Title	Regulation of Transcription in Animal Cells: A 50-year Journey Revealing an Expanding Universe of Factors and Mechanisms
Author(s)	Robert G. Roeder
Language	English
Event title	The 2021 Kyoto Prize Commemorative Lecture
Publisher	Inamori Foundation
Issue Date	10/01/2022
Start page	1
End page	12
URL	https://www.kyotoprize.org/wp-content/uploads/2022/10/2021_roeder_en.pdf

URL for Japanese translation: https://www.kyotoprize.org/wp-content/uploads/2022/10/2021_roeder_jp.pdf

The 2021 Kyoto Prize Commemorative Lecture

Robert G. Roeder

Regulation of Transcription in Animal Cells: A 50-year Journey Revealing an Expanding Universe of Factors and Mechanisms

I am both honored and humbled to receive the 2021 Kyoto Prize in Basic Sciences. It is especially pleasing in view of the ideals and philosophy of the Inamori Foundation, and I graciously thank all those associated with the Foundation. I will now describe several major discoveries over 50 years of biological research, as well as some personal background.

It is often stated that we are our genes—but perhaps more accurately, we are the products of our genes (Fig. 1). The central dogma states that DNA is transcribed into RNA copies that are then translated into proteins with various enzymatic, structural, and regulatory properties. Remarkably, embryonic stem cells can give rise to different cell types with the same set of genes (Fig. 2). Related, the normal formation and function of different cell types and many associated pathologies result from differential gene expression, which is controlled primarily at the level of transcription, the first step in gene expression. This makes it critical to understand the mechanisms that regulate transcription, and this has been my major objective, and passion, for over 50 years. Transcription is carried out by an enzyme called RNA polymerase, and Fig. 3 shows transcription in prokaryotes as a frame of reference for the eukaryotic transcription that I will describe. In essence, the DNA is transcribed by a single RNA polymerase, and is regulated by interacting gene-specific activators. As is true in eukaryotes as well, transcription begins at a specific site and makes single-stranded RNA copies of the template strand of the DNA.

As a preview, my major discoveries include eukaryotic RNA polymerases I, II, III, their distinct structures and functions, cognate general initiation factors, the prototype gene-specific transcriptional activator, general and gene-specific coactivators, causal roles for chromosomal histone modifications in transcription, and biochemically defined cell-free systems that accurately transcribe specific genes. Before discussing these in more detail, a little personal background.

