

시스템생물학 연구사업

Systems Biology

간염 바이러스의 간암유발 기작 규명을 위한 간세포  
단백질 network 구축과 systematic 분석

Construction and analysis of protein network in hepatocyte  
for investigation of liver cancer induced by hepatitis viruses

서울대학교

과학기술부

제 출 문

과학기술부 장관 귀하

본 보고서를 “간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석”과제 의 보고서로 제출합니다.

2005. 5.

주관연구기관명	: 서울대학교
주관연구책임자	: 정 구 흥
협동연구책임자	: 박 근 수
연구원	: 박 성 규
"	: 이 태 형
"	: 임 승 외
"	: 류 현 미
"	: 김 진 욱
"	: 김 이 용
"	: 노 강 호
"	: 이 선 호
"	: 최 승 걸
협동연구기관명	: 한림대학교
협동연구책임자	: 황 순 봉
연구원	: 최 수 호
"	: 김 소 연
"	: 김 연 주
"	: 김 태 광
"	: 전 현 정

보고서 초록

과제관리번호	M1-0309-06-00 00	해당단계 연구기간	2003.6.23 - 2005.3.31	단계 구분	(1) / (3)
연구사업명	중 사업명	국책연구개발사업			
	세부사업명	시스템생물학연구사업			
연구과제명	중 과제명	간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석			
	세부(단위)과제명				
연구책임자	정 구 홍	해당단계 참여연구원수	총 : 17 명 내부 : 17 명 외부 : 명	해당단계 연구비	정부: 370,000 천원 기업: 천원 계: 천원
연구기관명 및 소속부서명	서울대학교 생명과학부		참여기업명		
국제공동연구	상대국명 :		상대국연구기관명 :		
위탁연구	연구기관명 :		연구책임자 :		
요약(연구결과를 중심으로 개조식 500자이내)				보고서 면수	35
<p>HBV 및 HCV protein들에 의한 세포내 신호전달 체계의 변화를 일으키는 동역학적 의미를 밝히기 위하여, 이번 연구과제에서는 다음과 같은 단계를 거쳤다.</p> <ol style="list-style-type: none"> <li>1. HBV 및 HCV protein들이 직접적으로 영향을 끼치는 protein들을 yeast two hybrid system 및 immunoprecipitation &amp; MALDI-TOF analysis 등을 통해 10가지 이상을 발굴.</li> <li>2. 위 protein-protein binding에 의해 영향을 끼칠 것으로 예상되는 protein expression 변화 분석을 cDNA microarray 및 2D electrophoresis 방법을 이용하여 분석하였다. 이 과정을 통해 변화되는 유전자 180 가지 이상을 발굴.</li> <li>3. 위 분석에 의해 영향 받는 세포내 신호 전달 체계에 대한 pathway를 설정하기 위하여 우선 문헌정보 분석을 통해 pathway를 가장 최신의 data에 맞도록 설정.</li> <li>4. 설정된 pathway (NF-<math>\kappa</math>B pathway, IFN-<math>\gamma</math> pathway, TGF-<math>\beta</math> pathway, JNK pathway 등) 내 protein들의 초기 상태의 농도 및 activated form의 농도를 문헌에서 찾고, 찾을 수 없을 경우 실험을 통하여 그 값을 얻도록 함.</li> <li>5. in silico simulation을 하고, pathway에 대한 검증.</li> </ol>					
색인어 (각 5개 이상)	한 글	시스템생물학, B형 간염 바이러스, 네트워크, 모델링, 시뮬레이션			
	영 어	systems biology, hepatitis B virus, network, modeling, simulation			

요 약 문

I. 제 목

간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석

II. 연구개발의 목적 및 필요성

HBV 및 HCV는 다양한 viral protein들은 간세포내에서 다양한 protein과 서로 상호작용하며, 이것은 이 세포내에서 다양한 생리학적 변화를 일으킨다. 이 변화를 일으키는 원동력과 또 이것에 의해 변화되는 protein들의 활성화 및 발현 수준 변화를 분석하여 HBV 및 HCV에 의한 간세포에 영향을 전체적으로 분석할 수 있는 결과를 도출하는 것이다. 또, 이와 같이 도출된 데이터를 저장하고 관리할 수 있는 system을 구축한다. 지식 기반 경제로 전환하는 시점에 있어서 BT와 IT 기술은 핵심적인 역할을 하게 된다. 본 연구를 통해서 두 기술이 결합된 바이오인포매틱스 연구를 촉진시켜서 국가경쟁력 향상에 기여할 수 있다.

III. 연구개발의 내용 및 범위

본 연구팀은 우선 HBV 및 HCV의 감염을 모방하는 in vitro system을 구성하였고, 이 system을 이용하여, HBV 및 HCV에 의해 세포내에서 변화되는 요소를 선별하였다. 선별된 요소들은 HBV 및 HCV에 의해 영향받는 pathway의 구조를 결정하는 initial data로 활용되었으며, 이후 구조의 확정을 통해 in silico modeling을 시도하였다. 이 in silico modeling은 다시 실험적 접근 방법에 의하여 검증을 시도하였다.

IV. 연구개발결과

이 연구과제를 통해 Viral protein 10 가지 protein이상 expression, HBV 및 HCV에 의해 변화되는 유전자 180 가지 이상 탐색, NF-kB 활성화도 변화, JNK 활성화도 변화, apoptosis 영향 분석, 데이터베이스 및 인터페이스 설계, 4 가지 pathway 이상 설정 (20-30 가지의 pathway내 protein의 역할 정량적 분석), marker 50 종류 이상 발굴, 그리고, 3 종류 이상의 HBV 및 HCV와 연관된 in silico model을 완성하였다.

V. 연구개발결과의 활용계획

이 연구과제를 통해 구성된 모델은 간염 관련 in silico 약물 target 선정, 간염 치료 약물의 효과 및 부작용(side effect)에 대한 in silico 평가, 검증, 간질환 환자의 임상 데이터 적용 및 in silico 예측, 그리고, 환자 임상 데이터에 기반한 최적의 치료 전략 시뮬레이션 등과 같은 연구에 활용될 것으로 기대된다.

## S U M M A R Y

(영 문 요 약 문)

### I. Title

Construction and analysis of protein network in hepatocyte for investigation of liver cancer induced by hepatitis viruses

### II. The Present State of Technical Development

Both HBV and HCV induce various physiological changes through interacting cellular proteins. In this project, we analyzed the effect of HBV and HCV on hepatocyte through protein-protein interaction and differential gene expression analysis. This analyzed data is stored in database system. This fusion of BT and IT is contributed in increasing of national research power.

### III. Result of the Project

Signal pathways related to HBV and HCV are modeled in in silico system and wet-derived data are integrated to the model. With the model, in silico simulational analysis performed and the data are confirmed by wet-derived experiment.

### IV. Achievement of the Project

From this project, we obtained the results such as expression of viral proteins (over 10 proteins), differentially expressed genes by HBV and HCV (over 180 genes), HBV and HCV related pathway analysis (4 pathway), and construction of in silico model (3 pathway).

### V. Utilization of the Research Results

The constructed in silico model can be applied for verification of drug effect (NF-kB pathway inhibiting chemicals) in NF-kB pathway.

C O N T E N T S

(영 문 목 차)

The First. Summary of the Project.....	7
The Second. The Present State of Technical Development.....	9
The Third. Result of the Project.....	13
The Fourth. Achivement of the Project.....	27
The Fifth. Utilization of the Research Results.....	31
The Sixth. Information of Technical Development in Foriegn Country.....	33
The Senventh. References.....	34

목 차

제 1 장 연구개발과제의 개요.....	7
제 2 장 국내외 기술개발 현황.....	9
제 3 장 연구개발수행 내용 및 결과.....	13
제 4 장 목표달성도 및 관련분야에의 기여도.....	27
제 5 장 연구개발결과의 활용계획.....	31
제 6 장 연구개발과정에서 수집한 해외과학기술정보.....	33
제 7 장 참고문헌.....	34

## 제 1 장 연구개발과제의 개요

### 제 1 절 연구개발의 목적

1. HBV 및 HCV는 다양한 viral protein을 production한다. 이 protein들은 간세포내에서 다양한 protein과 서로 상호작용하며, 이것은 이 세포내에서 다양한 생리학적 변화를 일으킨다 (Seeger and Mason, 2000). 이 변화를 일으키는 원동력과 또 이것에 의해 변화되는 protein들의 활성 및 발현 수준 변화를 분석하여 HBV 및 HCV에 의한 간세포에 영향을 전체적으로 분석할 수 있는 결과를 도출하는 것 (직접 또는 간접적으로 상호작용하는 protein 180가지 이상의 도출)이다 (서울대 정구홍, 한림대 황순봉). 또, 이와 같이 도출된 데이터를 저장하고 관리할 수 있는 system을 구축 (서울대 박근수) 한다.

### 제 2 절 연구개발의 필요성 및 범위

1. 본 연구의 세부과제 책임자는 BT에서 2명 (간염 B, C 전문가), IT에서 한명 (Bioinformatics 전문가) 등 각 분야에서 국제적인 경쟁력을 갖춘 세 명의 과학자가 공동으로 team을 갖추고 있어 기술적 보완이 완전한 협동연구팀으로 구성되어 있다. 따라서 바이러스성 간염 유발의 원인을 규명하고 이에 대한 치료방법을 개발할 수 있는 연구는 국가의 정책적 지원이 선행될 경우 국내 기술에 의해 확립될 수 있다고 사료된다.

2. 간염 바이러스와 간세포내 단백질 네트워크와의 상호 작용을 컴퓨터를 통해 모델링하며, 실험을 통해 축적된 데이터로 데이터베이스를 구축하며, 데이터베이스에 저장된 상호 작용 과정을 연구자들이 직관적으로 쉽게 이해할 수 있도록 보여주는 인터페이스를 구축하는 것을 목표로 한다. 이 과제의 수행을 통해서 간염에 관련된 데이터베이스 구축 및 in silico 모델 시스템을 구축하게 되면 해당 분야에 대한 연구에 지대한 도움이 될 것으로 기대된다.

3. 전 세계적으로 간염 및 간암치료제 시장은 약7조원 이상으로 추정되고 있으며 여기에 HBV 및 HCV 치료제가 개발될 경우 그 규모는 더욱 클 것으로 기대된다. 따라서 C형 간염바이러스의 제어 기술을 간염치료제 개발로 연계할 경우 노동력 창출 및 제약산업의 활성화로 인한 산업발전 및 소득 증대와 수출에 의한 외화획득에 기여할 것이 확실시 되어 경제·산업적 기대치가 매우 높다고 사료된다.

4. 간염 바이러스는 우리 나라 전 인구의 10%가 감염되어 있을 뿐 아니라, 간암의 주요 원인이 되기도 한다 (Feitelson, 1999). 간염 바이러스에 대한 연구와 실험 데이터를 통한 모델링과 검증, 데이터베이스 개발 등을 통해서 치료제 개발에 직접적인 도움을 줄 수 있다. 이의 경제적 파급 효과는 매우 클 것으로 기대된다.



5. 우리나라에서 한해 암에 걸리는 사람들 중 간암에 걸리는 사람의 수는 위암에 이어 두 번째이다 (2000년 보건복지부). 간암은 한번 걸리면 5년 이내에 사망률이 췌장암에 이어 두 번째로 (2002년 국립암센터) 더욱더 심각하다. 또한, 대부분의 간암 환자가 40대 남성으로 사회 및 경제적으로도 매우 큰 부담요인이 되고 있다.
6. 우리나라의 경우 전 인구의 약 10% (450여 만명)가 HBV에 감염되어 있는 것으로 나타나 있고, B형 간염환자의 약 10%가 간암으로 진행되어 사망하는 것으로 보고되고 있다. 즉 간암의 주요 원인이 HBV에 의한 감염이다. 감염과 간암에 의한 사회, 경제적 손실이 막대한 실정으로 치료제 개발을 위한 연구가 매우 시급하다.
7. 전 세계적으로 C형 간염환자는 1억 7000만 명이상이고 매년 일백만 명 이상의 새로운 간암환자가 발생하여 이로 인한 사회적·문화적 문제점을 대두시키고 있으며 인류 보건에 지대한 위협요소가 되고 있다. 아직 HCV는 백신이 개발되지 않았으며 감염자도 지속적으로 증가하고 있다.
8. 지식 기반 경제로 전환하는 시점에 있어서 BT와 IT 기술은 핵심적인 역할을 하게 된다. 본 연구를 통해서 두 기술이 결합된 바이오인포매틱스 연구를 촉진시켜서 국가경쟁력 향상에 기여할 수 있다.

