

## 대두 저장단백질 유전자의 발현 조절 메카니즘

최 양 도 . 김 정 호  
(서울대학교 . 농화학과)

## Regulation Mechanism of Soybean Storage Protein Gene Expression

Yang Do Choi and Chung Ho Kim  
(Dept Agricultural Chemistry, Seoul National University)

### Abstract

Glycinin and  $\beta$ -conglycinin are the most abundant storage protein in soybean. These proteins are known to be synthesized predominantly during germination and cell expansion phase of seed development for short period, and synthesized not in other tissues. Genes encoding these storage proteins are useful system to study the mechanism of development stage and tissue specific gene expression in eukaryotes, especially plants, at the molecular level. The cDNA and genomic clones coding for glycinin have been isolated and regulation mechanism of the gene expression has been studied. Initially, development and tissue-specific expression of the glycinin gene is regulated at the level of transcription. Post-transcriptional processing is also responsible for delayed accumulation of the mRNA. Translational control of the storage protein gene has not been reported. Post-translational modification is another strategic point to regulate the expression of the gene. It is possible to identify positive and/or negative regulatory elements in vivo by producing transgenic plants after gene manipulation. Elucidation of activation and repression mechanism of soybean storage protein genes will contribute to the understanding of the other plant and eukaryotic genes at molecular level.

## 서      론

고등 식물의 발달 분화는 발달 특이성 유전자의 발현에 의해 유도되며 조직 특이성 유전자의 발현은 조직의 발달 및 특성 유지에 절대적인 역할을 한다 (Goldberg 1986, 1988). 대두의 조직 특이성 유전자의 예를 들면 표 1 과 같다. 녹색식물 특유의 잎을 포함한 광합성 조직에서 발현이 되어 광합성에 관여하는 ribulose 1,5-bisphosphate carboxylase small subunit (rbcS) 과 chlorophyll a/b 결합단백질 (cab) 의 유전자, 대두 특유의 질소고정균에 의해 생성된 뿌리혹에서 발현이 되는 nodulin 유전자, 그리고 대두의 종자에서 발현이 되는 종자단백질 유전자 등이 있다.

Table 1. Tissue-specific Genes of Soybean

| genes             | plant | organ/tissue | references               |
|-------------------|-------|--------------|--------------------------|
| rbcS              |       | green        | Shirley et al. 1987      |
| cab               |       | green        | Simpson et al. 1986      |
| leghemoglobin     |       | nodule       | Jensen et al. 1986       |
| uricase II (N35)  |       | nodule       | Nguyen et al. 1985       |
| sucrose synthase  |       | nodule       | Morell and Copeland 1985 |
| nodulin N23       |       | nodule       | Joergensen et al. 1988   |
| beta-conglycinin  |       | seed         |                          |
| alpha'-subunit    |       |              | Chen et al. 1986, 1988   |
| beta-subunit      |       |              | Barker et al. 1988       |
| glycinin (11S)    |       | seed         | Goldberg et al. 1981     |
| lectin            |       | seed         | Okamuro et al. 1986      |
| trypsin inhibitor |       | seed         | Jofuku et al. 1989       |

두과 식물의 종자에는 발아 및 초기 성장 단계에서 탄소 및 질소원으로 활용되는 저장 단백질이 다량 존재한다. 대두의 저장단백질의 유전자 발현은 대두의 생활 주기에 따라 매우 정확하게 조절되는데 단백질 생합성은 배형성 단계에서 종자의 숙성과 함께 시작되며 떡잎에서만 대량 발현이 되고 숙성한 종자 및 다른 기관 및 조직에서는 발현이 되지 않는 것으로 알려져 있다. 이와 같이 대두 저장 단백질 유전자는 대두의 분화 발달 과정에서 매우 정확하게 조절되며 유전자의 활성이 매우 왕성해서 식물 유전자의 조직특이성 및 분화 발달 특이성 발현 조절 메커니즘을 연구하는데 있어서 매우 바람직한 system 으로 인식되어진다.

이러한 특이성 유전자의 발현은 대체로 전사 수준에서 조절이 이루어지지만 전사 후 단계, 예를들면 pre-mRNA 의 가공 및 전달, 전이 및 전이 후 변형 단계에서도 조절이 이루어지기도 한다 (Goldberg *et al.* 1981, 1989; Walling *et al.* 1986). 여기서는 대두의 저장단백질 유전자를 예를 들어 그들의 발현 조절 메커니즘에 대해 언급하기로 한다.

유전자의 발현 조절 메커니즘은 최근 transgenic plant 의 생산과 함께 분자수준의 이해가 가능해졌다. 발달 및 분화에 따른 유전자의 발현 특성이 담배를 비롯한 다른 전입된 host plant 에서도 원래 식물 내에서 처럼 재현이 가능해졌기 때문이다. 유전자의 발현 조절 부위를 절단 혹은 치환 변형한 후 타식물체 내에 전입시켜 발달, 분화 및 조직에 따른 발현을 확인하는 과정은 유전자의 발현 조절 기능을 생체 (세포) 내에서 검정할 수 있는 가장 좋은 system 으로 인식된다 (Benfey and Chua 1989). 이는 다수의 유사한 유전자를 각각 분리하여 개개의 발현을 연구하기에도 좋은 system 이다.

## 본 론

대두는 예로부터 동양에서 단백질의 중요한 공급원으로 이용되어 왔다. 대두

저장단백질의 대부분은 5% 의 염용액으로 추출할 수 있으며 globulin 계통에 속한다 (Pusztai and Stewart 1980). 이것은 sucrose density gradient centrifugation 에 의해 침강 계수가 2S, 7S 그리고 11S 인 3개의 주된 단백질 부분으로 분리할 수 있다 ( Hill and Breidenbach 1974). 2S 부분은 여러가지 효소와 단백질 분해 효소 억제제 (proteinase inhibitor), lectin 등을 포함한 단백질 혼합체로 구성되어 있으며 7S 와 11S 부분은 콩의 주된 저장단백질인 conglycinin ( $\alpha, \beta, \gamma$ ) 과 glycinin 에 각각 해당된다. 이중 glycinin 은 legumin type 으로 가장 양이 많아서 전체 대두 단백질의 약 50 - 60% (종자 건물중의 20%) 를 차지하며 vicilin type 인  $\beta$ -conglycinin 이 그 다음이다. (Hill and Breidenbach 1974; Nielsen 1984, 1985).