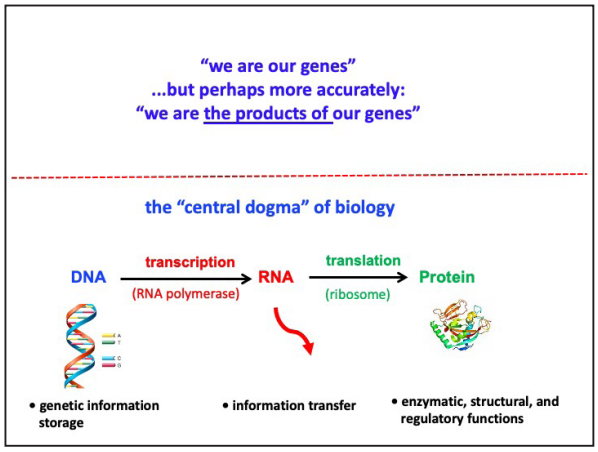

Fig. 1

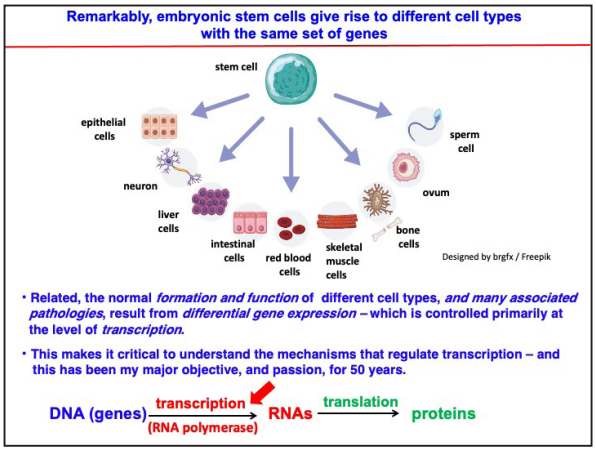

Fig. 2

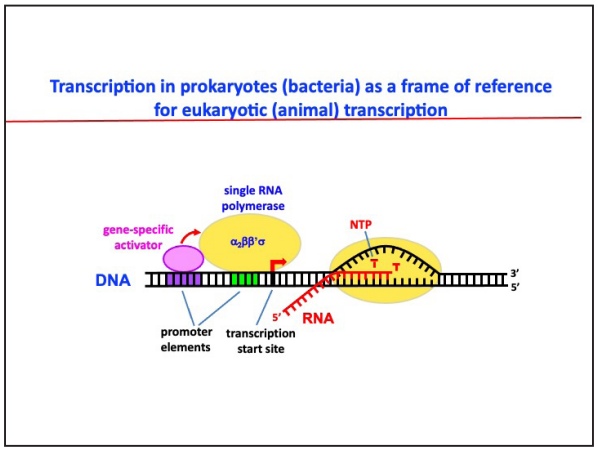

Fig. 3

Fig. 4

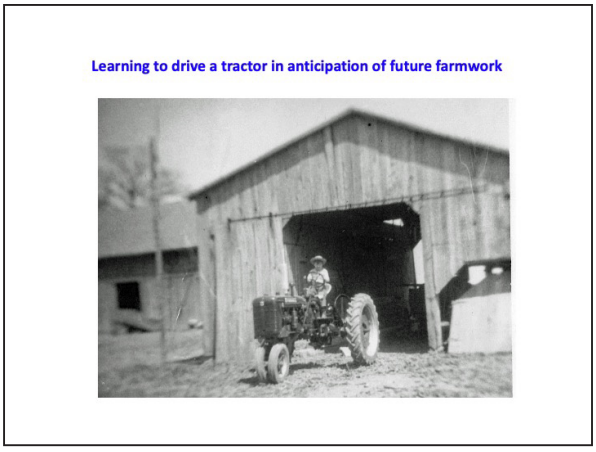

Fig. 5

Personal Background

I was born and raised on a farm in Southern Indiana. My parents had very limited formal educations but were caring parents who taught their four children to be honest, humble, and diligent. And although diligence in school studies was expected, education beyond high school was not considered, and I was expected to remain on the family farm. Fig. 4 shows a 1946 family photo on a Sunday afternoon outing. Note my older brothers and I, in the lower left, in overalls typically worn by farm boys. Fig. 5 shows me learning to drive a tractor in anticipation of future farm work. As a child, I had little exposure to science per se, but enjoyed trying to figure out how small mechanical devices worked, as well as building various small devices, such as crystal radios. In high school, I was especially interested in mathematics and chemistry, but had little time for extracurricular activities because of extensive farm chores. Fortunately, with a strong academic background, I received a full tuition scholarship to attend Wabash College, a small liberal arts college with strong science departments. Although I had a primary interest in chemistry, I also became intrigued by biochemistry and the emerging molecular biology during a course taught by a new assistant professor, Tom Cole, from Caltech. I was especially influenced by the classic 1961 Jacob-Monod paper on gene regulation in bacteria, leading me to think about future studies on gene regulation in animal cells. Those interests led me to a graduate program in biochemistry at the University of Illinois. At Illinois, I joined the laboratory of Bill Rutter, an inspiring mentor working on aldolase enzymes and pancreas development, albeit not transcription. Fortunately, he allowed me to initiate studies on transcription, which began after the laboratory moved to the

University of Washington in Seattle in 1965. At that time, virtually nothing was known about gene regulation in animal cells, except that, as in bacteria, there were three major classes of RNA: messenger, ribosomal, and transfer. My initial studies focused on quantitative measurements of RNA synthesis in isolated nuclei and in cells during hormonal responses in rat liver and during sea urchin development. However, it was not yet possible to monitor specific gene products in those pre-gene cloning days. Therefore, I decided to go to the heart of the transcription problem, and to first identify the enzyme that transcribes DNA. As detailed later, this led to my discovery of RNA polymerases I, II, III (Pols I, II, III). This was a true *Eureka!* moment in my career—and also guaranteed a noteworthy thesis. For postdoctoral studies, I joined the laboratory of Don Brown, another inspiring mentor, who had purified the large ribosomal RNA genes that I suspected were transcribed by Pol I because of colocalization in the nucleolus. Surprisingly, I failed to see specific transcription of these genes by purified Pol I, which led me to suspect that eukaryotic transcription would be more complicated than imagined and set the stage for subsequent studies in my own laboratory at Washington University.

Identification of the Eukaryotic Nuclear RNA Polymerases

Returning to my graduate work, and regarding the identification of the eukaryotic nuclear RNA polymerases, in 1959–1960, Sam Weiss showed NTP-dependent RNA synthesis both in isolated nuclei and in a derived chromatin aggregate. In 1964–1969, several labs reported only a single chromatographic peak of RNA polymerase activity, suggestive of a single enzyme, but had employed low salt, low yield extraction procedures. In 1968–1969, in the Rutter laboratory, I realized that, unlike the situation in bacteria, most RNA polymerase in eukaryotic cells was chromatin-bound, that is engaged in transcription, and systematically developed new extraction and purification procedures that included: high salt/sonication to dissociate the histone-DNA and polymerase-DNA-RNA complexes; dialysis to low salt to precipitate DNA-histone, leaving quantitatively solubilized RNA polymerase; and finally, ion exchange chromatography to resolve Pols I, II, III. Fig. 6 shows my laboratory notebooks and my thesis with volume 12 containing the original description and characterization of Pols I, II and III. Fig. 7 shows me collecting sea urchins, the organism in which Pols I, II and III were first discovered. Fig. 8 shows a chromatographic resolution of the three nuclear RNA polymerases in February of 1969, with the red lines showing the three different RNA polymerases that were chromatographically eluted. This work resulted in my first publication, a *Nature* article (Fig. 9). But of note, the paper was submitted to *Nature* on August 5, 1969 and initially rejected on the grounds that it was "not of general interest"—not a particularly happy moment for a graduate student. But happily, it was published on October 18 of

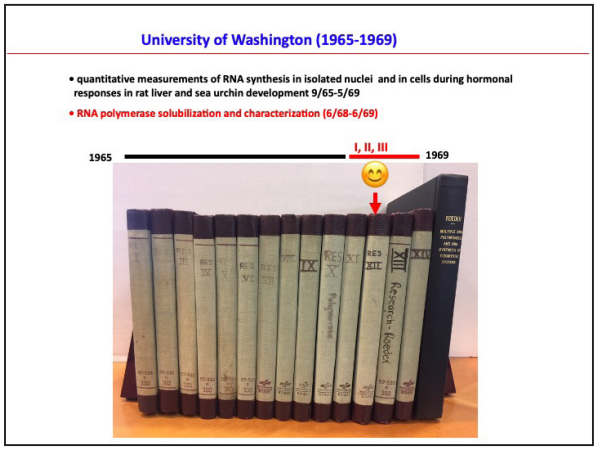

Fig. 6

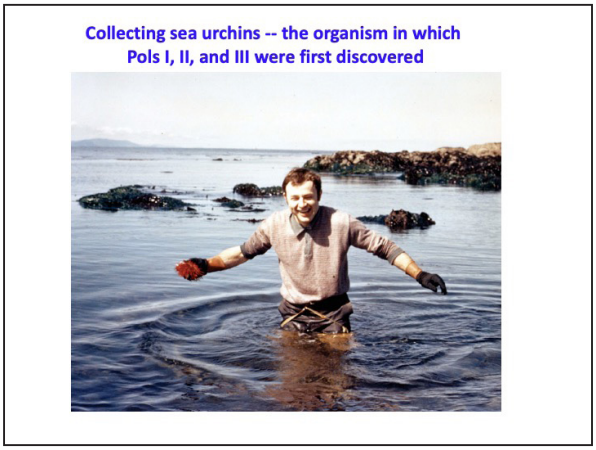

Fig. 7

that year as originally submitted. The identification of the distinct RNA polymerases in 1969 was foundational, but as it turns out, and as you will see, just the tip of the iceberg (Fig. 10)—as discussed in my 2003 Lasker Award Commentary entitled “The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen”.