## 제 2 장 국내외 기술개발 현황

### 제 1 절 연구사례의 조사

1. Systems Biology 연구는 국내외적으로 아직 개발기에 속해있다. 특히 전체 기술 분야를 주도할 수 있는 이렇다 할 구체적인 방법론이 개발되어 있지 않아 대상과 방법에 대한 연구의 여지가 많이 남아있다 하겠다. 이로 인해 최근 들어 다양한 목표를 가지고 Systems Biology 연구에 참여하는 국내외 연구진들이 하루가 다르게 늘어나고 있으며, 이는 비단 생명 과학 분야뿐만 아니라 컴퓨터 과학, 제어 공학, 물리학 등 다양한 학문 분야의 관심을 불러 일으켜 학제간(interdisciplinary) 학문 연구를 촉진시키고 있다.
2. 생물정보 데이터베이스와 관련하여 외국의 경우 기존에 밝혀진 생명 현상과 관련된 지식을 통합한 데이터베이스를 구축하는 작업이 진행되고 있다. 스위스 바이오인포매틱스 연구소의 단백질 데이터베이스인 ExPASy (Expert Protein Analysis System)나, 일본 교토대학교의 KEGG (Kyoto Encyclopedia of Genes and Genomes) 등이 그 대표적인 예이다.
3. 한편, 2000년대에 들어 포스트 지놈 시대 생명 과학 연구가 본격화됨에 따라 생명 현상에 대한 *in silico* 모델 구축을 통해, 축적된 *in vitro/in vivo* 실험 데이터를 통합하고, 생명 현상에 내재된 동역학적 특성을 밝히려는 움직임도 활발히 진행되고 있다. *in silico* 모델링은 궁극적으로, 세포 내 물질들의 상호 작용 경로를 수학적 모델로 기술하여, 세포의 상태 변화 및 물질 간 상호 작용이 가지는 동적 특성을 *in silico* 상에서 시뮬레이션하는 E-Cell을 구축하고자 한다. 이를 위하여 미국, 일본, 독일 등 선진 각국에서는 다양한 학문 분야를 기반으로 한 연구자들의 상호 교류 및 협력 연구가 활발히 시도되고 있다.
4. 특히 신호전달경로의 모델링과 관련한 연구로는, 경로를 이루는 주요 구조인 kinase cascade, receptor system 및 transcription factor 조절 시스템에 관한 연구가 다각적으로 진행되고 있다. 그러나 MAPK cascade 및 EGFR 시스템 등 일부 시스템에만 제한적으로 모델링이 이루어지고 있으며, 이를 종합한 전체 신호전달경로를 모델링하는 연구는 아직 연구 초기 단계에 있다 (Asthagiri et al., 2001).
5. 이와는 별도로 wet 실험의 target을 설정하기 위해, sensitivity analysis, Monte Carlo 시뮬레이션 등 공학적 분석 방법을 적용하여 *in silico* 모델 상에서 중요한 파라미터 및 molecule을 선별하는 연구도 진행되고 있다. 이러한 연구들은 향후 dry 시뮬레이션을 통한 drug target 발굴 등에 적극 활용될 전망이다. 그러나 아직 특정 질병, 질환과 관련된 신호전달경로의 *in silico* 모델 개발은 미진하다.
6. 선진 각국의 활발한 연구 동향과 비교해 볼 때 국내 *in silico* 모델 연구는 일부 분야를 제외하고는 아직 초기 단계에 머무르고 있다. 특히 wet 실험실과 dry 실험실 간의 실질

적인 연구 협력 체계 구축이 미비한 형편이다.

7. 본 연구에서는 인간 간세포에서 HBV, HCV와 관계된 신호전달경로를 모델링하고, 이에 대한 *in silico* 시뮬레이션 및 *in vitro* 검증을 시도하였다. 이는 국내외적으로도 선행 연구를 찾아보기 힘든 것으로, *in silico* 모델링을 통한 질병의 작용 기작에 대한 동역학적 원리 규명 뿐만 아니라, 치료 목적으로의 응용 가능성을 열었다는 점에서 큰 의미를 지니고 있다.

8. 또한, 본 과제를 수행하기 위해 wet/dry 실험실 간 구축해온 긴밀한 연구 협력 관계는 아직 걸음마 단계에 있는 국내 Systems Biology 연구 분야에서, wet/dry 연구 간의 성공적인 협력 사례를 제시할 수 있으리라 기대된다.

#### 9. 외국의 경우

가. 시스템 바이올로지에 대한 연구는 게놈프로젝트가 발표된 2000년에 하나의 조류로 자리잡았다. 이미 스탠포드, 캠브리지, 버지니아 공대, MIT, 하버드, 칼텍, UCLA서 20여 개 팀이 이 분야 연구에 뛰어들었고, 2002년부터 'System biology'를 주제로 국제 학술회의를 개최했다.

#### 나. 미국의 system biology의 선두 주자

(1) ISB (Institute for System Biology)의 리로이 박사팀

(가) DNA, mRNA, protein, 기능 protein간의 상호 연결 방식과 신호 전달 체계에 대한 이해를 통해 신약개발이 돌파구를 찾을 수 있을 것이라고 주장

(나) 질병의 원인이 세포간 신호전달 체계의 이상에서 기인하므로, 세포간 네트워크를 이해하고 이를 조절한다면 획기적인 신약을 개발할 수 있다는 설명

(다) 특히 암이나 관절염 에이즈와 같은 면역체계 관련 병에서 획기적인 치료법을 제시할 수 있을 것으로 기대

#### 다. 일본의 system biology 선두 주자

(2) 소니 컴퓨터 사이언스 연구소의 키타노 박사팀

(가) 생명현상을 이해하고, 이 연구결과를 컴퓨터 사이언스와 접목해 사이보그와 같은 인공생명체를 구현하는데 초점

(나) 이를 위해 키타노는 대장균과 초파리 연구를 통해 생명현상을 컴퓨터 시뮬레이션화하는 작업에 초점

(다) 휴먼로봇 시그(Sig)와 피노(PINO)도 이 연구의 일환이다.

#### 라. HBV에 대한 system biology

(1) 현재 system biology 체계를 통한 HBV 연구의 접근이 이루어지지 않음

10. 국내의 경우

가. KAIST의 이상엽 박사팀

- (1) prokaryote에서의 metabolic pathway에 대한 system biology
- (2) in silico 상에서의 metabolic pathway의 설정
- (3) 이에 따른 실험 설계 및 결과 분석

11. 조사연구개발사례에 대한 평가

- 가. 국외의 경우 개념의 설정이 이제 시작 되는 분야로 아직 개발기에 속함
- 나. system biology를 이용한 HBV에 대한 연구의 시작이 필요

## 제 2 절 세부 기술사항의 검토 분석

### 1. 국내·외 기술수준 비교표

	비 교 (%는 선진국대비)
Gene expression analysis tool	거의 모든 (> 99%) 기술이 선진국과 대등함
Proteom analysis tool	장비편을 제외한 기술이 선진국과 대등 (> 90%)
Signal pathway analysis tool	최근의 BT 연구와 더불어 이 분야의 연구 기술력이 상승 (> 80%)
In vitro reconstitution analysis tool	최근의 BT 연구와 더불어 이 분야의 연구 기술력이 상승 (> 80%)
Protein expression & purification system	전통적인 기술은 국내에도 기술이 축적됨 (> 90%)
High through put analysis tool	자동화 기술과 함께 발전 하지만, 장비 분야의 발전은 미흡 (< 80%)

### 2. 공정단위별로 주요 기술사항 및 그 기술수준의 분석평가를 다음 사항에 걸쳐 기술함

가. 외국의 경우

	국 외 (아직 미흡한점이 초점)
Gene expression analysis tool	고집적, 빠른 시간, 낮은 error율을 목표로 자동화에 초점
Proteom analysis tool	분석의 자동화에 초점
Signal pathway analysis tool	system biology를 통한 대량 분석으로의 접근이 초점
In vitro reconstitution analysis tool	Kit화와 정량화에 그 초점
Protein expression & purification system	빠른시간, 낮은 비용, 자동화에 그 초점
High through put analysis tool	극미량 handling 기술에 그 초점

나. 국내의 경우

	국 내
Gene expression analysis tool	Scanning 기술, 분석 software등에 그 초점
Proteom analysis tool	장비면의 발전이 미흡
Signal pathway analysis tool	system biology의 개발기에 참여
In vitro reconstitution analysis tool	선진국의 기술력의 이용, 국내에서의 bioassay의 kit화는 미흡
Protein expression & purification system	실험 기술면은 선진국과 대등하나 장비면에서의 발전은 미흡
High through put analysis tool	극미량 handling 기술에 그 초점

## 제 3 장 연구개발수행 내용 및 결과

1절. 본 연구팀의 연구 수행은 지난 2년간 다음과 같이 진행.

1. HBV 감염에 의한 세포내 변화 분석: 서울대학교 생명과학부
2. HCV 감염에 의한 세포내 변화 분석: 한림대학교 일송생명과학연구소
3. HBV 및 HCV 감염에 의한 세포내 변화의 in silico simulation: 서울대학교 컴퓨터 공학과

### 2절. 이론적 접근 방법

HBV 및 HCV에 의한 in silico system에서의 systematic 분석을 위해서는 우선 HBV 및 HCV 감염에 의해 세포내에서 변화되는 요소를 분석하는 것이 중요하다. 이를 위해 본 연구팀은 우선 HBV 및 HCV의 감염을 모방하는 in vitro system을 구성하였고, 이 system을 이용하여, HBV 및 HCV에 의해 세포내에서 변화되는 요소를 선별하였다. 선별된 요소들은 HBV 및 HCV에 의해 영향받는 pathway의 구조를 결정하는 initial data로 활용되었으며, 이후 구조의 확정을 통해 in silico modeling을 시도하였다. 이 in silico modeling은 다시 실험적 접근 방법에 의하여 검증을 시도하였다.

즉, 본 연구과제의 이론적 접근은 생물학적 data를 바탕으로 in silico modeling을 구성하였으며, 구성된 model에서의 문제 및 실험적으로 풀긴 힘든 문제를 in silico model을 통해 접근하면서, 다시 실험적 접근 방법에 의한 검증 과정을 거치게 되었다.

본 연구팀의 연구 과정은 단순히 in silico model 및 data 해석만을 목적으로 한 구성이 아니라, 본 과제를 통해 완성된 모델이 기타의 생물학적 연구 탐색에 응용될 수 있는 실험적 모델 (실제 실험과 가까운 모델) 이 될 수 있도록 최선을 다하였다.

### 3절. 실험적 접근 방법

HBV 및 HCV protein들에 의한 세포내 신호전달 체계의 변화를 일으키는 동역학적 의미를 밝히기 위하여, 이번 연구과제에서는 다음과 같은 단계를 거치도록 하였다.

1. HBV 및 HCV protein들이 직접적으로 영향을 끼치는 protein들을 yeast two hybrid system 및 immunoprecipitation & MALDI-TOF analysis등을 통해 찾아 내었다.
2. 위 protein-protein binding에 의해 영향을 끼칠것으로 예상되는 protein expression 변화 분석을 cDNA microarray 및 2D electrophoresis 방법을 이용하여 분석하였다.
3. 위 분석에 의해 영향 받는 세포내 신호 전달 체계에 대한 pathway를 설정하기 위하여 우선 문헌정보 분석을 통해 pathway를 가장 최신의 data에 맞도록 설정하였다.
4. 설정된 pathway (NF- $\kappa$ B pathway, IFN- $\gamma$  pathway, TGF- $\beta$  pathway, JNK pathway 등) 내 protein들의 초기 상태의 농도 및 activated form의 농도를 문헌에서 찾고, 찾을 수 없을 경우 실험을 통하여 그 값을 얻도록 하였다.
5. 협동 연구기관인 서울대학교 컴퓨터 공학과 팀과 공동으로 in silico simulation을 하

고, pathway에 대한 검증을 하였다.