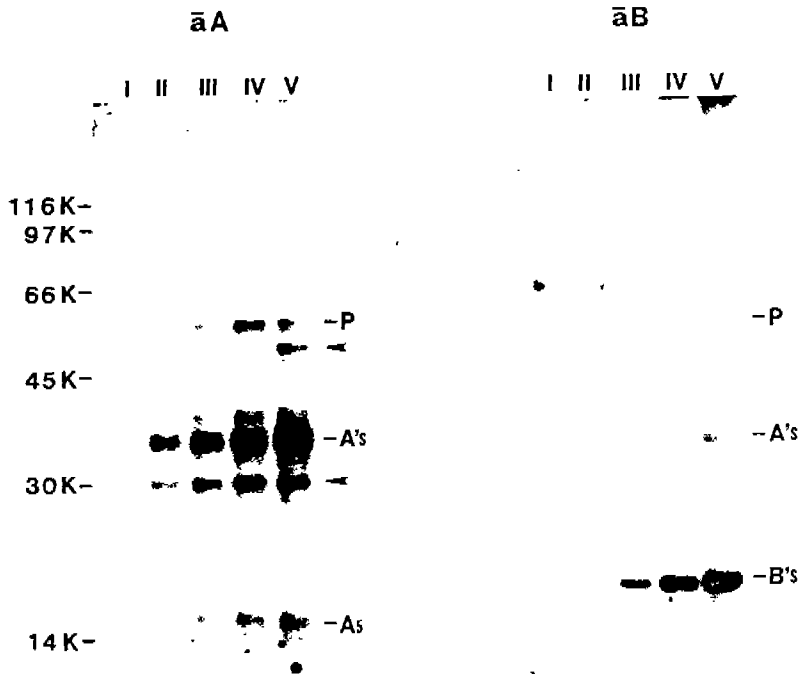


Fig. 1. Immunoblot analysis of glycinin in developing soybean seed. Immature soybean seed were collected and divided into 5 groups by size. Crude extract was analyzed by immunoblot with anti-glycinin A or anti-glycinin B polypeptide antibody. Panel aA, anti-A polypeptide antibody; panel aB, anti-B polypeptide antibody; lane I, extract from seed at stage I; lane II, extract from seed at stage II; lane III, extract from seed at stage III; lane IV, extract from seed at stage IV; lane V, extract from seed at stage V.

Glycinin 은 수정이 일어난 직후에 종자의 생성 및 숙성과정에 걸쳐 정해진 시기에 그리고 비교적 짧은 시간에 걸쳐 불과 몇 안되는 종류의 분자가 대량으로 생합성 되어지며 잎이나 기타 조직에서는 발현이 되지않는 것으로 알려져 있다 (Goldberg *et al.* 1981). 대두 종자의 배 형성 및 숙성과 더불어 생성되는 glycinin 의 양적인 변화는 그림 1 과 같다. 종자의 발달 및 숙성과정에서 분리한 단백질을 SDS-PAGE 와 immunoblot 으로 분석한 결과에 의하면 미숙종자에서는 생합성되는 glycinin 이 극히

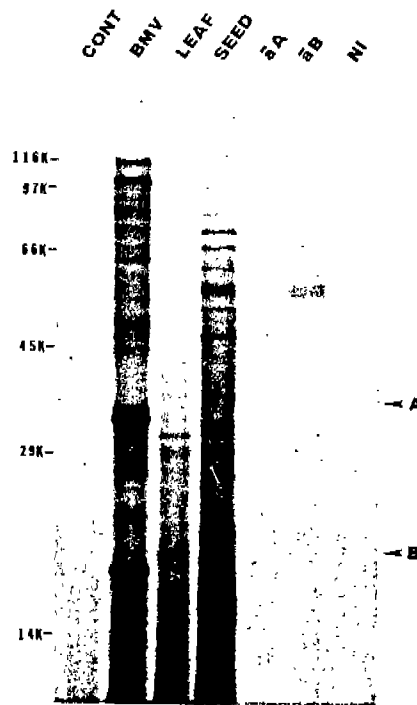


Fig. 2. *In vitro* translates of mRNA isolated from soybean. mRNA was isolated by phenol extraction and oligo(dT)-cellulose chromatography and *in vitro* translation was in wheat germ extract. They were separated by 12.5% SDS-PAGE and visualized by fluorography. [4,5-<sup>3</sup>H] leucine was used as a label. The position of glycinin A- and B-polypeptide are shown by arrow heads. Lane CONT, no exogenous mRNA was added; lane BMV, bromo mosaic virus RNA was added; lane LEAF, soybean leaf mRNA was added; lane SEED, soybean seed mRNA was added; lane aA, immunoprecipitation with anti-A antibody; lane aB, immunoprecipitation with anti-B antibody; lane NI, immunoprecipitation with non-immune serum.

적지만 종자의 속성과 더불어 급격히 증가하는 것을 알 수 있다. 이러한 경향은 유전자에 따라 다소의 차이는 있지만 종자 단백질의 공통된 특성이다 (Goldberg *et al.* 1981). 또한 이러한 합성 단백질량의 증가는 mRNA 수준의 증가와 대체로 비례하는 것으로 알려졌다 (Nielsen *et al.* 1989; Harada *et al.* 1989). 이러한 유전자의 발현 형태는 발달 분화 특이성 유전자 발현의 대표적인 예이다.

이러한 단백질 생합성의 조절 메커니즘을 이해하기 위하여 미숙 종자에서 mRNA를 분리하여 *in vitro* translation 을 실시한 결과는 그림 2 와 같다. 앞에서 분리한 mRNA 에 의해 합성되는 단백질의 종류와 대두 종자에서 분리한 mRNA 에 의해 합성되는 단백질은 근본적인 차이를 나타내고 있다. 대두 종자에서 분리한 mRNA 는 주로 분자량이 13.5K, 18K, 21.5K, 48K, 59K, 65K, 68K 인 polypeptide 를 code 하지만 앞에서 분리한 mRNA는 대체로 분자량이 14K 와 30K polypeptide 를 code 한다. 이러한 결과는 조직 특이성 유전자 발현 형태를 단적으로 증명하고 있다. *In vitro* translation 결과 합성된 glycinin 을 확인하기 위하여 anti-glycinin 항체를 사용하여 면역침전법을 실시한 결과 대두 종자에서 분리한 mRNA 에 의해 합성된 polypeptide 중에서 분자량이 58K인 polypeptide 가 항체와 반응하였다. 대두 세포 내에 존재하는 glycinin을 추출하여 SDS-PAGE 로 분석하면 분자량이 약 39K (A-polypeptide) 와 19K (B-polypeptide) 인 polypeptide로 이루어진 것을 알 수 있다 (그림 1). 그러나 *in vitro* translation 결과 mRNA 로 부터 합성된 polypeptide 에는 이에 해당하는 크기의 polypeptide 를 관찰할 수 없었다. 대신에 anti-A polypeptide 와 anti-B polypeptide 항체는 다 같이 분자량이 58K인 polypeptide 와 반응하였다. 이는 glycinin subunit polypeptide 는 mRNA 로 부터 분자량이 58K인 전구체 형태로 합성된 후 post-translational endopeptidic cleavage 에 의해 분자량이 39K인 A-polypeptide 와 분자량이 19K인 B-polypeptide 로 전환된다는 것을 의미한다 (Barton *et al.* 1982; Tumer *et al.* 1981; Kim and

Choi 1989a). 이러한 결과는 cDNA cloning 에 의한 mRNA 구조와 일치한다 (Marco *et al.* 1984; Kim and Choi 1989b). 그러나 이러한 post-translational modification 에 관여하는 효소 및 인자는 *in vitro* translation system 에는 존재하지 않는 것으로 보아 소포체, protein body, lysosome 등 세포 기관 혹은 기타 세포 인자를 필요로 하는 것으로 여겨진다.