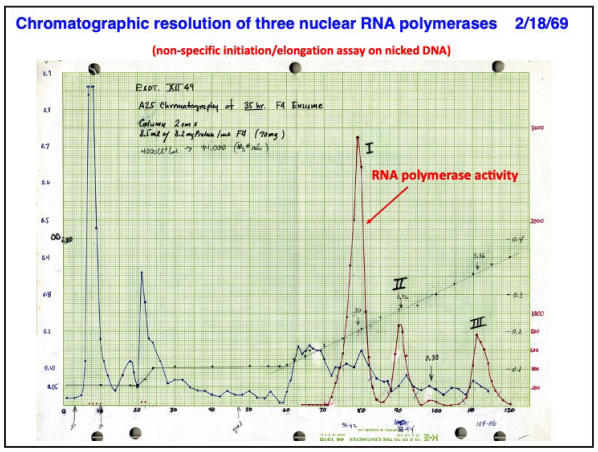

Fig. 8

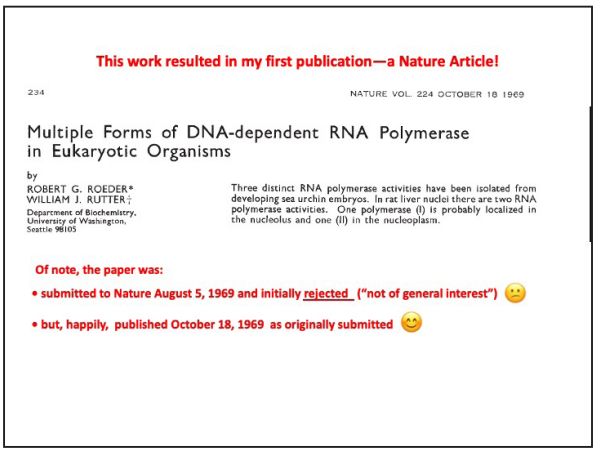

Fig. 9

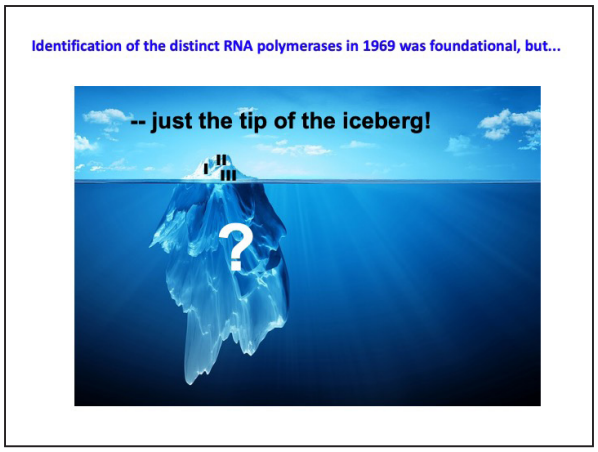

Fig. 10

Identification of Specific RNA Polymerase Functions

Having three enzymes, the next task was to identify their specific functions. The early functional assignments were based on comparisons of the differential sensitivities of the purified polymerases to the mushroom toxin α-amanitin with the sensitivities of the synthesis of specific RNAs by endogenous, engaged, polymerases in isolated nuclei (Fig. 11). As you can see in the panel on the right, the sensitivity of messenger RNA synthesis matched that of Pol II, the sensitivity of 5S and tRNA synthesis matched that of Pol III, and the insensitivity of ribosomal RNA synthesis matched that of Pol I. So, these studies led us to the following situation, namely that RNA polymerases I, II and III transcribe, respectively, the genes encoding ribosomal RNA, messenger RNA, and 5S and tRNA (Fig. 12)—and these RNAs converge on the ribosome for protein synthesis. This scenario is distinct from that in prokaryotes, which have one enzyme for all classes of RNA, and it provides a convenient means for independent regulation of the global synthesis of the major classes of RNA—for example, during growth state changes.

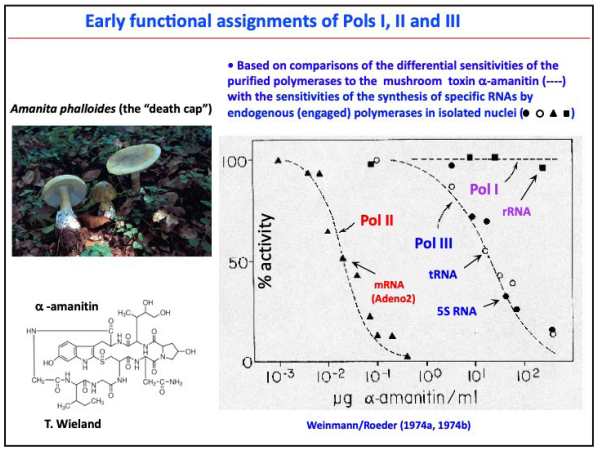

Fig. 11

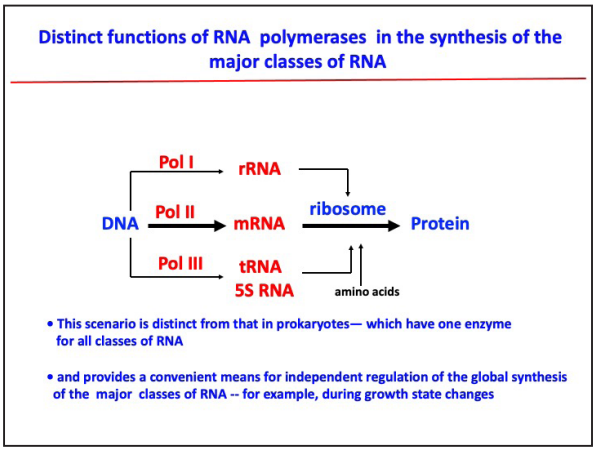

Fig. 12

Assessment of the Structural Basis for the Distinct RNA Polymerase Functions

Given these important findings, the next question was to assess the structural basis for the distinct RNA polymerase functions. By 1974, my students and I had purified to homogeneity the RNA polymerases from mouse tumor cells and resolved their subunits by electrophoresis—as shown in the left panel of Fig. 13. As you can see, the three enzymes have complex and distinct subunit structures. Twenty years later, the laboratories of Sentenac and Young had purified the homologous yeast enzymes and cloned the corresponding subunits (Fig. 13, right panel). Those studies revealed that some subunits were common to the three RNA polymerases, some were completely distinct, and others were highly related to each other and to subunits in the bacterial RNA polymerase. So, these results revealed a molecular basis for some common enzymatic properties, as well as the distinct specificities and regulation of the three enzymes.