#### 4절. 연구내용 및 연구결과

본 연구과제를 통해 1-2년차 연구 기간 동안 본 연구팀은 타겟의 설정 및 타겟이 된 pathway내 protein들을 이용한 in silico simulation을 할 수 있는 정량적 분석을 하였다. 그 연구는 아래와 같이 진행 되었다.

##### 1. HBV protein과 직접 binding하는 protein screen

###### 가. Yeast two hybrid system을 이용한 HBV polymerase binding protein screen

(1) HBV Polymerase는 in vitro system에서 상당히 expression하기 힘들며, expression이 되더라도 conformational state가 상당히 불안정하다. 이로 인하여, heterologous system에서 expression된 protein을 이용하여 binding하는 protein을 찾기는 매우 힘들다. 이와 같은 이유로 HBV polymerase binding protein은 yeast two hybrid system을 이용하여 탐색하였다. 그 결과 HBV polymerase와 binding하는 Apolipoprotein을 발굴하였다.

###### 나. Affinity chromatography를 이용한 HBV core binding protein screen

(1) Core protein과 binding하는 host protein들을 찾기 위한 affinity chromatography technique를 사용했다. Core protein은 direct covalent linkage를 통해 NHS-activated sepharose에 immobilization하였다. 이 실험에서 immobilized core antigens과 binding한 host protein은 MALDI-TOF MS에 의해 identification되었다. 그 결과 5가지의 candidate protein들을 찾아내었다. 이들 중 하나는 human nucleolar phosphoprotein B23 (nucleophosmin)이었다. 게다가, HepG2에서 cellular B23은 co-immunoprecipitation, sucrose density gradient, cell cycle synchronization 실험을 통해 core antigens과 binding함을 보였다.

###### 다. Mammalian expression system을 이용한 surface protein 및 HBx binding protein screen

(1) Surface protein은 lipoprotein으로 prokaryotic system에서 원래의 conformation을 유지하면서 expression하여 얻기 힘들다. 또한, HBx protein도 inclusion되는 경향이 강하다. 하지만, mammalian cell에서 expression시 다른 protein들에 비해 상대적으로 expression을 잘 할 수 있었다. 이 과정에서, surface protein을 expression하기 위한 pCMV-HA/HBs와 HBx expression 하기 위한 pCMV-HA/HBx를 construction하였다. Expression된 protein의 immunoprecipitation & MALDI-TOF analysis 결과 HBx protein에 binding하는 HBXIP와 TFIIB protein을 찾아내었고, surface protein과 binding하는 Hsp70를 찾아내었다.

2. Yeast two-hybrid system, baculovirus expressio system, 그리고 mammalian system에서 사용하기 위해 본 연구팀은 C형 간염바이러스의 단백질을 발현하는 부위를 expression vector에 클로닝하였으며, 이들의 sequence를 sequencing을 통해 확인하였다

(Fig. 1).

3. TGF- $\beta$ 에 의해 유도되는 signal transduction pathway는 Smad 단백질이 heterodimerization에 의해 nuclear로 translocation됨으로써 gene transcription을 유도한다. 이러한 과정에서 NS5A 단백질이 미치는 영향을 분석하고자 Smad3와 Smad4 단백질을 NS5A와 co-transfection한 후 Smad3와 Smad4의 complex형성을 co-immunoprecipitation으로 분석한 결과 NS5A 단백질이 TGF- $\beta$ 에 의해 유도되는 Smad3와 Smad4의 complex형성을 억제함을 관찰하였으며 (Fig. 17) 또한 Smad-DNA complex 형성도 억제됨을 확인하였다 (Fig. 18). 이러한 결과는 C형 간염바이러스의 NS5A 단백질이 세포내 단백질인 T $\beta$ R-I과 직접적으로 상호 작용함으로써 TGF- $\beta$ 에 의해 유도되는 signal transduction pathway를 억제할 것으로 사료된다.

#### 4. HBV에 의해 직·간접적으로 영향받는 유전자 screen

가. 본 연구실에서 확립된 primary normal human hepatocyte culture system을 이용한 HBV 감염 전후의 gene expression 변화 분석

(1) Culture cell line에는 HBV가 감염되지 않으므로, 실제 HBV 감염에 따른 분석을 하기 위해서는 primary culture가 필수적이다. 본 연구실에서는 이전 연구과제(생명현상 및 기능연구사업)를 통하여 human hepatocyte를 primary culture하는 기술을 확립하였다. 이번 연구과제에서도 이 system을 활용하여 HBV 감염 전후 변화되는 유전자 발현을 분석하였다. 그 결과 149 가지의 유전자가 의미성 있게 변화됨을 발견하였다.

나. 임상 sample 중 HBsAg(+) 간암 환자와 HBsAg(-) 간암 환자의 liver specimen에서의 2DE technique를 이용한 protein display 및 비교 분석

(1) NBNC-, B-type HCC 환자에서 얻은 tumorous tissue sample에서 total protein들을 two-dimensional gel을 통해 분리하여 1000 ~ 1200개의 spot을 찾았다. 각각 tissue sample로부터 얻은 two-dimensional gel image를 분석하여 protein spot들의 volume percentage를 nontumorous tissue sample과 tumorous tissue sample에서 서로 비교하였다. 이렇게 선택된 protein spot은 MALDI-TOF mass spectrometric analysis를 통해 protein identification되었다. 약 60개 정도의 protein이 clearly identification되었다.

significant variation을 보이는 60 protein들 중에 7개의 stress-associated protein들과 nucleophosmin, elongation factor 2는 모든 type의 HCC에서 증가하였다. 반대로 CPSASE I과 argininosuccinate synthase는 모든 type의 HCC에서 감소하였다

Hepatitis virus 감염에 의해 differentially expressed protein들의 analysis는 NBNC proteins뿐 아니라 HBV-associated protein들을 알 수 있게 해준다. SOD (Cu-Zn)의 expression은 B-type HCC (5 of 7,  $2.5 \pm 1.1$  ; 그림 4에서 decrease 경향을 보였고 dihydrolipoamide dehydrogenase는 NBNC-type HCC (7 of 7,  $2.4 \pm 0.5$ )에서 증가하는 경향을 보였다. ribosomal binding protein은 NBNC type HCC에서는 거의 변화가 없지만 B-type HCC (6 of 7,  $3.6 \pm 2.8$  ; Table 2)에서는 clearly decrease하였다.



5. C형 간염바이러스는 현재까지 적당한 cell culture system 및 animal model이 없는 실정임. 하지만 최근에 hepatoma cell line (Huh7)을 이용하여 viral RNA가 스스로 복제할 수 있는 system이 개발되었으며 이 cell을 이용하여 현재 많은 연구가 진행되고 있다. 이러한 cell을 이용한 microarray analysis는 한번의 실험을 통해 HCV-associated marker gene을 다량으로 확보할 수 있는 장점이 있다. 따라서 본 연구팀은 이 cell을 이용하여 microarray analysis를 수행한 결과 2배 이상 up-regulated gene 840개, 2배 이상 down-regulated gene 1,896 개, 그리고 4배 이상 fold-changed gene 142개를 확보하였다. 이러한 유전자들에서 cytokine-related gene은 약 200여개 정도 되며 이들 중 약 50개정도의 유전자에 본 연구팀이 관심을 가지고 있다.

6. C형 간염바이러스의 NS5A 단백질은 multifunctional한 단백질로써 C형 간염바이러스에 의한 pathogenesis에 중요한 역할을 함. 따라서 본 연구팀이 확보하고 있는 NS5A 단백질에 대한 transgenic mice의 liver를 이용해 간의 morphology를 관찰한 결과 liver에서 tumor가 형성됨을 관찰하였다 (Fig. 4). 위의 결과를 확인하고자 tumor 조직을 이용해 histological analysis를 수행한 결과 mice의 liver를 section하여 H&E staining한 결과 정상적인 mice와는 달리 간세포 핵의 크기가 다양하게 나타남을 관찰하였고, 또한 glycogen storage foci가 형성되어 있음을 관찰하였다 (Fig. 5). 이러한 결과로 미루어 볼 때 C형 간염바이러스의 NS5A 단백질이 C형 간염바이러스가 유도하는 간암 형성에 매우 중요한 역할을 할 것으로 생각됨. 또한 transgenic mice의 liver를 이용해 microarray analysis를 수행하였다. 그 결과 2배 이상 fold-changed gene 743개, 4배 이상 fold-changed gene 206개를 확보하였고 (Fig. 6), 특히 4배 이상 fold-changed gene들 중 약 23개의 유전자에 관심을 가지고 있는데 이러한 유전자들은 signal transducer, apoptosis regulator, cell cycle regulator, fibrosis, cancer, chaperone 등의 기능을 가지고 있다 (Table 2). 따라서 microarray data가 C형간염바이러스와 관련된 간암 후보 유전자 발굴에 상당히 중요한 database로 활용될 수 있을 것이다.

## 7. HBV 관련 신호전달경로 ODE 모델 구성 및 검증

### 가. NF- $\kappa$ B 신호전달경로

NF- $\kappa$ B 신호전달경로는 세포의 사멸 및 생장을 조절하는 중요한 신호전달경로이다. NF- $\kappa$ B는 TNF $\alpha$  등 proinflammatory cytokin에 의해 주로 activation 되는데, 이러한 세포 외부 신호가 전달되어, IKK가 phosphorylation되면, NF- $\kappa$ B와 binding된 I $\kappa$ B의 ubiquitination을 유도함으로써 NF- $\kappa$ B가 핵으로 이동한다는 메커니즘이 밝혀져 있다. NF- $\kappa$ B는 여러 질병과 치료 작용에 관련되어 있는데, 제1세부과제팀의 연구 결과 특히 HBV 감염에 의해 그 activation이 증가하는 것으로 밝혀졌다. 본 연구에서는 NF- $\kappa$ B 신호전달경로의 *in silico* 모델을 구성하고, 제 1세부과제팀의 wet 실험 결과를 바탕으로 HBV가 NF- $\kappa$ B activation 신호전달경로에 주는 동역학적 변화에 대한 모델링 및 시뮬레이션을 진행하였다.

### 나. ODE 모델의 구성

NF- $\kappa$ B activation과 관련하여, IKK에 의해 조절되는 NF- $\kappa$ B 신호전달 모듈이 Hoffmann 등에 의해 ODE로 모델링되었다[Hoffman et al., 2002]. 이후 IKK downstream에서 NF- $\kappa$ B activity가 보이는 다양한 동적 특성을 *in silico* 모델을 이용해 규명하려는 연구가 활발히 진행되고 있다. 한편, NF- $\kappa$ B를 activation시키는 주요 cytokine인 TNF $\alpha$ 로부터 시작되는 신호전달경로 모델을 구성하려는 연구도 진행 중에 있다. 그러나 아직까지 *in silico/in vitro* 실험 협력을 통한 모델 검증 및 특정 질병과 관련한 신호전달경로의 변화 모델링 등의 연구는 미흡한 실정이다.

본 연구에서는 먼저, 기존 연구를 통해 부분적으로 구성된 NF- $\kappa$ B 신호전달경로 모델을 종합하여 전체 신호전달경로 *in silico* 모델을 구성하고 이에 대한 시뮬레이션을 진행하였다.

*in silico* 모델의 설계 및 분석을 용이하게 하기 위하여, 전체 시스템을 기능 및 세포 내 위치에 따라 나누는 모듈화 원칙을 적용, 전체 신호전달경로 모델을 아래와 같이 3개의 서브시스템으로 구성하였다.