SDS-PAGE 에 의하면 7S 와 11S 단백질은 여러개의 subunit 으로 구성되어 있다. 7S 부분의 주된 단백질 beta-conglycinin 은 vicilin type 의 glycoprotein (4-5% 탄수화물 포함) 으로  $\alpha'$  (분자량 76K),  $\alpha$  (분자량 72K) 그리고  $\beta$  (분자량 53K) subunit 등 3개의 산성 polypeptide subunit 이 약 6 가지 종류의 다양한 조합으로 존재하며 분자량은 약 150K - 200K 이다 (Bray and Beach 1985; Doyle *et al.* 1986; Schuler *et al.* 1982). 11S 단백질 glycinin 은 분자량이 360K 로서 6 개의 subunit (분자량 약 60K) 으로 구성되어 있다. 각각의 subunit은 다시 분자량이 45K - 37K이고 주로 산성 아미노산으로 이루어진 A-polypeptide 와 분자량이 약 20K이고 주로 염기성 아미노산으로 이루어진 B-polypeptide가 disulfide 결합에 의해 이루어져 있다 (Moreira *et al.* 1981; Staswick *et al.* 1984). A-polypeptide 는 그 종류가 다양하며 아미노산 분석 및 peptide sequencing 결과 5 가지 정도의 전하가 다른 형태 (A1a, A1b, A2, A3, A4, A5) 가 존재하며, B-polypeptide 는 5 가지 정도의 전하가 다른 형태 (B1a, B1b, B2, B3, B4) 가 존재하는데 각 subunit 은 이들 A- 및 B-polypeptide 가 특정한 조합으로 존재하는데 A1aB2, A1bB1b, A2B1a (이상 group I), A3B4, A4B3, A5B3 (이상 group II) 등 이다 (Ereken-Tumer *et al.* 1982; Moreira *et al.* 1981; Staswick *et al.* 1984). 비록 구성 아미노산 서열의 미세한 차이가 있으나 각 subunit 상호간에는 매우 유사한 구조 (homology)를 갖고 있는데 각각의 group 내에서는 80 - 90% 의 유사성이, group 상호간에는 50% 정도의 유사성이 있다 (Moreira *et al.* 1981; Staswick *et al.* 1984). 이들 subunit 은 각각의 mRNA 에

의해 생합성 되어지지만 subunit의 미세한 다양성 (microheterogeneity)은 전이 후 변형에 의해서 이루어지기도 한다 (Fukazawa *et al.* 1985; Hirano *et al.* 1984; Marco *et al.* 1984). 이러한 전이 후 변형은 식물의 종자 단백질에서 매우 보편적으로 관찰된다.

Glycinin 유전자의 발현조절 및 단백질의 구조를 분자 생물학적 수준에서 연구하기 위하여 glycinin 의 cDNA 유전자와 genomic 유전자를 분리하였다. 그림 3 은 분리한 glycinin subunit A2B1a 의 cDNA 유전자의 염기서열 이다. 이 cDNA 유전자에는 1 개의 open reading frame 이 존재하는데 481 개의 아미노산으로 이루어진 polypeptide 를 code 한다. 대두 종자에서 분리한 glycinin subunit polypeptide 의 아미노산 서열과 cDNA 유전자로 부터 예상되는 polypeptide 의 아미노산 서열을 비교하면 그림 3과 같다. 대두 종자에서 분리한 glycinin polypeptide의 아미노산 서열과 그림 3에서 예상되는 아미노산 서열을 비교하면 이 cDNA 유전자는 5' 쪽에 278 개의 아미노산으로 구성되어 있는 A2-polypeptide 를, 3' 쪽에 180 개의 아미노산으로 구성되어 있는 B1a-polypeptide 를 code 하며 이 2 부분은 동일한 reading frame 에 존재하고 2 부분 사이에는 연결부위 (linker) 에 해당하는 4 개의 아미노산이 존재한다. 이 결과는 앞서 *in vitro* translation 결과와 일치하는 것을 알 수 있다. mRNA의 전이 결과 전구체가 합성되며 이것은 전이 후 변형 단계인 endopeptidic cleavage 에 의해 A- 및 B-polypeptide 가 생성되는 것을 알 수 있다. 따라서 이들 A- 및 B-polypeptide 는 항상 특정한 조합으로 존재하게 되며 이들은 동일한 mRNA 에서 생산된 것들이다. 이 전구체는 N-terminal 쪽에 18 개

Fig. 3. Nucleotide sequence of the glycinin A2B1a cDNA clone. Nucleotide sequence and deduced amino acid sequence are shown. Parts of polypeptides processed out posttranslationally are shown as shadows. N-terminal part of the open reading frame corresponds to the A2-polypeptide and C-terminal part to the B1a-polypeptide.





아미노산으로 이루어진 signal polypeptide 가 존재하며 C-terminal 쪽에 5 개 아미노산이 더 존재한다. 따라서 glycinin 의 생합성 과정에는 매우 여러단계의 가공 과정이 존재하는 것을 알 수 있으며 이 과정은 glycinin 의 발현에 매우 중요한 인자로 작용하는 것으로 생각된다 (Dickinson *et al.* 1987,1989; Negoro *et al.* 1985).

Signal peptide 의 존재로 볼때 glycinin 은 소포체 표면에서 합성되어지는 것으로 여겨진다. 이들 저장단백질은 조면소포체 (rER) 의 표면에서 전구체 형태로 생합성되면서 rER 내부로 전달이 이루어지며 이때 signal peptide 가 cotranslational 하게 절단된다 (Chrispeels *et al.* 1982; Dickinson *et al.* 1987, 1989).  $\beta$ -conglycinin의 경우 여기서 glycosylation 이 일어나며 7S 형태로 조합이 이루어진다. pro-glycinin 은 여기서 9S (180K, trimer) 조합을 형성하는 것으로 알려져 있다 (Barton *et al.* 1982). 이 trimer 는 Golgi apparatus 를 거치면서 가공이 이루어진 후에 vacuolar protein body 에서 proteinase 에 의해 endopeptidic cleavage 가 일어나서 A- 및 B-polypeptide 가 생성되며 이들 사이에 disulfide 결합이 이루어지며 이들 trimer 가 모여서 11S 의 hexamer 형태로 저장되는 것으로 알려져 있다 (Dickinson *et al.* 1987,1989). 11S glycinin hexamer의 구조는 전자 현미경 및 X-ray scattering 에 의해 알려진 바에 의하면 대체로 그림 4와 같이 glycinin trimer 가 2 층으로 배열된 모습으로 추측된다. 그러나 *in vitro* assembly 결과에 의하면 proglycinin trimer 는 결코 hexamer 를 형성하지 않는것으로 나타났다. 따라서 이러한 전사 후 변이 과정은 합성된 단백질이 목표하는 세포 내 위치에 저장되는 과정에서 매우 유기적으로 작용하는 것을 알 수 있으며 중자단백질의 합성 및 저장 과정에서 중요한 조절 단계의 역할을 할 것으로 여겨진다 (Dickinson *et al.* 1987,1989). 이러한 현상은 대두의 trypsin inhibitor (Kunitz type) 에서도 관찰되는 것으로 C-terminal processing 과 세포 내 최종 위치 결정 및 수송과정의 단백질 활성화와의 유기적인 관계를 예측할 수 있다.