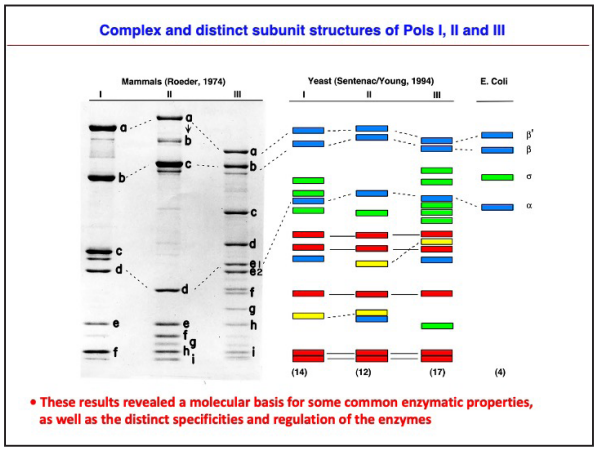

Fig. 13

Establishment of Accurate Transcription of Specific Genes by Purified RNA Polymerases for Mechanistic Analyses

Given distinct enzyme structures and functions, the next task was to establish accurate transcription of specific genes by purified RNA polymerases for mechanistic analyses. These studies resulted in the identification of general initiation factors for specific transcription (Fig. 14). Prior to 1979, incubation of purified RNA polymerases I, II, III with corresponding purified genes failed to yield any specific transcription. But in 1979, we showed that incubation of purified RNA

polymerase II or III with corresponding purified genes in the presence of a cellular extract resulted in specific transcription. In 1980, the subsequent chromatographic fractionation of this extract yielded two general initiation factors, TFIIIB and TFIIIC, as well as a gene-specific factor, TFIIIA, for RNA polymerase III. In the case of RNA polymerase II, the chromatographic fractionation resulted in the identification of multiple (partially purified) factors that were later resolved into six independent factors called TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. By 1992, studies in my own and other labs (indicated in Fig. 14) had resulted in the complete purification of these factors and in the cloning and validation of the individual subunits. Fig. 15 shows some of the original data for accurate transcription initiation by purified Pol II. Incubation of the indicated DNA fragment containing the adenovirus major late promoter with Pol II and with the extract led to the synthesis of a specific 536 -nucleotide RNA transcript that indicated specific initiation at the natural initiation site. This RNA was not seen when RNA polymerase II was inhibited with α-amanitin or when the extract was omitted from the reaction. In later studies with purified factors (Fig. 14, lower panel), robust transcription of the same promoter was seen when all factors were present, but not with the omission of any single factor or RNA polymerase II, clearly establishing their requirements for promoter-specific transcription initiation. An important point from these studies was that our results showed accurate, but promiscuous, transcription of the cell-specific Adenovirus2 and β -globin promoters by the ubiquitous Pol II and general initiation factors. This led to the prediction, and later discovery, of a general repression mechanism, namely DNA assembly into chromatin, and gene- and cell-specific transcriptional regulatory factors, activators, that reverse the repression.

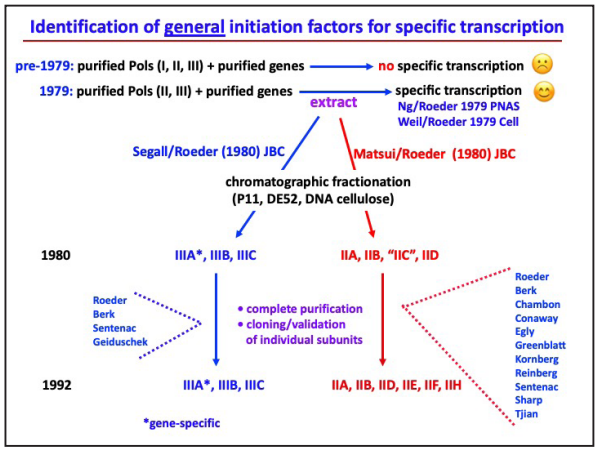

Fig. 14

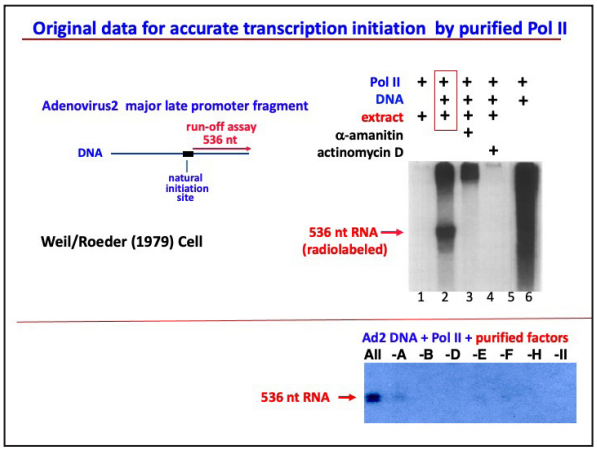

Fig. 15

Elucidation of the Mechanism Involved in Specific Transcription Initiation

Having identified general initiation factors, the next task was to establish the mechanism involved in specific transcription initiation by the polymerase and these factors. Fig. 16 (left panel) summarizes our initial results on a Pol III-transcribed tRNA gene. Promoter recognition is established by TFIIIC, and TFIIIC in turn results in the stepwise recruitment of TFIIIB and Pol III, resulting in the formation of a preinitiation complex containing about 25 polypeptides. Similar results were obtained for a Pol II-transcribed promoter (Fig. 16, right panel). Our early studies showed an initial recognition of the promoter by the initiation factor TFIID. Subsequent studies by my own and other laboratories (indicated in the figure) showed the stepwise assembly of the remaining factors in the formation of a preinitiation complex containing 44 polypeptides. Incubation of either of these preinitiation complexes with nucleoside triphosphates resulted in specific transcription initiation and subsequent transcription elongation.