#### 다. TNF $\alpha$ receptor 시스템

세포 외부에서 전달되는 cytokine TNF $\alpha$ 가 세포막에 닿아 TNFR1 receptor 및 TRADD, RIP 등의 receptor adaptor 단백질과 차례로 complex를 이루는 과정을 ODE 모델로 구성하였다. 최종적으로 TRAFs 단백질이 complex에 참여하면, 이를 통해 다음 단계인 kinase cascade가 activation 된다. FADD 단백질은 세포 사멸과 관련된 또 다른 신호전달경로를 매개하는 물질로 본 모델에서는 receptor complex에 binding되는 부분만 모델 되었다. Receptor 시스템과 관련하여 EGFR receptor 시스템 연구(Kholodenko et al., 1999) 등 선행 연구의 파라미터들을 참고하였다.

#### 라. Kinase cascade 시스템

Kinase cascade는 일련의 kinase 단백질들이 연쇄적으로 다음 단계 protein을 phosphorylation을 통해 activation시키는 신호전달 구조이다. 지금까지 대표적인 kinase cascade인 MAPK cascade에 대한 연구가 활발히 진행되어 왔으며, amplification, switching behavior, ultrasensitivity 등 kinase cascade 구조가 갖는 동적 특성들도 많이 연구되고 있다. 본 연구에서는 1단계 TNF $\alpha$  receptor 시스템에서 형성된 TRAFs complex를 입력 kinase로 IKKK(RIPK, NIK 등), IKK가 차례로 activation되는 모델을 구성하였다.

또한, kinase cascade 모델에는 kinase를 deactivation시키는 phosphatase도 포함되어 있다. 본 연구에서는 이 phosphatase와 kinase의 상대적 비율에 따라 kinase cascade의 동적 특성이 결정됨을 *in silico* 시뮬레이션을 통해 확인하였다.

#### 마. NF- $\kappa$ B 조절 시스템

2단계 kinase cascade에서 activation된 IKK는 NF- $\kappa$ B와 binding된 I $\kappa$ B isoforms를 phosphorylation함으로써 I $\kappa$ B의 ubiquitination 및 proteolysis를 유도한다. I $\kappa$ B의 감소는 세포질 내에 묶여 있던 transcription factor NF- $\kappa$ B가 핵으로 이동하는 것을 허용함으로써 DNA 상에서 NF- $\kappa$ B와 binding하는 여러 가지 유전자들을 transcription한다. 이 유전자들 중 핵 내에 있는 NF- $\kappa$ B를 다시 세포질로 이동시키는 I $\kappa$ Ba가 포함되어 있어, NF- $\kappa$ B를 조절하는 핵심적인 역할을 담당하

고 있다. 이러한 negative feedback loop 시스템은 NF- $\kappa$ B 조절 시스템의 동역학 모델링에 있어 중요한 연구 주제로 주목받아 왔다. 본 모델에서는 Hoffmann의 NF- $\kappa$ B 조절 모델 중 NF- $\kappa$ B 동역학적 변화에 결정적인 역할을 담당하는 I $\kappa$ B $\alpha$ 와 관계된 구조와 파라미터를 사용하였다.

위와 같이 서브시스템 모델을 구성하고, 각각의 동역학 특성을 분석한 후, 이를 종합하여 28개의 단백질과 39개의 파라미터로 구성된 전체 신호전달경로 모델을 구성하였다.

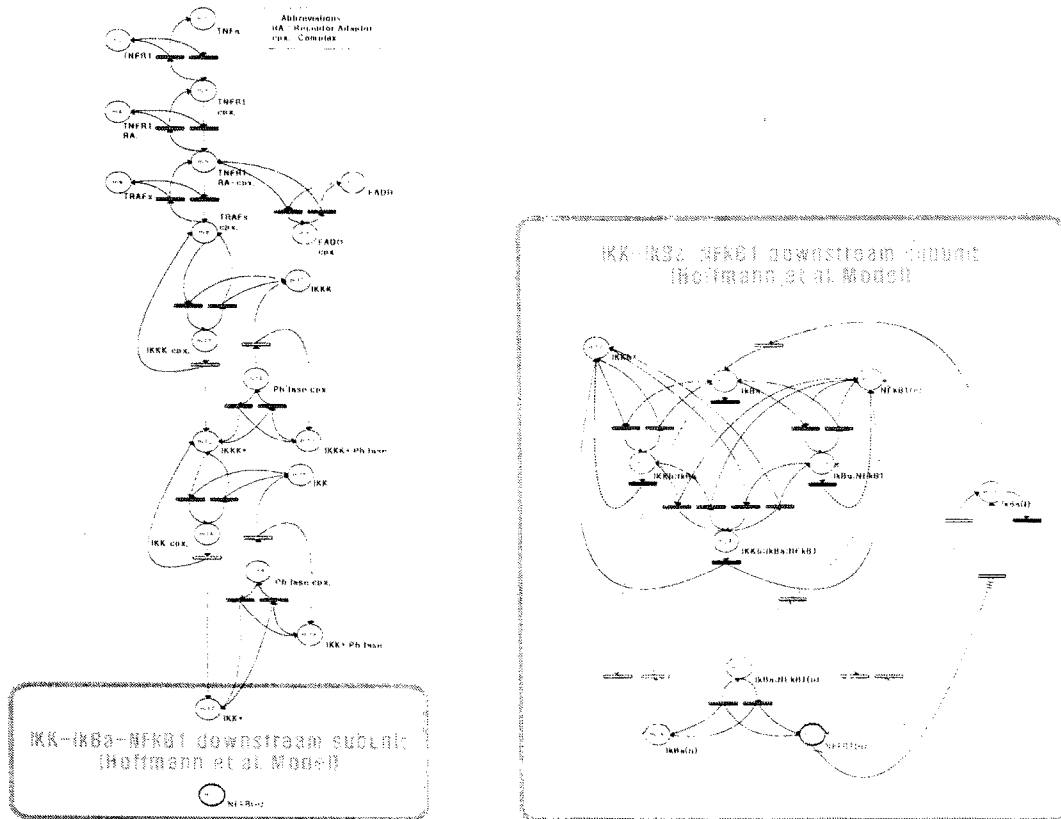


그림 13. NF- $\kappa$ B 조절 시스템(우)과 전체 NF- $\kappa$ B 신호전달경로 모델 구성(좌)

## 8. 초기 농도 및 파라미터 결정

신호전달경로 ODE 모델은 기본적으로 반응 상수와 초기 조건에 매우 민감한 영향을 받는 비선형 시스템이다. 따라서 간세포 내에서 일어나는 생화학 반응을 정확히 모델링하기 위해서는 실제 *in vitro* 실험을 통한 간세포 특성적인 파라미터 및 주요 단백질의 초기 농도 측정이 필수적이다.

제1세부과제팀의 *in vitro* 실험을 통하여 다음과 같은 모델 파라미터 값들을 결정하였다.

구분	TNF $\alpha$	TRAF2	NIK	IKK $\alpha$	IKK $\beta$
HepG2	0.00039	0.33389	0.04522	0.02417	0.02278
HepG2.2.15	0.00039	0.73653	0.06714	0.07243	0.05999

표 14 *in vitro* 실험을 통한 모델 초기 조건 측정 (단위  $\mu$ M)

기타 모델 파라미터는 *in vitro* 실험값과 literature를 참고하여 결정하였다.



Transcription factor NF- $\kappa$ B(n)가 갖는 다양한 출력 패턴은 유전자의 선택적 발현과 관련이 있다고 추정된다. 즉, NF- $\kappa$ B의 binding 시간 및 강도에 따라 서로 다른 발현율을 갖는 유전자들을 조절하는 메커니즘이 바로 negative feedback loop를 통해 구현되는 것이다.

서브시스템별 시뮬레이션 분석 결과를 종합하여 볼 때, 세포 외부에서 가해지는 TNF $\alpha$  신호에 대하여, receptor 서브시스템 및 kinase cascade 서브시스템의 구성 조합이 NF- $\kappa$ B 조절 시스템에 서로 다른 입력 IKK\* 입력 패턴을 유도하여 선택적으로 유전자 발현을 조절함을 유추해 볼 수 있다. 특히 kinase cascade 시뮬레이션에서 살펴보았듯이 kinase 농도 및 phosphatase 비율 등 모델 파라미터 변화가 서브시스템의 출력 패턴에 영향을 준다는 사실과, *in vitro* 실험을 통해 HBV 감염이 신호전달 경로 구성 단백질의 농도에 변화를 준다는 사실은 바이러스에 의한 신호전달 경로의 교란이 NF- $\kappa$ B activity의 현저한 변화를 일으킬 수 있음을 시사해주고 있다. 이에 앞서 구성한 신호전달 경로 모델을 기반으로 HBV 영향에 관한 모델링 및 시뮬레이션을 진행하였다. 또한 이와는 별도로 신호전달 경로 상의 주요 단백질을 억제하였을 때 NF- $\kappa$ B의 activation 변화를 측정하는 시뮬레이션도 진행하였다.

#### 나. inhibition 효과 시뮬레이션

신호전달 경로에 대한 *in silico* 모델은 생명 현상의 단순한 재현뿐만 아니라, 필요에 따라 실험자가 임의로 구성된 환경 하에서 신호전달 경로에 발생하는 변화에 대한 예측을 가능하게 한다. 예를 들어 약물이나 바이러스 또는 독성 물질에 의해 단백질이 inhibition 또는 activation되었을 때 신호전달 경로에 나타나는 동적 특성 변화를 *in silico* 시뮬레이션을 통해 어느 정도 예측해 볼 수 있다. 이는 특히 신약 개발 시, 최적의 투약 효과를 낼 수 있는 target 단백질 후보군을 *in silico* 시뮬레이션을 통해 미리 선정하여, 엄청난 비용과 시간이 드는 *in vitro/in vivo* 실험을 줄일 수 있게 해준다는 점에서 경제적으로 큰 의미를 가진다.

본 연구에서는 NF- $\kappa$ B 신호전달 경로를 inhibition하는 최적의 단백질 노드를 찾기 위하여 inhibition 효과 시뮬레이션을 시행하였다. 이를 위해 앞서 구성한 기본 신호전달 경로 모델을 바탕으로 kinase cascade의 주요 단백질 TRAFs complex와 IKKK, IKK에 대하여 dominant negative competitor type의 inhibitor를 모델링 하였다.

특정 단백질을 inhibition 하여 얻은 NF- $\kappa$ B inhibition 효과를 측정하기 위해 Area Under Inhibition Curve(AUIC)를 다음과 같이 정의한다.

$$AUIC = \left[ 1 - \frac{\text{mean}(NF\kappa B_n)}{\text{(steady state } NF\kappa B_n \text{ in the absence of inhibition)}} \right] \times 100 (\%)$$

즉, AUIC는 특정 단백질을 일정 시간 inhibition 하였을 때, 그 시간 동안 정상 상태보다 감소하게 되는 NF- $\kappa$ B activity의 비율을 나타낸다.