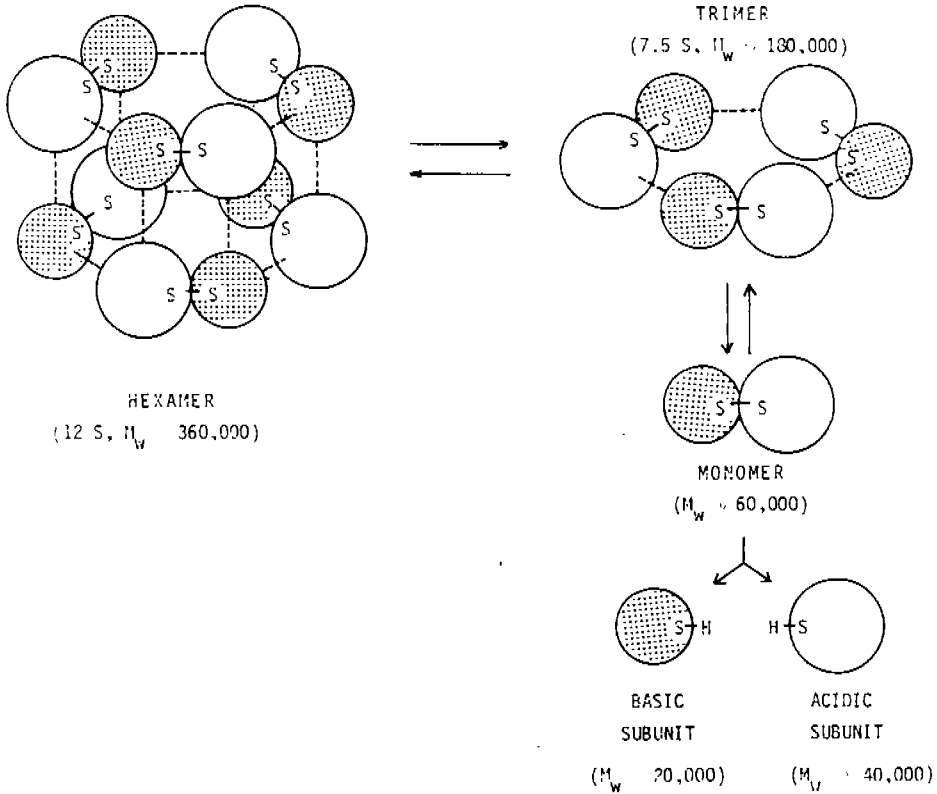
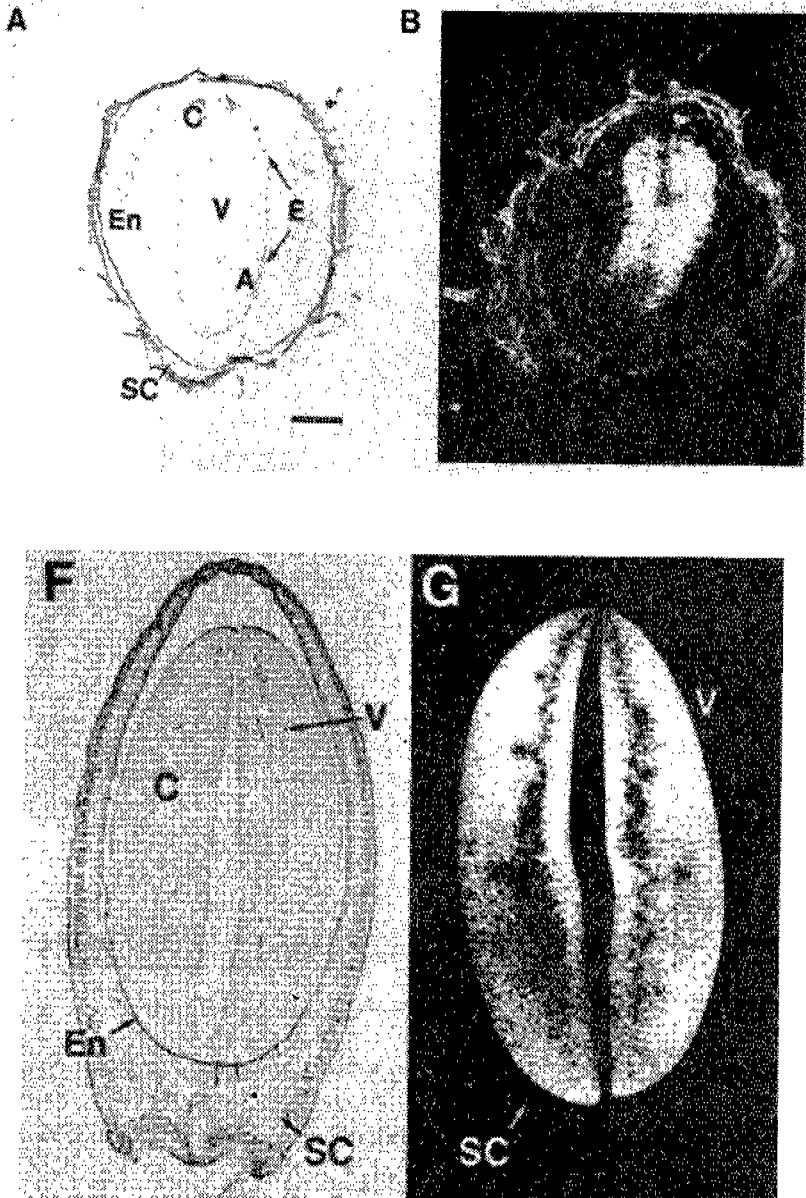


Fig. 4. Schematic model of glycinin quaternary structure. Figure is taken from Pernollet and Mosse (1983).

저장 단백질 유전자는 종자의 발생과 더불어 발현된다. 종자의 생성과정에서 저장 단백질의 생합성 부위를 검정하기 위하여  $\beta$ -conglycinin 의  $\beta$ -subunit 유전자를 담배에 도입하여 transgenic plant 를 얻고 이 담배 종자의 발생 과정에서  $\beta$ -conglycinin 유전자의 mRNA 존재 부위를 *in situ* hybridization 에 의해 결정하였다 (그림 5A) (Barker *et al.* 1988). Beta-conglycinin mRNA 는 담배씨의 내배유 (endosperm) 에서는 검출되지 않았지만 배세포 (embryonic cell) 인 자엽 저장 유세포 (cotyledon storage parenchyma cell) 에서 대량으로 관찰되었다. 이는 종자 발달 과정에서 종자 단백질 유전자의 조직 특이성 발현 형태를 의미한다. 이러한 결과는 대두 종자에서도 확인되었다 (그림 5B) (Barker *et al.* 1988; Goldberg *et al.* 1989). 그러나 담배에



**Fig. 5.** Localization of  $\beta$ -conglycinin mRNA in developing embryo. Embryo was from transgenic tobacco transformed with  $\beta$ -conglycinin (A) and soybean (B). *In situ* hybridization was carried out with  $\beta$ -conglycinin probe. Figures are taken from Barker *et al.* (1988) and Goldberg *et al.*

