proteinase inhibitor II gene from potato. *Nucleic Acids Res.*, 14, 5641-5650.
- McGurl, B., Pearce, G., Oronzca-Cardenas, M., Ryan, C.A., 1992. Structure, expression and antisense inhibition of the systemin precursor gene. *Science*, 255, 1570-1573.
- Narvaez-Vasquez, J., Pearce, G., Orozco-Cardenas, M. L., Franceschi, V. R., Ryan, C. A., 1995. Autoradiographic and biochemical evidence for the systemic translocation of systemin in tomato plants. *Planta*, 195, 593-600.
- Pearce, G., Strydom, D., Johnson, S., Ryan, C. A., 1991. A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins, *Science*, 253, 895-898.
- Perez, J. J., Sharkey, M. and Centeno, N. B., 1996. On the bioactive conformation of a small peptide and its set of thermodynamically accessible conformations. *J. Biomol. Struct. Dynam.*, 14, 185-191.
- Russell, D.J., Pearce, G., Ryan, C. A., Satterlee, J.D., 1992. Proton NMR assignments of systemin. *J. Prot. Chem.*, 11, 265-274.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning. A laboratory manual, second edition, Cold Spring Harbor Laboratory Press.
- Shaller, A., Ryan, C. A., 1994. Identification of a 50 - kDa systemin - binding protein in tomato plasma membranes having Kex2p - like properties. *Proc. Natl Acad. Sci. USA*, 91, 11802-11806.
- Shaller, A., Ryan, C. A., 1996. Systemin - a polypeptide defence signal in plants. *BioEssays*, 18, 27-33.
- Ślósarek, G., Kalbitzer, H. R., Mucha, P., Rekowski, P., Kupryszewski, G., Giel-Pietraszuk, M., Szymański, M., Barciszewski, J., 1995. Mechanism of the activation of proteinase inhibitor synthesis by systemin involves β -sheet structure, a specific DNA-binding protein domain. *J. Struct. Biol.*, 115, 30-36.
- Toumadje, A., Johnson, W.C., 1995. Systemin has the characteristics of a poly(L-proline) II type helix. *J. Amer. Chem. Soc.*, 117, 7023-7024.
- Van de Sande, K., Pawlowski, K., Czaja, I., Wieneke, U., Schell, J., Schmidt, J., Walden, R., Matvienko, M., Wellink, J., Van Kammen, A., Franssen, H., Bisseling, T., 1996. Modification of phytohormone response by a peptide encoded by ENOD40 of legumes and a nonlegume. *Science*, 273, 370-373.

INDEX OF AUTHORS

Alexandratos Jerry	89
Andrake Mark	89
Apostoluk Włodzimierz	49
Banecki Bogdan	119
Barciszewski Jan	141
Bierzyński Andrzej	107
Buczek Olga	49
Bujacz Grzegorz	89
Błaszczak Adam	119
Chaubet Nicole	27
Cholawska Liliana	49
Ciarkowski Jerzy	101
Czaplewski Cezary	101
Dadlez Michał	49
Dukowski Piotr	125
Dąbrowska Renata	37
Dłużniewska Joanna	27
Erdmann Volker A.	141
Esaki Nobuyoshi	31
Giel-Pietraszuk Małgorzata	141
Gigot Claude	27
Grankowski Nikodem	125
Grzesiak Agnieszka	49
Hayashi Hideyuki	11
Jaskólski Mariusz	85
Jerzmanowski Andrzej	27
Kagamiyama Hiroyuki	11
Kamio Yoshiyuki	95
Kato Nobuo	65
Katsube Yukiteru	79
Katsuragi Naruto	79
Katz Richard	89
Kawata Yasushi	113
Kaźmierkiewicz Rajmund	101
Kitagawa Yasuyuki	79
Konieczny Igor	119
Kościelska Katarzyna	49
Krokoszyńska Izabela	49
Krowarsch Daniel	49
Kupryszewski Gotfryd	141
Kuraś Mieczysław	27

Kurihara Tatsuo	31
Kuźnicki Jacek	131
Legocki Andrzej B.	71
Leśniak Wiesława	131
Liberek Krzysztof	119
Lim Seong-Il	79
Makuch Robert	37
Marszałek Jarosław	119
Masaki Takeharu	79
Merkel George	89
Mucha Piotr	141
Nariya Hirofumi	95
Nishiyama Akihito	95
Norioka Shigemi	79
Ohta Shigeki	79
Ostrowski Włodzimierz S.	15
Otlewski Jacek	49
Pilecki Marek	125
Pokrywińska Gertruda	131
Polanowski Antoni	59
Politowska Ewa	101
Prymakowska-Bosak Marta	27
Przewłoka Marcin	27
Rekowski Piotr	141
Sakai Yasuyoshi	65
Sakiyama Fumio	79
Sansom Clare E.	85
Sashi Kazumichi	79
Shiraki Kentaro	79
Siemion Ignacy Z.	21
Sikorski Michał M.	71
Skalka Anna Marie	89
Soda Kenji	31
Specht Thomas	141
Spiker Steven	27
Stachowiak Damian	49
Szymański Maciej	141
Ślósarek Genowefa	141
Ślusarczyk Joanna	27
Tomaszewski Radosław	27
Wawrzynów Alicja	119
Wilusz Tadeusz	59
Włodawer Alexander	89
Yagi Kunio	45
Żylicz Maciej	119