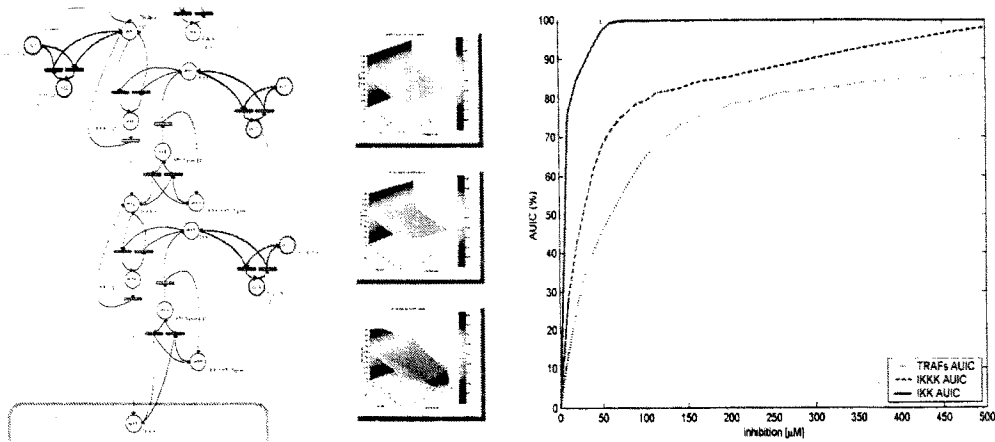


그림 16. 신호전달경로 상 단백질의 inhibition 모델(좌) 및 효과 측정AUIC 분석(우)

시뮬레이션 결과 IKK를 inhibition하였을 때, NF- $\kappa$ B가 최대로 inhibition되는 것을 볼 수 있다. 특히 IKK inhibition은 IKK나 TRAFs에 비하여 상대적으로 빠른 시간 내에 100% inhibition에 도달함을 그래프를 통해 살펴볼 수 있다. 이는 치료 목적으로 NF- $\kappa$ B activation을 감소시킬 필요가 있을 때, IKK를 inhibition하는 molecule을 약물로 설계하는 것이 효과적임을 시사해 준다.

#### 다. HBV 감염 효과 시뮬레이션

HBV 감염으로 인한 NF- $\kappa$ B activation의 증가를 *in silico* 상에서 모델링하기 위하여 기본 모델을 다음과 같이 확장하였다.

(1) 제1세부과제팀의 *in vitro* 실험 결과를 통하여, HBV가 간세포 내 phosphorylated IKK를 증가시키는 것을 확인하였다. 여기에 HBV가 NIK를 통하여 NF- $\kappa$ B activity가 현저히 증가시킨다는 보고를 바탕으로, 모델에 바이러스로부터 기인하는, NIK를 통한 IKK activation 경로를 추가하였다. 또한 literature를 근거로 TRAFs complex를 통한 NIK 억제 작용도 모델링 하였다.

(2) 제1세부과제팀의 HBV 감염 세포에 대한 TNF $\alpha$  반응 실험은 HepG2 및 HepG2.2.15 세포를 대상으로 이루어졌다. 정상 간암 세포인 HepG2 세포와 간암 세포인 HepG2.2.15 세포를 정확히 모델링하기 위하여, *in vitro* 실험에 근거, HepG2 및 HepG2.2.15의 미세한 IKK activation 경로를 추가하였다.

(3) *in vitro* 실험을 통해 밝혀진 HepG2와 HepG2.2.15의 서로 다른 초기 농도 구성을 모델에 적용하였다.

즉, 간암으로 기인한 미세 IKK activation은 임의의 단백질 X로 인한 IKK phosphorylation으로, HBV로 기인한 NIK 및 IKK activation은 임의의 단백질 Y로 인한 NIK phosphorylation과 phosphorylated NIK로 인한 IKK phosphorylation으로 모델링하였다.

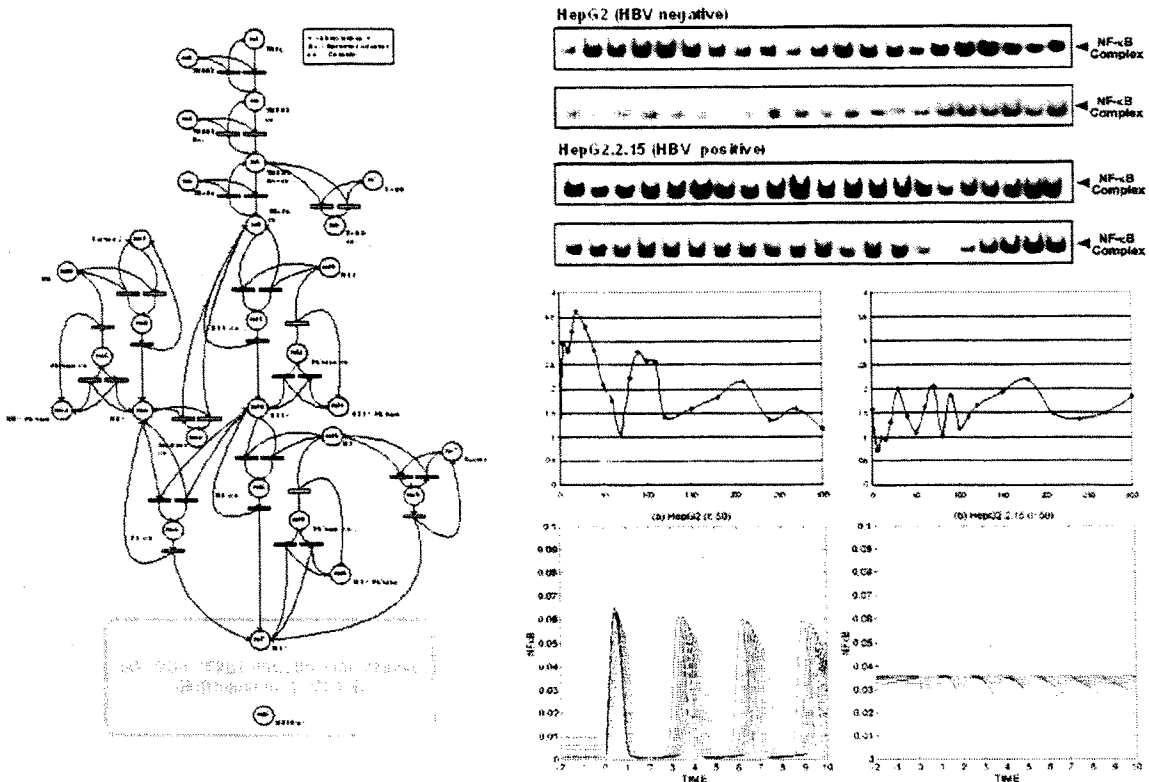


그림 17. HepG2 및 HepG2.2.15 모델과 *in vitro/in silico* 실험 (A) TNF $\alpha$  투여에 의한 NF- $\kappa$ B 반응 EMSA 이미지 (B) EMSA 결과를 정량적으로 표현한 그래프 (C) *in silico* 시뮬레이션 결과. 서로 다른 phosphatase 비율에 대해 시뮬레이션을 진행하였다. (Phosphatase 비율이 적을 수록 진한 선으로 표시)

*in vitro* 실험은 TNF $\alpha$  투여 후 5시간 동안 세포 내 NF- $\kappa$ B activity를 일정 시간 간격으로 EMSA 분석하는 방식으로 이루어졌다. 그 결과 정상 세포인 HepG2에서는 NF- $\kappa$ B activity가 진폭 및 주기가 큰 뚜렷한 oscillation을 띄는 형태로 나타났다. 반면에 HepG2.2.15에서는 진폭 및 주기가 적은 oscillation이 발생함을 볼 수 있다. 특히 TNF $\alpha$  투여 초기 NF- $\kappa$ B activity가 오히려 감소하는 undershoot 현상이 관측됨을 살펴볼 수 있다.

*in silico* 시뮬레이션을 통해서도 이와 유사한 패턴의 NF- $\kappa$ B activity가 나타났다. HepG2의 경우 크고 뚜렷한 진폭 및 주기를 갖는 oscillation 형태의 출력이 보이며, 반면에 HepG2.2.15의 경우 작고 좁은 진폭과 주기를 갖는 oscillation 출력이 나타나고 있다. 이는 NF- $\kappa$ B 조절 시스템에 들어오는 입력 phosphorylated IKK가 HepG2.2.15 세포에서는 HBV의 영향으로 정상 세포보다 많이 activation 되어 NF- $\kappa$ B 조절 시스템의 출력 형태가 변화될 수 있음을 시사한다.

또한, HBV 감염 세포인 HepG2.2.15에서 특징적으로 발견되는 undershoot 현상은 TNF $\alpha$  투여 초기 신호전달경로 상에 있는 TRAFs complex의 증가가 바이러스에 의해 활성화된 NIK를 억제하고, 이로 인해 IKK activation이 감소하는 동적 변화에 기인하는 것으로 보인다. 이는 *in silico* 모델을 통하여, 감염 세포의 신호전달경로의 동역학 특성 변화에 대한 의미 있는 가설을 설정하고, 시뮬레이션과 *in vitro* 실험을 통하여 이를 검증할 수 있음을 보여주는 hypothesis driven research의 좋은 사례라 할 수 있다.

## 10. HCV 관련 TGFβ 신호전달경로에 대한 ODE 모델 표현

### 가. TGFβ 신호 전달 경로

TGFβ는 receptor 1, 2와 상호 작용하여 TGFβ/TGFβ receptor 1/2 complex를 이루며 이것이 R-Smad인 Smad2, 3를 phosphorylation시킨다. 다음에는 이것이 Co-Smad-Cyto Smad 4와 반응하여 p-R-Smad/Co-Smad Complex-Cyto를 이루게 된다. p-R-Smad/Co-Smad Complex-Cyto이 세포질에서 핵 속으로 이동하여 p-R-Smad/Co-Smad complex-nucleus가 된다. 다음에는 p-R-Smad/Co-Smad complex-nucleus가 target 유전자의 promoter binding하여 target 유전자의 transcription/translation이 이루어지고, R-Smad와 Co-Smad는 다시 세포질로 translocation 되는 과정을 거친다. 이렇게 하여 생겨난 Smad 7이 Smad 2,3의 competitive inhibitor로 작용하게 된다. 이와 같은 과정으로 거쳐 TGFβ 신호 전달 경로가 완성된다.

### 나. ODE 모델의 구성

앞에서 말한 HBV에서 모델을 구성했던 것과 비슷한 방법으로 HCV 네트워크의 전체 신호전달 경로 *in silico* 모델을 완성하였다. 그 다음 이 모델을 MATLAB을 이용하여 시뮬레이션 작업을 수행하고, 그 결과를 제2세부과제 팀과 연계하여 필요한 실험을 진행하고 그 결과를 반영하여 더 나은 모델을 만드는 작업을 수행하였다.

### 다. 초기 농도 및 파라미터 결정

위와 같은 모델은 반응 상수와 초기 농도에 따라 서로 다른 시뮬레이션 결과를 보인다. 따라서 적절한 파라미터 값과 초기 값을 정하는 것이 매우 중요하다. 본 세부연구과제에서는 제2세부과제팀과의 협력을 통하여 다음과 같은 초기 농도 값을 정하였다. 그 외에 모델에서 사용된 파라미터 값들은 literature를 참고하였다.

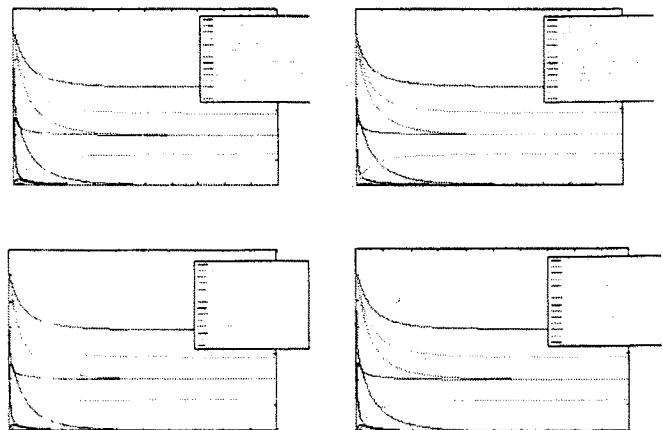
TGFβ	TβR-2	TβR-1	R-Smads	Co-Smad
0.2	0.3	0.3	0.3	0.3

*in vitro* 실험을 통한 모델 초기 조건 측정 (단위 nM)

### 라. *in silico* 시뮬레이션

위의 모델을 MATLAB으로 구성하여 실험한 결과 옆과 같은 결과를 얻을 수 있었다.

이 결과는 모델에 사용된 파라미터 값을 다양하게 변화시켜 가면서 실험한 결과 그림들이다. 파라미터 값이 바뀔 때마다 시뮬레이션 결과가 바뀌지만, 대부분의 경우 모든 물질이 평형 상태에 이르는 것으로 알아볼 수 있었다. 현재 제2세부과제팀과의 긴밀한 연계를 통하여 위의 모델을 개선해나가는 작업을 진행하고 있으며, HBV 모델에서 진



HCV 관련 신호전달경로 모델 *in silico* 시뮬레이션 결과



행했던 것과 유사한 방향으로 실험을 진행될 것이다.