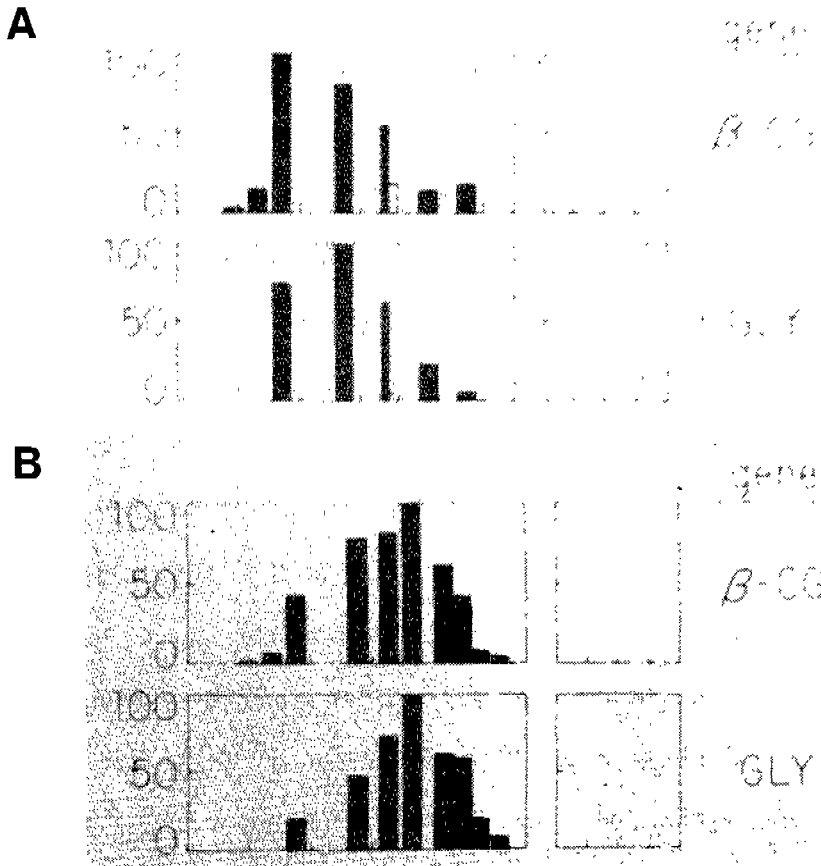


Fig. 6. Relative transcription rates and accumulation of seed protein genes during the soybean life cycle. A, relative seed protein gene transcription rates during development; B, changes in seed protein mRNA level during development. Figures is taken from Walling *et al.* (1986).

비해 대두 종자는 거의 대부분이 자엽 배 세포로 이루어져 있으며 내배유는 종자 발생 초기에만 일부 존재하며 beta-conglycinin 발현 단계인 배 형성의 숙성 단계에는 거의 존재하지 않는다 (Barker *et al.* 1988).

Glycinin 의 조직특이성 발현을 결정하는 가장 중요하고 근본적인 단계는 결국 transcription 단계인 것을 *in vitro* translation 과 Northern blot analysis 결과 알 수 있다. Glycinin mRNA 는 배 형성 초기에 대량 생산되어 숙성 중기에서 최대에 이르고 완숙 종자에서는 감소하는 경향을 보인다 (Goldberg *et al.* 1981a,b ; 1989)

(그림 6). Glycinin mRNA 수준의 변화를 보면 개화 후 35 일에 최초로 발견된 후 70일 - 87일 사이에 최대치에 이르고 그 후 차츰 감소하여 완숙 및 건조 과정인 100일이 경과한 후에는 전혀 인지되지 않는다. Glycinin subunit 유전자 군의 발현은 모두가 이와 유사한 동시에 유기적으로 발현 조절되는 것으로 알려졌다 (Nielsen *et al.* 1989; Harada *et al.* 1989). 그러나  $\beta$ -conglycinin 의 경우 2.5 kb mRNA ( $\alpha/\alpha$ ) 의 수준은 1.7 kb mRNA ( $\beta$ -subunit) 및 glycinin mRNA 에 비해 약 2주 먼저 최대치에 도달하는 것으로 알려졌다. 그러나 이들  $\beta$ -conglycinin subunit 유전자들의 transcription 은 이 기간 동안 같은 속도로 일어나는 것으로 알려졌다 (Goldberg *et al.* 1989). 또한 이들 mRNA 수준과 transcription 속도의 관계는 분리한 핵의 run-off transcription 실험 결과를 보면 대체로 전사 속도가 먼저 최대가 되며 mRNA의 수준은 그 후 약 2주 정도 지난 후 최대가 되는 것으로 알려졌다. 따라서 post-transcriptional control에 의해 유전자 발현이 조절된다는 것을 의미한다 (Nielsen *et al.* 1989; Goldberg *et al.* 1989; Walling *et al.* 1986). 이것은 배 세포 핵 내 인자가 있어서 선택적 운반을 도우거나 배 세포질에 인자가 있어서 배 형성 단계에서 mRNA 를 일정화시키는 것으로 여겨진다 (Walling *et al.* 1986; Goldberg 1986). 또한 저장 단백질 mRNA 는 종자의 숙성과 함께 최대치에 이른 후 차츰 감소하며 이 단계에서는 transcription이 선택적으로 억제되는 것을 알 수 있다.

이러한 전사 조절 메카니즘을 이해하기 위하여 glycinin 의 genomic 유전자를 분리하였다 (Nielsen *et al.* 1989; Kim and Choi 1989c). Glycinin 유전자는 beta-conglycinin (약 15개) 과 달리 5개의 유전자에 의해 code 되는 비교적 간단한 구조를 하고 있으며 이들 모두 glycinin 유전자는 4개의 exon 과 3 개의 intron 으로 구성되어 있다. Glycinin subunit A2B1a를 code하는 Gy2 유전자의 염기서열을 결정한 결과는 그림 7과 같다. 유전자의 발현을 조절하는 것으로 생각되는 5' 주변을 살펴보면 매우 흥미있는 모습을 볼 수 있다 (표 2).

; pBH400  
: 5' UPSTREAM REGION OF GLYCININ GY2

-328 GGATCCCACTCTATTGTCACACGGTGTATTATATAATTCATATGTCATAAACC-276

-275 CGCCGAACATGAAATGAAAGCATTGTCCCCCTCCTCCACCAGCGTTTTCTGGCAAT -221

-220 TGCATGCAATACAACACACTTGGTATTTGTCACATAATGTTGATGTCGAACTGT -166

-165 CGAAGCCACCTCAGACCCATGAACTTAATGAGGTGTAACACACAGGCTTCATA -111

-110 GCAATGCAACTGAAGAATGTCTCAAGCTCAGCACCCCACTCCTGTGACGTGTCC -56

-55 CTCACCCACCTTCCTCTCTCCCTATAAATAACCAGCCTCAGGTTCTCCGCTTC -1

+1 ACAACACAAACATTCTCTCCATTGTCTTTGAACACTCATCACC ATG GCC AAG CTT  
Met Ala Lys Leu

Fig. 7. Nucleotide sequence of the genomic clone of glycinin subunit A2B1a, Gy2. The 5' flanking region of the gene is shown. Nucleotide numbering starts from transcriptional initiation site, +1, and to backward. Several consensus sequence elements are noted.