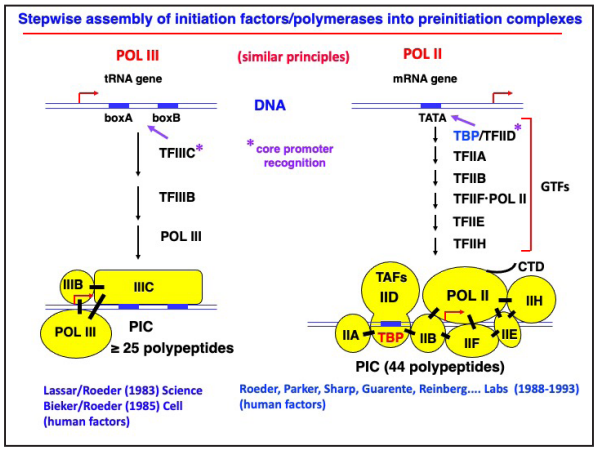

Fig. 16

Discovery of Gene/cell-specific Transcriptional Activators

The next seminal event in this journey was the discovery of gene- and cell-specific transcriptional activators. These factors were predicted based on precedent from bacterial studies and on the promiscuity of the general transcription machinery, necessitating some mechanism to achieve gene- and cell-specific transcription. The first of these factors to be identified was the 5S gene-specific TFIIIA, which was purified in 1980 on the basis of a functional transcription assay and shown to bind and activate the 5S RNA gene (Fig. 17). TFIIIA was the prototype DNA-binding transcriptional activator in eukaryotes. The left panel in Fig. 17 shows the transcription assay with purified Pol III, TFIIIB, and TFIIIC, which result in the robust transcription of the control tRNA gene, but not the 5S RNA gene. Robust transcription of the 5S gene is achieved in the presence of TFIIIA and these same factors as further shown in the left panel. The center panel simply shows the electrophoretically resolved, highly purified TFIIIA, and the panel on the right shows promoter binding of TFIIIA by a DNase footprint assay. These studies in 1980 were followed by our subsequent cDNA cloning of TFIIIA, providing the first protein sequence of a transcription initiation factor and leading Aaron Klug to deduce the zinc finger motif, which is the most common DNA binding motif in eukaryotic transcription factors. Mechanistically, we showed that TFIIIA binding to the promoter facilitates the subsequent recruitment of TFIIIC, which otherwise does not bind to this particular promoter, and that TFIIIC, in turn, facilitates the stepwise recruitment of TFIIIB and Pol III, as described for the tRNA gene (Fig. 18). This represents the first defined mechanism of action for any gene-specific transcriptional activator in eukaryotes, and is distinct from the prokaryotic mechanism in which activators directly bind the RNA polymerase (Fig. 3).

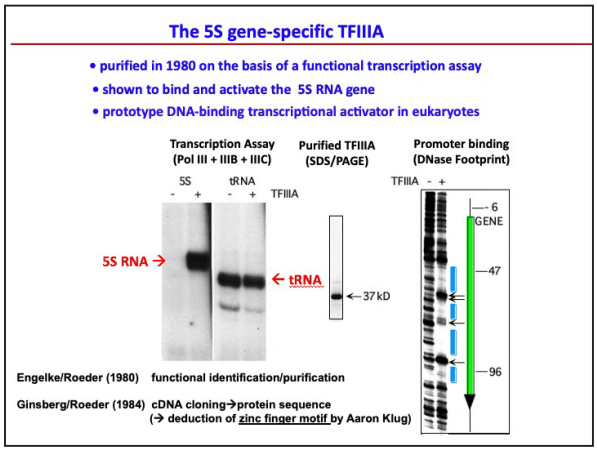

Fig. 17

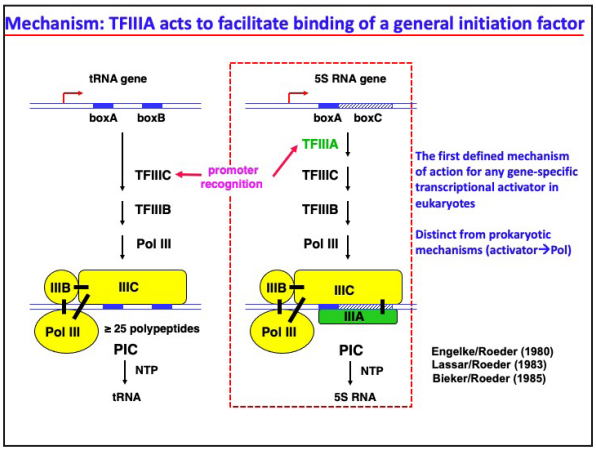

Fig. 18

Subsequent studies in my own and other laboratories in the next four years led to the identification of four gene-specific activators for Pol II-transcribed genes (Fig. 19), and currently we know there are about 1,600 of these gene-specific, cell-specific factors. They are typified by a DNA binding domain and a so-called activation domain. Many of these factors are master transcriptional regulators of cell fate and differentiation, as summarized in Fig. 20. Weintraub, in 1987, showed that MyoD could convert a fibroblast to a muscle cell. Yamanaka, in 2006, showed that a fibroblast could be converted to a pluripotent stem cell by the ectopic expression of only four other factors. These remarkable studies were recognized by the Kyoto Prize, the Nobel Prize, and others. These studies and other studies with master transcriptional regulators, some of which are shown in Fig. 20, emphasize both the physiological significance and the power of these transcription factors—their ability to change cell fate.