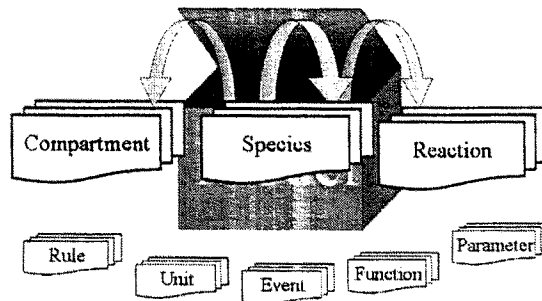
## 11. 관련 데이터베이스 설계

현재 HBV, HCV 모델과 같은 생물학 네트워크 모델을 표현할 수 있는 다양한 프로그램이 존재한다. 이 때 각 프로그램에서 네트워크 모델을 표현하는데에 서로 다른 방법을 이용한다면 한 프로그램의 결과물을 다른 프로그램에서는 사용할 수 없을 것이다. 따라서 생물학 모델을 표현하는 표준 방법이 필요하게 되었다. 이와 같은 목적으로 개발 된 것이 SBML이다. SBML은 Systems Biology Markup Language의 약자로써, biochemical network 모델을 표현하는 표준 방법을 제공한다.

가. SBML은 다음과 같은 장점을 가지고 있다.

- (1) XML을 기반으로 만들어져 있어서 확장성이 우수하다.
- (2) Biochemical network을 표현하는 표준을 제시함으로써, 다양한 소프트웨어의 결과물을 서로 다른 소프트웨어에서 사용하는 것이 쉽게 이뤄질 수 있다.

나. SBML은 다음 그림과 같은 여러 구성 요소로 이루어져 있다.



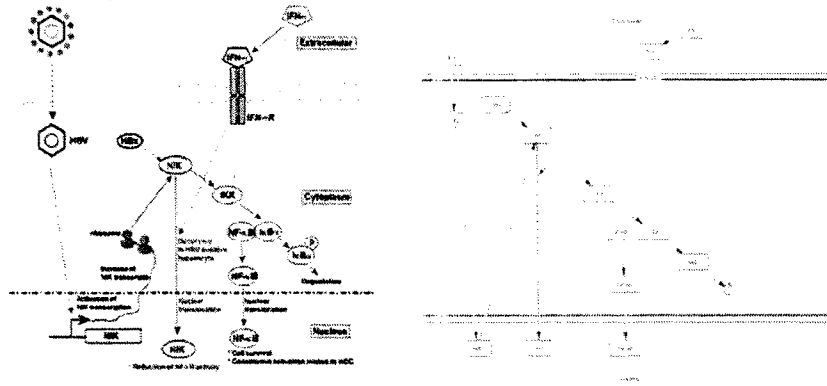
- (1) Unit : SBML에서 사용되는 각종 원소들의 단위 (몰, mg 등)
- (2) Compartment : species가 위치하는 볼륨을 나타냄
- (3) Species : network 내의 proton, calcium와 같은 간단한 ion, glucose, ATP과 같은 간단한 molecule 그리고 RNA, protein과 같은 큰 molecule을 표현함.
- (4) Parameters : network을 표현하는 데 필요한 값들을 선언
- (5) Rules : network 내의 species들은 상태가 변할 수 있다. 이때 제약 조건이 있다면 이것을 선언할 수 있는데, 이런 제약조건을 Rule이라고 한다.
- (6) Reaction : species들끼리 영향을 주고 받는 관계를 나타낸다. 즉 어떤 species의 양이 변하거나 compartment를 옮기거나 species의 상태가 바뀌는 관계를 표현한다.
- (7) 이외 Function, Event와 같은 요소가 존재한다.

SBML은 현재 계속 개선되고 있다. 그리고 50가지가 넘는 다양한 소프트웨어에서 biochemical network을 표현하는데 SBML을 사용하고 있으며, 앞으로는 더 많은 소프트웨어에서 SBML이 사용

될 것으로 예상된다.

본 연구에서는 제1세부과제와 제2세부과제의 HBV, HCV 네트워크를 design 한 후, SBML tool인 Cell-designer를 이용하여 그림을 그려보았다.

다음은 제1세부과제의 network model과 그것을 SBML로 표현한 그림이다.



제1세부 과제의 HBV 네트워크 그림과 SBML로 표현한 그림

위와 같이 그림을 그리면, Cell-designer와 같은 소프트웨어에서는 다음과 같은 SBML을 생성하여 준다.

```
<?xml version="1.0" encoding="UTF-8" ?> <!-- Created by SBML API 2.0(a15.2) -->
<sbml xmlns="http://www.sbml.org/sbml/level1" level="1" version="2" xmlns:html="http://www.w3.org/1999/xhtml"
xmlns:cellDesigner="http://www.sbml.org/2001/ns/celldesigner">
<model>
<annotation>
<cellDesigner:modeVersion>2.2</cellDesigner:modeVersion>
<cellDesigner:modeIDisplay sizeX="1000" sizeY="700" />
<cellDesigner:listOfCompartmentAliases>
<cellDesigner:compartmentAlias id="ca3" compartment="c3">
<cellDesigner:class>SQUARE_CLOSEUP_SOUTH</cellDesigner:class>
<cellDesigner:point x="416.0" y="124.0" />
</cellDesigner:compartmentAlias>
<cellDesigner:compartmentAlias id="ca4" compartment="c4">
<cellDesigner:class>SQUARE_CLOSEUP_NORTH</cellDesigner:class>
<cellDesigner:point x="516.0" y="569.0" />
</cellDesigner:compartmentAlias>
</cellDesigner:listOfCompartmentAliases>
<cellDesigner:listOfSpeciesAliases>
<cellDesigner:speciesAlias id="s1" species="s1">
<cellDesigner:activity>inactive</cellDesigner:activity>
<cellDesigner:bounds x="145.0" y="155.0" w="70.0" h="30.0" />
</cellDesigner:speciesAlias>
<cellDesigner:speciesAlias id="s12" species="s14">
```

### HBV 네트워크 모델의 SBML 명세

다음은 제2세부과제의 결과를 제1세부과제와 마찬가지로 SBML로 표현한 그림이다.

## 12. EST 정보를 이용한 패스웨이 검색 툴 개발

### 가. 연구 목적 및 내용

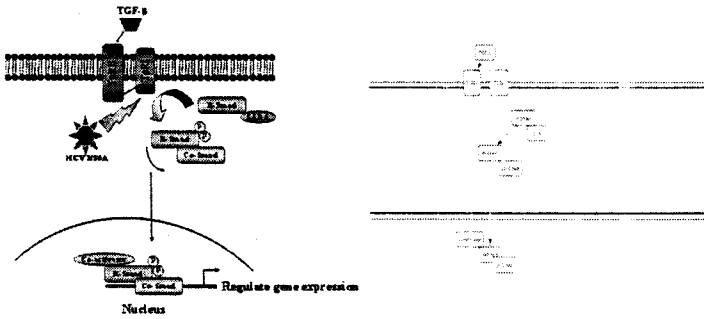
본 연구에서는 KEGG PATHWAY 데이터베이스에서 제공하는 pathway 그래프를, EST data와 융

```

<?xml version="1.0" encoding="UTF-8" ?> <!-- Created by SBML API 2.0(a15.2) -->
<sbml xmlns="http://www.sbml.org/sbml/level1" level="1" version="2" xmlns:html="http://www.w3.org/1999/xhtml"
xmlns:celldesigner="http://www.sbml.org/2001/ns/celldesigner">
<model>
<annotation>
<celldesigner:modeVersion>2.2</celldesigner:modeVersion>
<celldesigner:modeDisplay size="1000" sizeY="700" />
<celldesigner:listOfCompartmentAliases>
<celldesigner:compartmentAlias id="ca3" compartment="ca3">
<celldesigner:class>SQUARE_CLOSEUP_SOUTH</celldesigner:class>
<celldesigner:point x="468.0" y="191.0" />
</celldesigner:compartmentAlias>
<celldesigner:compartmentAlias id="ca4" compartment="ca4">
<celldesigner:class>SQUARE_CLOSEUP_NORTH</celldesigner:class>
<celldesigner:point x="428.0" y="439.0" />
</celldesigner:compartmentAlias>
<celldesigner:listOfCompartmentAliases>
<celldesigner:listOfSpeciesAliases>
<celldesigner:speciesAlias id="s18" species="s17" compartmentAlias="ca3">
<celldesigner:activity>inactive</celldesigner:activity>
<celldesigner:bounds x="283.0" y="156.0" w="50.0" h="50.0" />
</celldesigner:speciesAlias>
<celldesigner:activity>inactive</celldesigner:activity>

```

HCV 네트워크 모델의 SBML 명세



제2세부과제의 HCV 네트워크 그림과 SBML로 표현한 그림

하고, 이를 다른 tissue source의 결과와 비교함으로써, tissue source에서의 pathway 그래프 특성을 살펴볼 수 있다. 다음 그림은 이 과정을 진행하기 위해 데이터베이스를 검색하는 과정을 나타내고 있다.

KEGG PATHWAY 데이터베이스 내 MAPK signalling pathway 상의 단백질들이 정상 간세포 (liver)와 간암 세포(hepatocellular carcinoma)에서 어떤 발현 패턴을 보이는지, 위의 그림과 같이 EST 개수를 검색한 결과는 다음과 같다. 아래 그래프에서 노드가 가지는 숫자는 총 EST 개수와 지정된 tissue source로 검색된 EST 개수를 나타내고 있다

합하여, 그래프 상의 각 노드에 해당하는 단백질들의 발현 패턴을 tissue source별로 선택, 비교·분석할 수 있는 소프트웨어 도구를 개발하였다.

KEGG PATHWAY 데이터베이스에는 현재까지 알려진 metabolic pathway, regulatory pathway, molecular complex 등이 그래프 형태로 저장되어 있다. 현재의 데이터베이스는 생물체 (organism) 별로 pathway 그래프를 제공하며, 노드에 해당하는 단백질에 관련된 외부 데이터베이스 링크를 제공하고 있다.

본 연구에서는 각 단백질 노드에 관련된 외부 데이터베이스를 자동으로 탐색하여, 해당 단백질과 관련된 EST sequence를 조사함으로써, 각 단백질의 발현 패턴을 EST의 개수로 정량화하여 보여줄 수 있도록 하였다. 특히, 특정 tissue source를 선택하여 EST를 계수

## 제 4 장 목표달성도 및 관련분야에의 기여도

### 1절. 연구개발의 최종목표

HBV 및 HCV의 viral protein과 hepatocyte내의 protein 네트워크의 상호작용을 분석하여 180 가지 이상의 상호작용하는 protein을 발굴 (서울대 정구홍, 한림대 황순봉)한다. 이 protein이 속한 간세포내 protein 네트워크 연구를 협동 연구기관의 database system (서울대 박근수) 을 통해 질병 (간염, 간경변, 간암등을 포함하는 간질환) 유도 기작을 수식적으로 예측할 수 있는 모델 4가지 이상을 제시 (서울대 정구홍, 한림대 황순봉, 서울대 박근수)한다. 이 protein들의 이 연구 결과를 토대로 journal of biological chemistry 및 gastroenterology등의 저널에 5편 이상 게재한다.

또한, 본 연구팀에서는 간염 바이러스와 간세포내 단백질 네트워크와의 상호 작용이라는 전체 과제에서, 실험을 통한 데이터와 이의 해석을 저장하는 데이터베이스를 구축 (서울대 박근수)하는 한편 이 데이터베이스를 연구자가 쉽게 사용할 수 있고 직관적으로 이해할 수 있는 형태로 제공하기 위한 인터페이스를 개발 (서울대 박근수)한다.