Eukaryotic 유전자에서 공통으로 존재하는 TATA box (Proudfoot 1979)와 유사 CAT box (AGGA box, Messing *et al.* 1982) 가 존재하며 종자 단백질 유전자에서 발견되는 RY repeat (CATGCAT; Dickinson *et al.* 1988) 가 한번, HAACACAMH (Goldberg 1986) 가 두번 존재하고 legumin consensus sequence (Gatehouse *et al.*

Table 2. Sequence elements found in glycinin gene Gy2

| sequence   | position                                 | characteristics       | references            |
|--|--|-----------------------|-----------------------|
| CATGCAT<br>(CATGCAT)   | -105                                     | seed protein specific | Dickinson et al. 1988 |
| AAACACACT<br>TAACACACA<br>(HAACACAMH)                                  | -204<br>-126                             | seed protein specific | Goldberg 1986         |
| TCCATAGCCATGCA<br>TACTGAAGAATGTC<br>(TCCATAGCCATGCA<br>TGCTGAAGAATGTC) | -109                                     | legumin box           | Gatehouse et al. 1986 |
| GCCACCTC<br>(GCCACCTC)   | -158                                     | vicilin box           | Gatehouse et al. 1986 |
| ATCCCA<br>AACCCG<br>CCCCCT<br>CACCCA<br>ACCCCA<br>(AVCCCA)             | -324<br>-276<br>-246<br>-149, -50<br>-74 | conglycinin enhancer  | Chen et al. 1986      |
| TATT ATATAA<br>AATTCAT AT  | -299<br>-291                             |                       |                       |



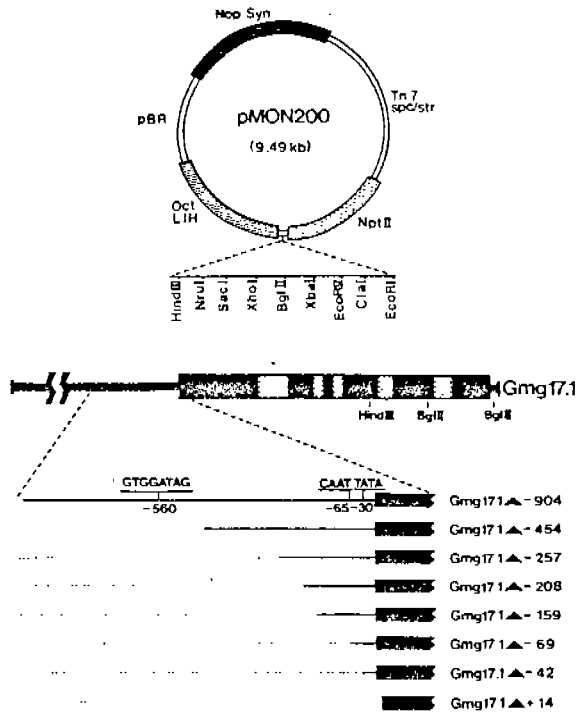
|               |      |                           |                     |
|---------------|------|---------------------------|---------------------|
| AACTAATGAG    | -140 |                           |                     |
| (AATTTAATTAA) |      | embryo factor binding     | Jofuku et al. 1987  |
| GTGTATTA      | -303 |                           |                     |
| TTGGTATT      | -196 |                           |                     |
| GAGGTGTA      | -133 |                           |                     |
| GTGTAACA      | -129 |                           |                     |
| GTGACGTG      | -63  |                           |                     |
| (GTGGWWWG)    |      | SV 40 enhancer            | Weiher et al. 1983  |
| TATAAAT       | -29  |                           |                     |
| (TATAAAT)     |      | TATA Box                  | Proudfoot 1979      |
| GAAT          | -94  |                           |                     |
| (GAAT)        |      | AGGA Box                  | Messing et al. 1982 |
| CACCAUGGC     | +45  |                           |                     |
| (AACAAUGGC)   |      | translation<br>initiation | Lutcke et al. 1987  |

---

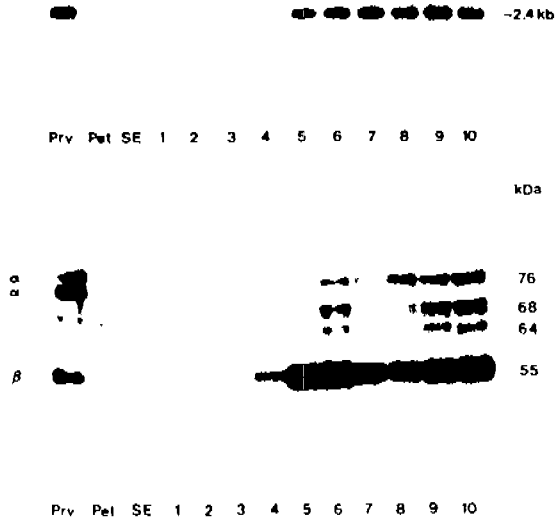
1986) 가 한번, embryo factor binding sequence (Jofuku *et al.* 1987) 가 3 번 존재하였다. 이외에도 SV40 enhancer sequence (GTGGWWWG; Weiher *et al.* 1983)가 5번 반복하고 conglycinin enhancer (AVCCCA; Chen *et al.* 1986, 1988) 가 5 번 반복하였다. 그러나 동시에 vicilin consensus sequence (GCCACCTC; Gatehouse *et al.* 1986) 가 존재하기도 한다. 그러나 이러한 인식 부위는 그 자체로서는 그리 큰 의미를 지니지 못한다.

이러한 조절 인자의 기능을 검증하기 위하여 transgenic plant 를 사용한 *in vivo* 실험이 가능해 졌다.  $\beta$ -conglycinin 의  $\alpha'$ -subunit 유전자의 발현을 조절하는 것으로 여겨지는 5' 부위에 대해 그 길이가 각각 다른 deletion mutant 를 재조합 한 후 petunia 에 도입하여 beta-conglycinin 의 발현을 검증하였다. 그림 8A 에서와 같이 이 부위에는 몇가지 중요한 인식 부위가 보이는데 TATA box, CAAT box 그리고 SV40 enhancer 등이 눈에 뜨인다. 그러나 이러한 transgenic plant 에 의한 *in vivo* 검증 결과 B-conglycinin 유전자의 경우 발달 및 조직 특이성 transcription 을 조절하는 부위는 transcription 이 시작되는 부위에서 5' 쪽으로 159 bp (-159) 내에 존재하지만 그 발현 정도는 매우 낮았다 (그림 8B). 그러나 추가로 49 bp (-208)를 더했을때 발현 정도는 매우 증가하는 것을 알 수 있었으며 257 bp 까지를 포함했을때 최대치를 보였다. 이 결과는 2가지 인식 부위의 존재를 의미하는데 세포 혹은 조직 특이성 발현을 결정하는 부위와 발현 정도를 양적으로 증가시키는 부위이다. 이 2가지 부위를