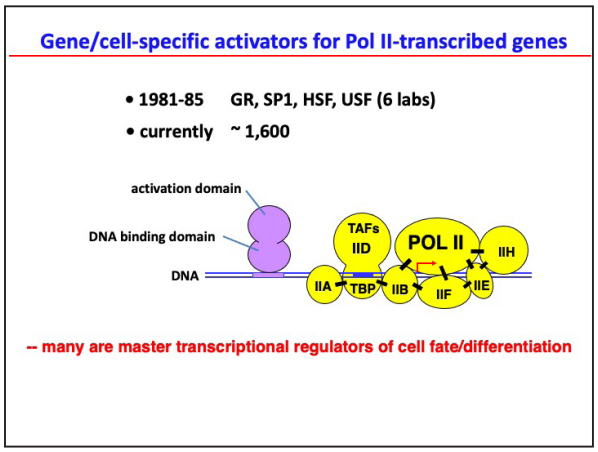

Fig. 19

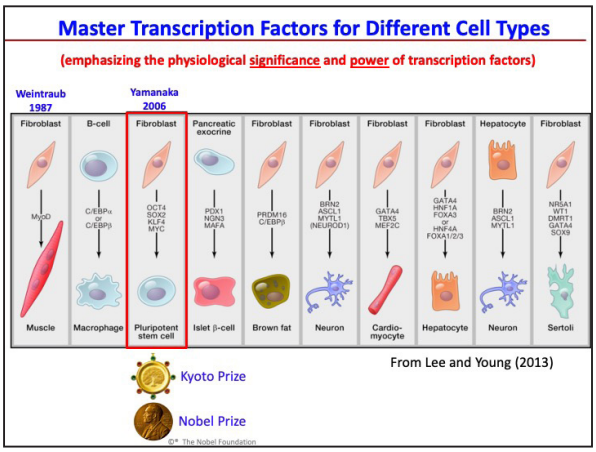

Fig. 20

Establishment of the Mechanism of Action of Gene/cell-specific Activators

Given their extreme biological significance, the next task was to establish the mechanism of action of gene- and cell-specific activators for Pol II-transcribed genes. The basic question here was once an activator is bound to the promoter or enhancer site, how does it effect the formation and function of the preinitiation complex on the target gene (Fig. 21)? Quite surprisingly, given the structural complexity of these entities, the activators failed to function with highly purified Pol II and initiation factors. Functional biochemical assays then identified essential coactivators in the early 1990s in several laboratories, some of which are summarized in Fig. 22. This introduces a group of cofactors that operate directly on the general transcription machinery. These cofactors include the TAF subunits of TFIID, which were initially described in Drosophila by the Tjian lab and in human cells by the Roeder and Berk labs. The 30-subunit Mediator complex, probably the most important coactivator complex, was initially described in yeast based on genetic and biochemical assays by the Young and Kornberg labs, and shown to interact directly with RNA polymerase II. It was first described in human cells by our laboratory on the basis of biochemical assays, which also showed a direct binding to transcriptional activators. Both of these groups of coactivators, the TAFs and the Mediator, are generally required for activator function. We also identified OCA-B, a B cell-specific coactivator that is selective for the OCT 1 - and OCT 2 -bound genes in B cells and the first representative of an expanding class of cell- and gene-specific coactivators. Fig. 23 shows more detail on the Mediator mechanism, which basically acts as a bridge between diverse enhancer-bound activators and the basal transcription machinery. As we initially showed, liganded nuclear hormone receptors like TR and PPARγ interact with the Mediator through the MED 1 subunit, according to the detailed interactions shown in the figure.

This direct interaction serves to recruit the Mediator to activator-bound enhancers for subsequent interactions with the general transcription machinery at the promoters. Again, this model was based on biochemical assays. It was validated by a MED 1 knockout in a mouse embryo fibroblast-based model of PPARγ-dependent adipogenesis, as shown in Fig. 24. As shown originally in the Spiegelman lab, fibroblasts can be differentiated to adipocytes, fat cells, with inducing factors and the master regulator PPARγ; and those cells, differentiated adipocytes with lipid droplets stained by Oil Red O, are shown in the left panel. As shown in the right panel, the MED 1-deficient MEFs fail to differentiate into adipocytes under these same conditions, and they further show impaired PPARγ target gene expression. So, this analysis provided a validation of the function and physiological relevance of the Mediator, and specifically the MED 1 subunit.

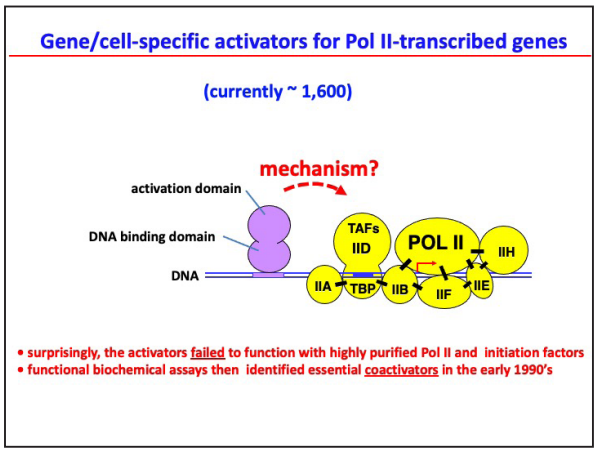

Fig. 21

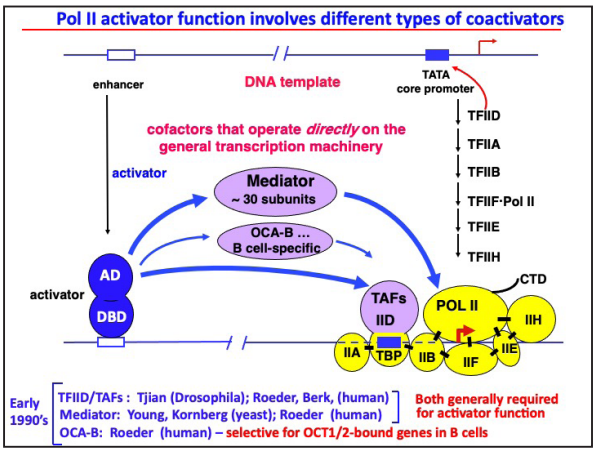

Fig. 22

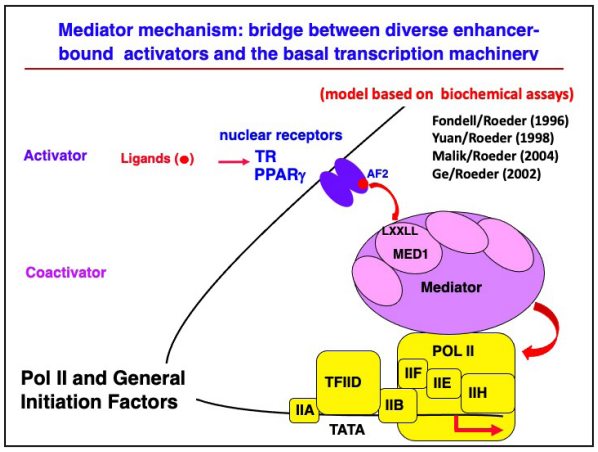

Fig. 23

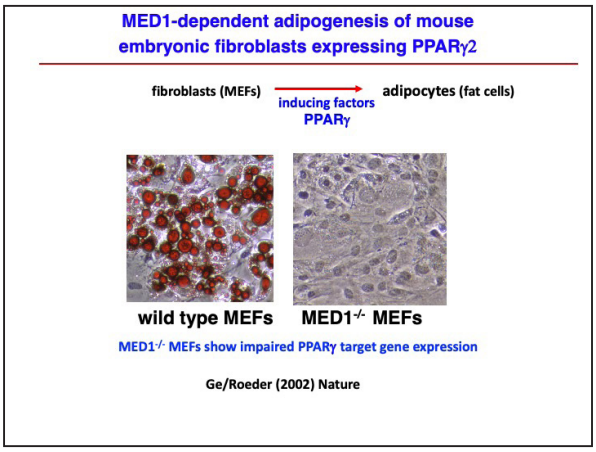

Fig. 24

Transcriptional Regulation in a Chromatin Context

Given its natural intracellular locale, we next analyzed transcriptional regulation in the context of chromatin. As is well known, genomic DNA interacts with core histones to form nucleosomal structures that both package DNA and, as you will see, serve to repress transcription (Fig. 25). Remarkably, many laboratories showed that the histone tails, on a nucleosome, can be acetylated, methylated, phosphorylated, and ubiquitylated at very specific sites by a diverse group of enzymes. This brings us to a group of cofactors that act through modifications of the chromatin template (Fig. 26). These cofactors include ATP-dependent chromatin remodeling factors and histone modifying factors that generate the indicated modifications. These factors were discovered by