#### 1. 제 1차년도 (2003. 6. 1 - 2004. 5. 31)

HBV 및 HCV는 다양한 viral protein을 production한다. 이 protein들은 간세포내에서 다양한 protein과 서로 상호작용하며, 이것은 이 세포내에서 다양한 생리학적 변화를 일으킨다. 이 변화를 일으키는 원동력과 또 이것에 의해 변화되는 protein들의 활성 및 발현 수준 변화를 분석하여 HBV 및 HCV에 의한 간세포에 영향을 전체적으로 분석할 수 있는 결과를 도출하는 것 (직접 또는 간접적으로 상호작용하는 protein 180가지 이상의 도출)이다 (서울대 정구홍, 한림대 황순봉). 또, 이와 같이 도출된 데이터를 저장하고 관리할 수 있는 system을 구축 (서울대 박근수) 한다.

#### 가. HBV 및 HCV의 system biology 연구를 위한 data 창출

HBV 및 HCV viral protein들은 간세포내에서 다양한 protein들과 interaction하여 세포내 생리적 변화를 일으킨다. 이 단계에서는 변화를 일으키는 원동력 및 이에 의한 변화를 분석을 시도한다.

이 분석에서는 본 연구팀은 같은 기술을 이용한다.

##### (1) microarray 기술

(가) 각 input (HBV 및 HCV protein들)에 의한 세포내 transcriptional regulation 분석

##### (2) 2DE 기술

(가) 각 input(HBV 및 HCV protein들)에 의한 protein production regulation 분석

(나) 각 input(HBV 및 HCV protein들)에 의한 protein modification 분석

(3) affinity chromatography기술

(가) in vitro system에서의 HBV 및 HCV protein들과 hepatocyte내 protein들과의 protein-protein binding의 분석

(4) yeast two hybrid 기술

(가) in vivo system에서의 HBV 및 HCV protein들과 hepatocyte내 protein들과의 protein-protein binding의 분석

나. HBV 및 HCV의 연구 결과를 통해 창출된 data를 저장하는 데이터베이스 및 사용자 인터페이스 설계

위 연구 결과의 활용 및 분석을 위해 computer science section에서는 실험 결과를 저장할 데이터베이스 및 사용자 인터페이스의 설계를 진행한다.

2. 제 2차년도 (2004. 6. 1 - 2005. 5. 31)

HBV 및 HCV에 의한 간세포의 영향을 제 1차년도에서 분석한 결과를 in silico system에서 database 확립 및 systematic analysis를 통해 얻어진 결과를 분석 (서울대 박근수) 하여, 정량적으로 분석할 타겟을 설정 (서울대 정구홍, 한림대 황순봉, 서울대 박근수) 한다. 설정된 타겟의 특성을 HBV 및 HCV에 의한 질병의 유도과 연관지어 정량적으로 분석하는 것 (서울대 정구홍, 한림대 황순봉) (4가지 이상의 systematic flow에 속한 protein들의 분석, 약 20-30가지 protein들이 2차년도의 주요 목표이다.

가. HBV 및 HCV의 연구를 통해 얻어진 data 중 system을 이루는 data의 선정 및 정량적 분석

제 1차년도의 연구 결과 중 data 분석을 통하여 세포내에서 system을 이루는 protein 및 gene group를 설정하고, 이에 대한 정량적 분석을 시도한다. 각 정량적 분석은 전체 systematic flow를 이루는 block을 형성하게 된다.

나. HBV 및 HCV의 연구를 통해 얻어진 정량적 data의 분석을 통한 모델화 과정

각각의 정량적 분석의 결과인 block 들이 완성되게 되면, 이 block 들을 종합하게 된다. 이 과정을 통해 간 질환과 연관된 모델을 세우게 된다.

이 모델은 간염바이러스와 간질환과의 연관관계 모델이지만, 실제 다른 pathway의 도입이 있으므로 보편화된 모델을 도출할 수 있을 것으로 기대된다.

또한, 이 모델은 간염 바이러스 인자의 간세포에 대한 influx와 이에 따른 세포의 efflux의 관계를 분석한 것이나, 내부의 인자들 (pathway 관련 protein)에 대한 각각의 요인을 분석하게 되므로 다

큰 모델과의 정보를 교환할 수 있을 것으로 기대된다.

다. 데이터베이스 및 사용자 인터페이스의 구현

Computer science section에서는 또한 1차년도에 설계된 따라 데이터베이스 및 사용자 인터페이스의 prototype을 개발한다. 이 prototype은 지속적인 feedback을 통해서 발전하게 된다.

## 2절. 연차별 연구개발 목표 및 내용

구분	연구개발 목표	연구개발 내용 및 범위
1차년도 (2003)	<ul style="list-style-type: none"> <li>● HBV 및 HCV viral protein과 상호 작용하는 protein 발굴</li> <li>● HCV protein 및 replicon을 이용한 cytokine network 연구</li> <li>● 데이터베이스 및 인터페이스의 설계*</li> </ul>	<ul style="list-style-type: none"> <li>● HBV (polymerase, core protein, surface protein, HBx protein) 및 HCV ( core protein, E1/E2 envelope protein, NS2 protein, NS3 protein, S4A/4B protein, NS5A/5B protein) viral protein expression (10 가지 protein이상)</li> <li>● HBV 및 HCV viral protein과 interaction하는 protein 탐색 (180 가지 이상 탐색)</li> <li>● HCV viral proteins 과 replicon의 TNF alpha 및 TGF beta signaling 조절연구</li> <li>● 각 연구로부터 얻어진 결과를 위한 데이터베이스 및 인터페이스 설계*</li> </ul>
2차년도 (2004)	<ul style="list-style-type: none"> <li>● 1차년도에 탐색된 HBV 및 HCV viral protein과 interaction하는 protein 군을 선별하여 정량적 분석을 시도</li> <li>● HBV 및 HCV-associated marker 단백질 탐색</li> <li>● 위 연구로부터 도출된 실험 결과 분석 및 단백질 네트워크 모델 개발로 Prototype 완성*</li> </ul>	<ul style="list-style-type: none"> <li>● 1차년도의 결과 분석을 통해 systematic flow를 완성할 수 있는 타겟 설정 (4 가지 pathway 이상 설정)*</li> <li>● 설정된 타겟 protein들과 HBV 각 protein들과의 상호작용을 정량적으로 분석 (in vivo system &amp; in vitro system)</li> <li>● HBV 연구실의 선행 연구인 TNF-b, TRAF2, NIK, SePP와 HBV 각 protein들과의 상호작용관계를 정량적으로 분석 (4가지 protein 분석)</li> <li>● HCC 임상 sample의 분석을 통해 marker 유전자 탐색 (50가지 이상)</li> <li>● 위 연구 결과를 토대로 HBV 및 HCV에 의해 영향을 받는 또는 영향을 주는 circuit을 4 종류 이상 완성</li> <li>● 위 연구 결과의 분석과 저장을 위한 데이터베이스 및 인터페이스의 prototype 완성*</li> </ul>

### 3절. 계획대비 달성도

번호	세부연구목표	달성내용	달성도(%)
1	○ HBV (polymerase, core protein, surface protein, HBx protein) 및 HCV (core protein, E1/E2 envelope protein, NS2 protein, NS3 protein, S4A/4B protein, NS5A/5B protein) viral protein expression	Viral protein 10 가지 protein이상 expression	100
2	○ HBV 및 HCV viral protein과 interaction하는 protein 탐색	관계된 protein 180 가지 이상 탐색	100
3	○ HCV viral proteins 과 replicon의 TNF alpha 및 TGF beta signaling 조절연구	NF-kappaB 활성화도 변화, JNK 활성화도 변화, apoptosis 영향 분석	100
4	○ 각 연구로부터 얻어진 결과를 위한 데이터베이스 및 인터페이스 설계*	데이터베이스 및 인터페이스 설계*	100
5	○ 1차년도에의 결과 분석을 통해 systematic flow를 완성할 수 있는 타겟 설정	4 가지 pathway 이상 설정* (NF-κB, IFN-g, TGF-b, SePP pathway 설정)	100
6	○ 설정된 타겟 protein들과 HBV 각 protein들과의 상호작용을 정량적으로 분석 (in vivo system & in vitro system)	20-30가지의 protein들 분석	100
7	○ HBV 연구실의 선행 연구인 TNF-b, TRAF2, NIK, SePP와 HBV 각 protein들과의 상호작용관계를 정량적으로 분석	4가지 protein 분석	100
8	○ HCC 임상 sample의 분석을 통해 marker 유전자 탐색	marker 50 종류 이상	100
9	○ 위 연구 결과를 토대로 HBV 및 HCV에 의해 영향을 받는 또는 영향을 주는 circuit을 3 종류 이상 완성	3 종류 이상 완성	100
10	○ 위 연구 결과의 분석과 저장을 위한 데이터베이스 및 인터페이스의 prototype 완성*	논문 1편 기고	100

연차	변경 전 (선정연도 연구계획서)			변경 후			비고
	연구목표	연구내용	연구범위	연구목표	연구내용	연구범위	
							승인여부

## 제 5 장 연구개발결과의 활용계획

### 1절. 추가 연구의 필요성

1. 본 연구과제는 1단계에 속하는 3년 계획으로 짜여진 연구계획을 갖추고 시작되었다. 따라서, 실제 연구가 마무리 되지 않은 상황에서 종료되는 결과가 되었다.

당초 본 연구 계획은

1년차: *in silico* modeling을 위한 target 탐색

2년차: 탐색된 target과 관련된 parameter estimation; prototype의 *in silico* modeling 구축

3년차: *in silico* modeling의 실험을 통한 검증 및 상호 조정 그리고 적용이었다.

2. 본 연구팀은 실제 진행 속도보다 빠르게, *in silico* modeling의 parameter estimation을 시작하였고, 또한 실제의 *in vivo* 문제를 해결하기 위한 적용단계를 수행하는 과정이었다. 따라서, 2년차에 미쳐 끝내지 못한 제 3년차의 나머지에 해당하는 부분이 추가로 연구 되어야 할 대상이다. 실제 *in silico* modeling의 결과는 실제 실험과 연동되어 변환되고 수정되어야 하며, 실제 문제를 풀 수 있는 모델이어야 하기 때문에 추가 연구가 더 요구되어진다.

추가 연구 과제	연구 내용
○ <i>in silico</i> 모델 구조의 정립	Literature 및 <i>in vitro/in vivo</i> 실험 결과 분석을 통하여, 구성 단백질 간의 상호 작용 관계 구조를 규명하고, 이를 통해 보다 정확한 <i>in silico</i> 모델을 정립한다.
○ <i>in silico</i> 상 주요 파라미터 선별 기법 개발	Sensitivity analysis 등을 공학적 기법을 활용하여 모델의 동역학적 변화에 가장 큰 영향을 주는 주요 파라미터를 선별하고, 이에 대한 <i>in vitro</i> 실험을 통해 모델을 보완한다.
○ <i>in silico</i> 모델의 동역학 특성 분석 기법 개발	비선형 시스템 분석에 사용되는 phase plane analysis, bifurcation analysis 등을 활용하여 모델이 가진 동역학적 특성을 발견한다.

### 2절. 타 연구에의 응용

1. Hoffman model (2004년 science)에 따른 NF-κB inhibitor들인 drug에 대한 inhibitory effect는 검사는 IKK-p 이후의 pathway에 그 초점이 맞추어져 있었다. 하지만, 본 연구과제를 통해 얻은 NF-κB activation 모델은 HBV infection시를 적용했을 뿐만 아니라, IKK-p의 up stream 의 모델을 포함하고 있다. 이 모델은 HBV infection시의 NF-κB activation저해를 위한 drug 개발과 기타의 IKK-p up stream pathway에 대한 drug effect를 검증하는데 중요한 모델로 사용될 수 있다. 즉, 기존의 pharmacokinetics 연구의



모델로 응용될 수 있는 연구 결과 이다.