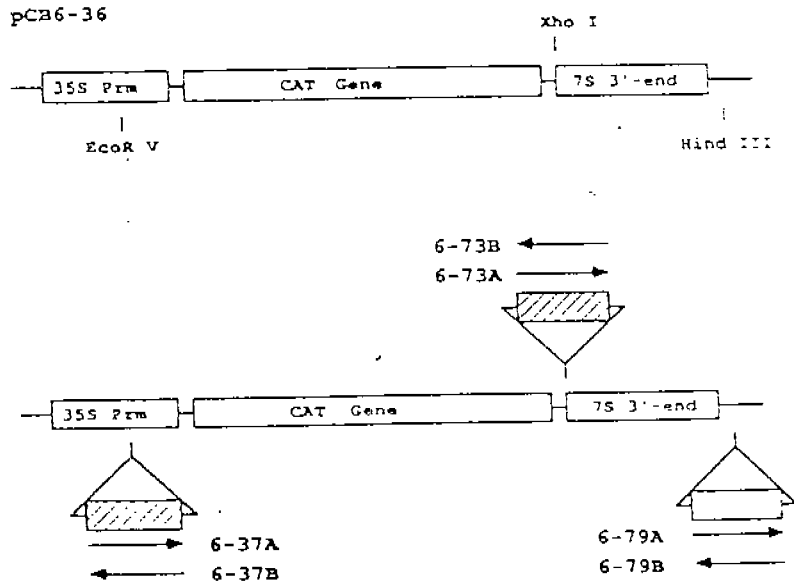
A

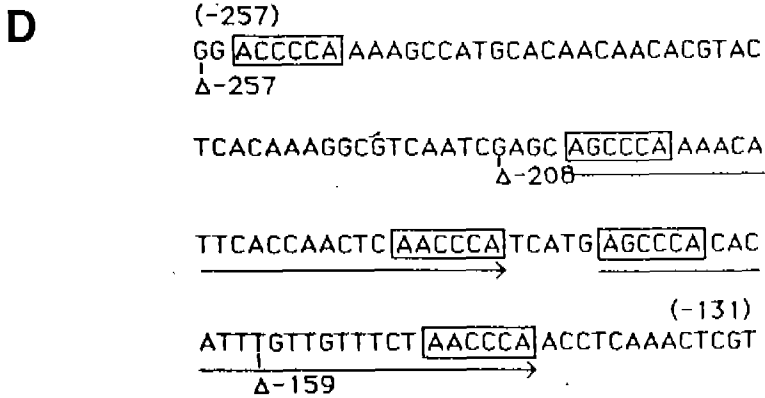


**B**



**C**





**Fig. 8.** Examination of transcriptional control activity of the 5' flanking region of  $\beta$ -conglycinin  $\alpha$ '-subunit gene in transgenic plant. A, construction of deletion mutants at 5' region; B, analysis of the level of expression of  $\beta$ -conglycinin gene in transgenic petunia; C, position effect analysis of the regulatory element of 5' upstream region in transgenic tobacco; D, nucleotide sequence in the 5' region of the  $\beta$ -conglycinin gene. Figures are taken from Chen *et al.* (1986, 1988).

포함하는 (-78 - -257) 부위를 CAT reporter 유전자와 재조합된 35S gene promoter 의 -90 위치에 삽입했을 때 orientation 에 상관없이 담배 씨의 발달 과정에서 예상된 시기에만 CAT 활성이 25 - 40 배 증가하였다 (그림 8C). 뿌리, 줄기, 잎 등 다른 조직에서는 CAT 활성의 변화가 없었다. 그러나 CAT coding sequence 의 끝에 삽입했을 때는 2 - 5 배의 활성 증가가 있었으나 poly(A) addition 위치의 3' 에 삽입했을 때는 별다른 증가가 없었다. 따라서 이 부위는 종자 발달 과정에서 유전자의 발현 조절에 특이하게 관여하는 인식 부위를 포함하는 것이 분명하다. 이 부위에는 CATGCAT sequence 와 CACA sequence 외에도 AVCCCA sequence 가 5번 반복하고 있어서 이 motif 가 작용하는 것으로 여겨진다 (그림 8D). 이 부위 (-78 - -257)는 orientation 에 상관없이 같은 효율로 작용하는 것으로 보아 조직 및 분화 발달 특이성 enhancer의 역할을 하는 것으로 추정된다. 따라서 이러한 cis-acting elements와 다른 trans-acting element가 상호작용하여 저장단백질의 조직특이성,

분화 발달 특이성 메카니즘을 나타내는 것으로 여겨진다 (Jofuku *et al.* 1986; Goldberg *et al.* 1989). Glycinin 유전자 Gy1 의 경우 -64 - +1 까지 65 bp 가 배 특이성 발현을 결정하는데 필요하며 -454 - -64 부위가 정량적 발현을 결정하는 것으로 알려졌다 (Goldberg *et al.* 1989; Nielsen *et al.* 1989). 여기에는 2개의 CATGCAT 부위 (-109 와 -253)가 있으며 1개의 CACA (-427) 부위가 있으며 이 CACA가 embryo nuclear factor 와 결합하는 것으로 알려졌다. Lectin 유전자의 경우 embryo nuclear factor는 분자량이 60K로 알려졌다 (Jofuku *et al.* 1987). 이와같이 두개의 다른 유전자의 발현을 조절하는 부위에 공통으로 존재하는 짧은 공통 염기서열 (consensus sequence)과 이들 인식 부위와 결합하는 특이한 단백질과의 상호 작용이 이들 종자 단백질의 발현을 transcription 수준에서 조절하는 중요한 역할을 할 것으로 여겨진다. 그러나 이러한 발달 분화 및 조직 특이성 발현 메카니즘이 enhancer 에 의해 종자에서 양성으로 조절되는 것과 함께 silencer 에 의해 다른 조직에서 음성으로 조절되는 경우도 배제할 수 없다. 완두콩의 cab AB80 gene 의 5' 247 bp (-347 - -100) 부위는 이 유전자가 잎에서 빛에 의해 발현이 유도되는 과정에서 enhancer 와 같은 양성 인자로 작용하지만 뿌리에서는 이러한 발현을 억제하는 silencer 의 기능도 갖고 있는 것이 보여졌다 (Simpson *et al.* 1986). 이와같이 transgenic plant 는 조직 특이성 및 분화 발달 특이성 유전자 발현을 조절하는 인자를 파악하는데 매우 유용하게 활용될 수 있다.

## 결 론

대두 종자의 주된 단백질은 glycinin과  $\beta$ -conglycinin이 있다. 이러한 대두의 저장

단백질 유전자는 종자의 숙성 과정에서 제한된 숫자의 조직 특이성 유전자가 한정된 시간에 매우 높은 수준으로 발현되다가 숙성 말기에는 다시 억제 상태가 된다. 이러한 왕성한 유전자의 활성화는 유전자의 발현 메카니즘을 분자 수준에서 이해하는데 매우 큰 도움을 준다. 이러한 대두 저장 단백질 유전자의 발현 조절은 우선 전사 단계에서 조절이 된다. 그러나 배 형성 과정에서 전사 활성화가 이루어진 유전자의 발현은 전사 후 mRNA로의 전환 과정에서 1차 조절이 이루어지는 것으로 여겨지며 mRNA가 전이 과정을 거쳐 단백질로 발현되는 과정에서 전이 후 변형에 의해 다시 한 번 조절이 이루어지는 것을 알 수 있다. 종자의 숙성과 더불어 발현이 억제되는 메카니즘도 역시 전사 단계에서 주로 조절이 이루어지는 것으로 여겨진다. 이러한 전사 조절은 배 세포 핵 내의 단백질 인자와 유전자의 발현 조절 부위가 상호 작용하여 이루어지는 것으로 알려져 있다. 이러한 유전자의 발현 조절 부위를 세포 내에서 검색하고 검정하는 과정은 transgenic plant의 생산으로 가능해졌다. 따라서 이러한 발현 조절 부위와 상호 작용하여 유전자의 발현을 활성화시키고 억제시키는 메카니즘은 일반적인 식물 유전자의 발현 조절을 이해하는데 도움을 줄 것이다.

## 참 고 문 헌

1. Barker, S.J., Harada, J.J. and R.B. Goldberg (1988). Proc. Natl. Acad. Sci. USA. 85: 458.
2. Barton, K.A., J.F. Thompson, J.T. Madison, R. Rosenthal, N.P. Jarvis and R.N. Beachy (1982). J. Biol. Chem. 257: 6089.
3. Baumlein, H., U. Wobus, J. Pustell, and F.C. Kafato (1986) Nucl. Acid Res. 14: 2707.
4. Benfey, P.N. and N.H. Chua (1989). Science 244: 174.
5. Bray, E.A. and R.N. Beachy (1985). Plant Physiol. 79: 746.