other laboratories, but we were basically interested in establishing a defined cell-free system to investigate the integrated functions of these cofactors for both biochemical and mechanistic studies. The general strategy is shown in Fig. 27. As already described, incubation of a DNA template with Pol II and initiation factors results in robust—albeit promiscuous—accurate transcription. My laboratory and the laboratories of Kornberg and Luse showed that the assembly of this DNA into a chromatin structure with histones resulted in the repression of this promiscuous transcription activity. So, basically, this established the general repression mechanism that had been predicted. These templates and our assays were then utilized in conjunction with Pol II, initiation factors, various activators, and various cofactors to establish mechanisms involved in the reactivation of the native, repressed chromatin template. More detail for this protocol is shown in Fig. 28. A DNA template with activator sites and a core promoter is assembled into chromatin with bacterially expressed and purified histones according to the procedure of Jim Kadonaga. This chromatin, which resembles a "beads-on-a-string" structure, is then incubated with activators and with acetyl and methyl transferases to effect chromatin modifications. Those modified templates are then incubated either with nuclear extract as a source of all of the other general transcription factors or, alternatively, with purified Pol II, initiation factors, coactivators, and elongation factors to provide a quite defined in vitro transcription system, with a chromatin template, that contains in excess of a hundred different polypeptides. Notably, the chromatin can be assembled with wild-type histones, with mutant histones that cannot be modified, or with pre-modified histones. This system has allowed us to establish causal effects of histone modifications as well as direct effects and mechanisms of action of various coactivators. Some of these results are summarized in Fig. 29.

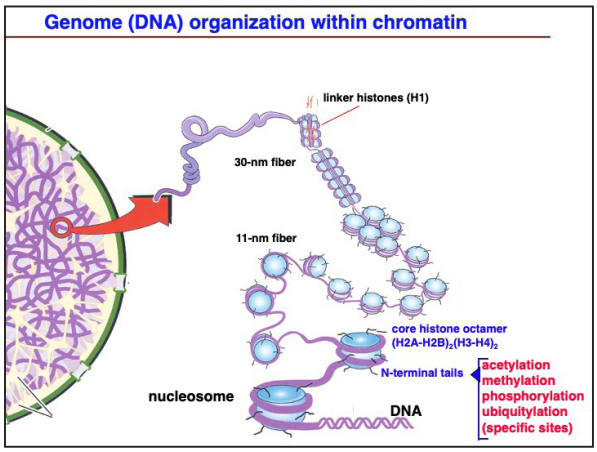

Fig. 25

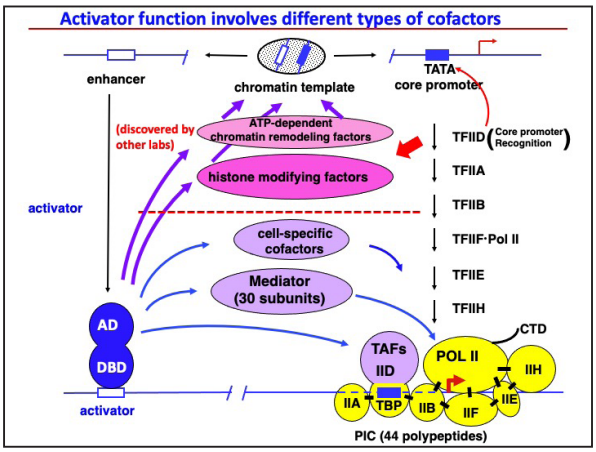

Fig. 26

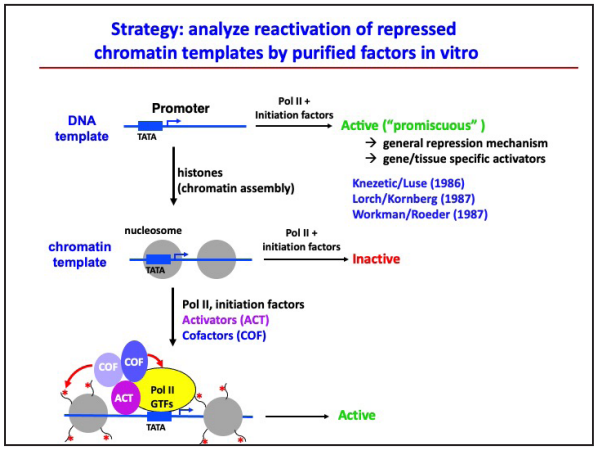

Fig. 27

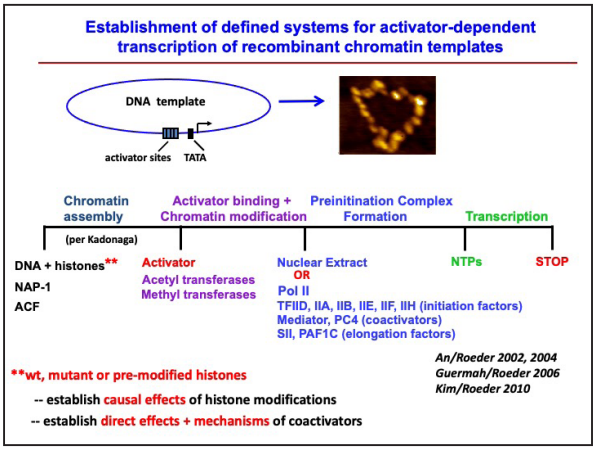

Fig. 28

In early studies, we analyzed the function of p300, a histone H 3/H 4 acetyltransferase, in conjunction with two different arginine methyltransferases that effect specific modifications. We later analyzed the function of p300 in conjunction with the SET1 complex, a major lysine histone methyltransferase that modifies lysine 4 in the promoter regions. In general summary of these studies (Fig. 29), they showed ordered, cooperative interactions and functions of coactivators in activator-dependent transcription along with specific histone modifications. Perhaps most importantly, they established causal effects of histone modifications on transcription. This was indicated by the observation that mutations in the modified histone acetylation and methylation sites eliminated the cofactor functions. This observation was critical, since we, in 1997, and later others, showed that histone modifying factors can also functionally modify many transcription factors. And the more common cell-based and genetic assays show only correlations of histone modifications with transcription and do not identify the essential substrates. So, again, the studies were important for establishing the anticipated causal effects of histone modifications.