6. Chen,Z.L., M.A.Schuler, and R.N.Beachy (1986). *Proc.Natl.Acad. Sci.USA.* 83: 8560.
7. Chen,Z.L., N.S.Pan and R.N.Beachy (1988). *EMBO J.* 7: 297.
8. Chrispeels,M.J., T.J.V.Higgins and D.Spencer (1982). *J.Cell Biol.* 93: 306.
9. Croy,R.R.D., G.W.Lycett, J.A.Gatehouse, J.N.Yarwood and D.Boulter (1982). *Nature* 295: 76.
10. Dickinson,C.D., L.A.Floener, G.G.Lilley and N.C.Nielsen (1987). *Proc.Natl.Acad.Sci.USA* 84: 5525.
11. Dickinson,C.D., R.P.Evans and N.C.Nielsen (1988). *Nucl.Acid Res.* 16: 371.
12. Dickinson,C.D., E.H.A.Hussein and N.C.Nielsen (1989). *Plant Cell* 1: 459.
13. Doyle,J.J., M.A.Schuler, W.D.Godette, V.Zenger, R.N.Beachy and J.L.Slighton (1986). *J.Biol.Chem.* 261: 9228.
14. Ereken-Tumer,N., J.D.Richter and N.C.Nielsen (1982). *J.Biol.Chem* 257: 4016.
15. Fukazawa,C., T.Momma, H.Hirano, K.Harada and K.Udaka (1985). *J.Biol.Chem.* 260: 6234.
16. Gatehouse,J.A., I.M.Evans, R.R.D.Croy and D.Boulter (1986). *Philos.Trans.Roy.Soc.(Lond.)* B314: 367.
17. Goldgerg,R.B., G.Hoschek, S.H.Tam, G.S.Ditta and R.W.Breidenbach (1981a). *Dev.Biol.* 83: 201.
18. Goldberg,R.B., G.Hoschek, G.S.Ditta and R.W.Breidenbach (1981b). *Dev.Biol.* 83: 218.
19. Goldberg,R.B. (1986). *Phil.Trans.Roy.Soc.(Lond.)* B314: 343.
20. Goldberg,R.B. (1988). *Science* 240: 1460.
21. Goldberg,R.B., S.J.Barker and L.Perez-Grau (1989). *Cell* 56: 149.
22. Harada,J.J., S.J.Barker and R.B.Goldberg (1989). *Plant Cell* 1: 415.
23. Heidecker,G. and J.Messing (1986). *Ann.Rev.Plant Physiol.* 37: 439
24. Hill,J.E. and R.W.Breidenbach (1974). *Plant Physiol.* 53: 742.
25. Hirano,H., C.Fukasawa and K.Harada (1984). *J.Biol.Chem.* 259: 14371.
26. Jensen,J.S., K.A.Marcker, L.Otten and J.Schell (1986). *Nature* 321: 669.
27. Joergensen, J.-E., J. Stougaard, A. Marcker and K. A. Marcker (1988). *Nucl.Acid Res.*16: 39.
28. Jofuku,K.D., J.K.Okamuro and R.B.Goldberg (1987). *Nature* 328: 734
29. Jofuku,K.D., R.D.Schipper and R.B.Goldberg (1989). *Plant Cell* 1: 427.
30. Kim,C.H. and Y.D.Choi (1989a). *Kor.J.Botany* 32: 51.
31. Kim,C.H. and Y.D.Choi (1989b). *Kor.Biochem.J.*, in press.

32. Kim, C.H. and Y.D. Choi (1989c). manuscript, in preparation.
33. Lutcke, H.A., K.C. Chow, F.S. Mickel, K.A. Moss, H.F. Kern, and G.A. Scheele (1987). *EMBO J.* 6: 43.
34. Marco, Y.A., V.H. Thanh, N.E. Tumer, B.J. Scallan and N.C. Nielsen (1984). *J. Biol. Chem.* 259: 13436.
35. 35. Messing, J., D. Geraghty, G. Heidecker, N.-T. Hu, J. Kridl and I. Rubenstein (1983). In *Genetic Engineering of Plants*, eds.
36. T. Kosuge Yet al. Q New York, p211.
37. Momma, T., T. Negoro, H. Hirano, A. Matsumoto, K. Utaka and C. Fukasawa (1985). *Eur. J. Biochem.* 149: 491.
38. Moreira, M.A., M.A. Hermodson, B.A. Larkins and N.C. Nielsen (1981). *Arch. Biochem. Biophys.* 210: 633.
39. Morell, M. and L. Copeland (1985). *Plant Physiol.* 78: 149.
40. Negoro, T., T. Momma and C. Fukasawa (1985). *Nucl. Acids Res.* 13: 6719.
41. Nguyen, T., M. Zelechowska, V. Foster, H. Bergmann and D.P.S. Verma (1985). *Proc. Natl. Acad. Sci.* 82: 5040.
42. Nielsen, N.C. (1984). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 304: 287
43. Nielsen, N.C. (1985). *J. Am. Oil. Chem. Soc.* 62: 1680.
44. Nielsen, N.C., C.D. Dickenson, T.J. Cho, V.H. Thanh, B.J. Scallan, T.L. Sims, R.L. Fischer and R.B. Goldberg (1989). *Plant Cell* 1: 313.
45. Okamuro, J.K., K.D. Jofuku and R.B. Goldberg (1986). *Proc. Natl. Acad. Sci.* 83: 8240.
46. Pernollet, J.-C. and J. Mosse (1983). In *Seed Proteins*, eds. J. Daussant Yet al. Q press, New York, p155.
47. Proudfoot, N.T. (1979). *Nature* 279: 376.
48. Pusztai, A. and J.C. Stewart (1980). *Biochim. Biophys. Acta.* 623: 418.
49. Schuler, M.A., B.F. Ladin, J.C. Pollaco, G. Freyer and R.N. Beachy (1982). *Nucl. Acids Res.* 101: 8245.
50. Shirley, B.W. (1987). *Nucl. Acids Res.* 15: 6501.
51. Simpson, J., J. Schell, M.V. Montagu and L. Herrera-Estrella (1986). *Nature* 323: 551.
52. Staswick, P.E., M.A. Hermodson and N.C. Nielsen (1984). *J. Biol. Chem.* 259: 13424.



53. Tumer, N.E., V.H.Thanh and N.C.Nielsen (1981). J.Biol.Chem. 256: 8756.
54. Walling, L., G.M.Drews and R.B.Goldberg (1986). Proc.Natl.Acad. Sci.USA 83: 2123.
55. Weiher, H., M.Konig and P.Gruss (1983). Science 219: 626.

## 저 자 약 력

### 최 양 도 (崔良燾) 박사

- 1953. 3. 15. 생
- 1976 서울대학교 농화학과 (농학사)
- 1978 한국과학원 생물공학과 (이학석사)
- 1985 미국 Northwestern대학교 분자세포생물학과 (이학박사)
- 1985 - 86 미국 Northwestern대학교 분자 세포생물학과 박사후과정
- 1986 - 현재 서울대학교 농화학과 교수