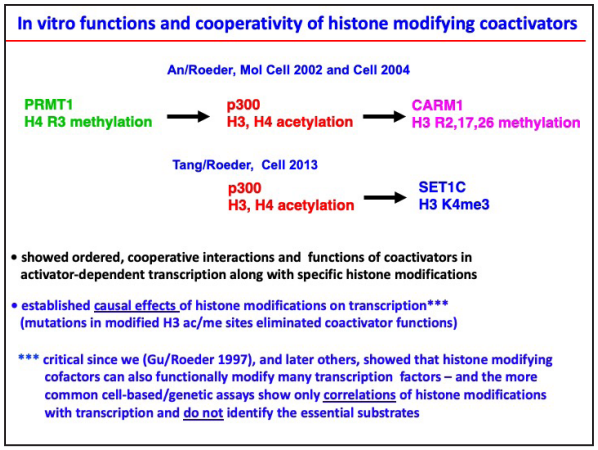

Fig. 29

Principal Discoveries and Achievements

Fig. 30 provides a general summary of these discoveries, which include: structurally and functionally distinct RNA polymerases, cognate general initiation factors, gene- and cell-specific transcriptional activators, general and gene- or cell-specific cofactors, mechanisms involved in the regulation of transcription through the use of biochemically defined systems with purified factors and recombinant DNA and chromatin templates, and a chromatin-based repression mechanism as well as causal roles for histone modifications in transcription. The diagram in Fig. 30 gives you a visual perception of the overall complexity of the factors and complexes required for the transcription of a simple single gene, requiring at least a hundred or more distinct polypeptides dispersed among the various factors.

These discoveries have been foundational for subsequent and future studies of the spectacular high resolution X-ray/cryo-EM studies of the transcriptional machineries, genomic analyses of gene activation mechanisms, cellular imaging studies of gene activation mechanisms, mechanisms underlying distal enhancer-promoter interactions and functions, an emerging role of phase separation (biological condensates) in gene activation, transcriptional regulatory circuits, and, importantly, the molecular basis and the therapeutic manipulation of aberrant transcription factor functions and transcriptional regulatory circuits found in many human pathologies.

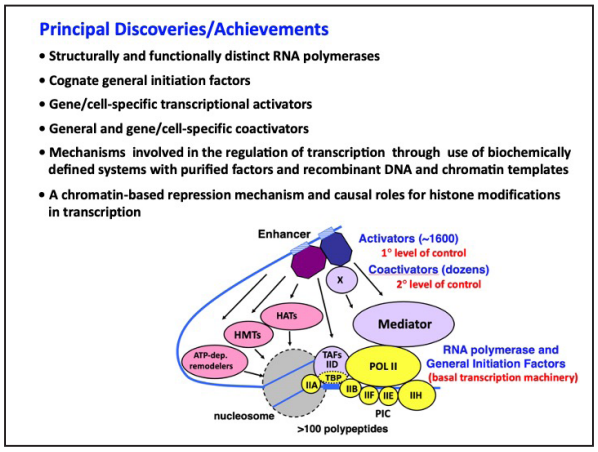

Fig. 30

Acknowledgements

In closing, I would like to acknowledge over 120 graduate and postdoctoral trainees who have contributed to this work, my undergraduate, graduate and postdoctoral mentors, my many colleagues in the transcription field, and my family—who have unfailingly supported my dedication to science. Fig. 31 shows three of my inspiring mentors in my early career development: my Wabash College professor Tom Cole, my Ph.D. mentor Bill Rutter, and my postdoctoral mentor Donald Brown. Fig. 32 shows some of my extraordinary students and postdoctoral trainees on the occasion of my 70 th birthday celebratory symposium in 2012 at The Rockefeller University. And lastly, I would again like to acknowledge the Inamori Foundation for this truly remarkable award.

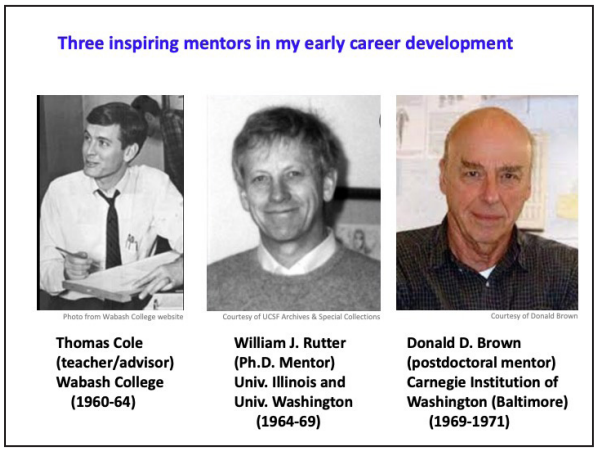

Fig. 31

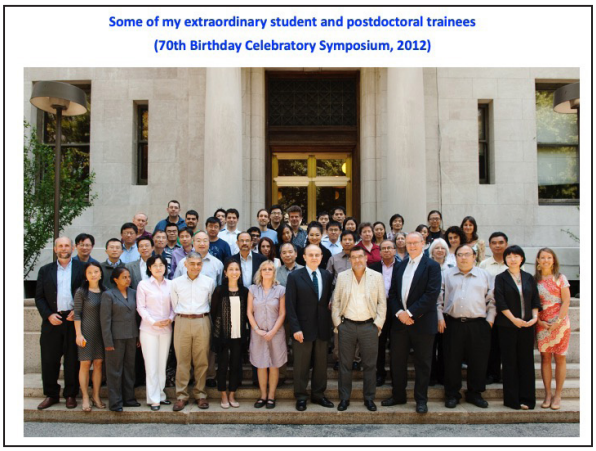

Fig. 32