2. 본 연구팀은 시스템 생물학 연구사업을 통해 전혀 활용된 바 없는 세포내 신호전달과정과 C형 간염바이러스간의 연계성을 찾아내어 좀 더 명확한 바이러스성 간암의 유발 mechanism을 밝히고자 한다. 이러한 연구가 성공적으로 진행된다면 C형 간염바이러스와 유사한 RNA 바이러스에 이러한 결과를 적용시켜 바이러스를 연구하는 학자들에게 암유발 mechanism에 대한 좀더 정확한 기전을 설명할 수 있으며 또한 세포생물학자들에게 있어서는 직접적인 실험을 수행하지 않고도 세포내 신호전달과정을 computer로 예상할 수 있어 타연구에 많은 응용이 이루어질 것으로 생각된다.

3. 위와 같은 연구를 통해 구성된 모델은 다음과 같은 연구에 응용 가능하다.

가. 간염 관련 *in silico* 약물 target 선정

나. 간염 치료 약물의 효과 및 부작용(side effect)에 대한 *in silico* 평가, 검증

다. 간질환 환자의 임상 데이터 적용 및 *in silico* 예측

라. 환자 임상 데이터에 기반한 최적의 치료 전략 시뮬레이션

### 3절. 기업화 추진방안

1. 본 연구팀에서 제안되어진 NF- $\kappa$ B activation 모델은 아직 정확한 검증 및 특정 문제에 대한 적용을 할 수 있는지에 대한 검증이 이루어지지 않았다. 실제 제약회사등에서 사용할 수 있는 *in silico* simulation model을 구축하기 위해서는 본과제의 실험 data와 다른 세부 과제의 *in silico* 모델과의 상호 연동한 검증 및 추가 연구가 요구된다.

2. 현재의 실험 결과로는 산업체가 참여할 시기는 아니지만 추가연구의 수행을 통해 많은 data가 축적된다면 향후 산업체가 참여하여 공동 연구를 수행하고 systemic한 접근 방식을 통해 바이러스성 간염 및 간암의 발병 경로가 예측 가능하며 이를 통해 간염 및 간암에 대한 제어가 가능함으로 치료후보물질이 도출될 수 있으리라 사료된다.

## 제 6 장 연구개발과정에서 수집한 해외과학기술정보

### 1절. 해외 유용 웹사이트

▶ (HBV 관련) TNF $\alpha$ 신호전달경로 모델링 연구 그룹		
Wolkenhauer group, 독일	<a href="http://www.sbi.uni-rostock.de/default.htm">http://www.sbi.uni-rostock.de/default.htm</a>	
MPI, 독일	<a href="http://www.mpi-magdeburg.mpg.de/en/research/projects/1010/1022/1016">http://www.mpi-magdeburg.mpg.de/en/research/projects/1010/1022/1016</a>	
Universität Stuttgart, 독일	<a href="http://www.sysbio.de/projects/tnf/StudentProjects.shtml">http://www.sysbio.de/projects/tnf/StudentProjects.shtml</a>	
▶ 모델 repository		
DOQCS	<a href="http://doqcs.ncbs.res.in">http://doqcs.ncbs.res.in</a>	주요 모델 및 파라미터, reference 데이터 베이스
SBML model repository	<a href="http://sbml.org/models/">http://sbml.org/models/</a>	각종 모델의 SBML 명세 파일 제공
SBML Simple ODE Solver oS	<a href="http://www.tbi.univie.ac.at/~raim/">http://www.tbi.univie.ac.at/~raim/</a>	ODE 시뮬레이터 및 MAPK 관련 모델 제공

### 2절. 연구 과정에서 수집한 참고 문헌

저자	저널	년도	제목
A. Kremling	-	2002	A benchmark for problems in reverse engineering, parameter estimation and model discrimination for cellular systems
Bentele et al.	JCB	2004	Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis
Bhalla & Iyengar	Science	2002	MAP Kinase Phosphatase As a Locus of Flexibility in a Mitogen-Activated Protein Kinase Signaling Network
G.Nelson et al.	J.Cell Science	2002	Multi-parameter analysis of the kinetics of NF- $\kappa$ B signalling and transcription in single living cells
Gardner et al.	Science	2003	Inferring Genetic Networks and Identifying Compound Mode of Action via Expression Profiling
Ghoah & Karin	Cell	2002	Missing Pieces in the NF- $\kappa$ B Puzzle
Hayen & Ghosh	Genes & Dev.	2004	Signaling to NF- $\kappa$ B
Hoffmann et al.	Science	2002	The I $\kappa$ B-NF- $\kappa$ B Signaling Module: Temporal Control and Selective Gene Activation
Huang & Ferrell	PNAS	1996	Ultrasensitivity in the mitogen-activated protein kinase cascade
Issacs et al.	PNAS	2003	Prediction and measurement of an autoregulator genetic module
Iyengar et al.	BioEssays	2002	Modeling of signaling networks
J.Ferrell	SKTE	2000	What Do Scaffold Proteins Really Do?
Kalir & Alon	Cell	2004	Using a Quantitative Blueprint to Reprogram the Dynamics of the Flagella Gene Network
Kholodenko	Eur. J. Biochem	2000	Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades
Kholodenko et al.	JBC	1999	Quantification of Short Term Signaling by the Epidermal Growth Factor Receptor
Levchenko et al.	PNAS	2000	Scaffold proteins may biphasically affect the levels of MAPK signaling and reduce its threshold properties
M.Santoro et al.	EMBO	2003	NF- $\kappa$ B and virus infection: who controls whom
Nils Bluthgen	Diploma thesis	2002	Dynamical Models of Signal Transduction and the Influence of Feedback Loop
Nils Bluthgen & Hanspeter Hezel	JTB	2003	How robust are switches in intracellular signaling cascades?
Shapiro et al.	draft	2000	Automatic model generation for signal transduction with applications to MAP-KINASE pathway
Sung & Simon	draft	draft	In Silico Simulation of Inhibitor Drug Effects on NF- $\kappa$ B Pathway Dynamics
Ting & Endy	Science	2002	Decoding NF- $\kappa$ B Signaling
Wiley et al.	TRENDS in C.Bio	2003	Computational modelling of the EGF-receptor system: a paradigm for systems biology
Cho et al.	GenomeResearch	2003	Investigations Into the Analysis and Modeling of the TNF $\alpha$ -Mediated NF- $\kappa$ B Signaling Pathway
Cho et al.	-	2003	Mathematical Modeling of the Influence of RKIP on the ERK Signaling Pathway
Kholodenko et al.	PNAS	2002	Untangling the wires: A strategy to trace functional interactions in signaling and genetic networks

## 제 7 장 참고문헌

- Seeger, C. and W.S. Mason. 2000. Hepatitis B virus biology. *Microbiol Mol Biol Rev* **64**: 51-68.
- Hoffmann, A., A. Levchenko, M.L. Scott, and D. Baltimore. 2002. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **298**: 1241-1245.
- Cho, K.H., S.Y. Shin, H.W. Lee, and O. Wolkenhauer. 2003. Investigations into the analysis and modeling of the TNF alpha-mediated NF-kappa B-signaling pathway. *Genome Res* **13**: 2413-2422.
- Feitelson, M.A. 1999. Hepatitis B virus in hepatocarcinogenesis. *J Cell Physiol* **181**: 188-202.
- Bonizzi, G. and M. Karin. 2004. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends Immunol* **25**: 280-288.
- Bonizzi, G., M. Bebien, D.C. Otero, K.E. Johnson-Vroom, Y. Cao, D. Vu, A.G. Jegga, B.J. Aronow, G. Ghosh, R.C. Rickert, and M. Karin. 2004. Activation of IKKalpha target genes depends on recognition of specific kappaB binding sites by RelB:p52 dimers. *Embo J* **23**: 4202-4210.
- Asthagiri, A.R. and D.A. Lauffenburger. 2001. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnol Prog* **17**: 227-239.
- Kholodenko, B.N. 2000. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur J Biochem* **267**:1583-1588.

주 의

1. 이 보고서는 과학기술부에서 시행한 특정연구개발사업의 연구보고서입니다.
2. 이 보고서 내용을 발표할 때에는 반드시 과학기술부에서 시행한 특정연구개발사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니됩니다.

제 출 문

과학기술부 장관 귀하

본 보고서를 “간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석”과제 (세부과제 “B 형간염바이러스 감염에 의한 간세포 단백질 네트워크 변화 분석”) 의 보고서로 제출합니다.

2005.

주관연구기관명 : 서울대학교

주관연구책임자 : 정구홍

연 구 원 : 박성규

” : 임승외

” : 류현미

보고서 초록

과제관리번호	M1-0309-06-00 01	해당단계 연구기간	2003.6.23 - 2005.3.31	단계 구분	(1) / (3)
연구사업명	중 사업명	국책연구개발사업			
	세부사업명	시스템생물학연구사업			
연구과제명	중 과제명	간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석			
	세부(단위)과제명	B 형간염바이러스 감염에 의한 간세포 단백질 네트워크 변화 분석			
연구책임자	정 구 홍	해당단계 참여연구원수	총 : 4 명 내부 : 4 명 외부 :    명	해당단계 연구비	정부: 147,000 천원 기업:            천원 계:                천원
연구기관명 및 소속부서명			참여기업명		
국제공동연구	상대국명 :	상대국연구기관명 :			
위탁연구	연구기관명 :	연구책임자 :			
요약(연구결과를 중심으로 개조식 500자 이내)				보고서 면수	25
<p>HBV의 genome으로부터 coding되어 translation된 protein들은 직접 또는 간접적으로 세포내의 protein들과 interaction하여 정상적인 세포의 신호전달 체계를 교란시킨다. 이 신호전달 체계의 교란은 HBV에 의한 간질환의 유도에 원인을 제공한다. 지금까지의 연구는 HBV의 특정 단백질에 의해 변화되는 세포내 신호전달 체계에 대한 분석을 시간적 변화에 따른 양상을 분석한 동적인 분석이 아니라, 이번 과제에서는 동적인 실험 분석뿐만 아니라, 이 분석을 바탕으로 in silico model을 완성하고, 이 model을 이용한 simulation을 통해 HBV protein에 의한 동역학적 의미를 밝히고자 노력하였다.</p> <p>이 연구를 통하여 얻은 결과는 아래와 같다.</p> <p>첫째, 총 10가지의 HBV viral protein과 interaction하는 세포내 protein 발굴.</p> <p>둘째, 총 197 가지의 HBV 감염에 의해 변화되는 유전자 발굴.</p> <p>셋째, 총 3가지의 HBV에 의해 활성화되는 세포내 pathway의 정량적 분석</p> <p>이 pathway는 NF-kB activation pathway, Sepp down regulation pathway, 그리고 IFN-g에 의한 NF-kB inhibition pathway이다.</p> <p>이 결과는 in silico simulation을 위한 data로 활용되었다.</p>					
색인어 (각 5개 이상)	한 글	B형 간염 바이러스, 간, 간질환, 시스템 생물학, 네트워크			
	영 어	hepatitis B virus, liver, liver disease, systems biology, network			

## 요 약 문

### I. 제 목

#### B 형 간염바이러스 감염에 의한 간세포 단백질 네트워크 변화 분석

### II. 연구개발의 목적 및 필요성

우리나라의 경우 전 인구의 약 10% (450여만 명)가 HBV에 감염되어 있는 것으로 나타나 있고, B형 간염환자의 약 10%가 간암으로 진행되어 사망하는 것으로 보고 되고 있다. 시스템 바이올로지에 대한 연구는 게놈프로젝트가 발표된 2000년에 하나의 조류로 자리 잡았다. 이 분야의 적용은 현재 세계 유수의 제약회사에서 그 기반을 확립하려고 하고 있다. 따라서, 위 질병 유도에 대한 'system biology'를 통한 접근은 매우 시급하다.

### III. 연구개발의 내용 및 범위

이 신호전달 체계의 교란은 HBV에 의한 간질환의 유도에 원인을 제공한다. 지금까지의 연구는 HBV의 특정 단백질에 의해 변화되는 세포내 신호전달 체계에 대한 분석을 시간적 변화에 따른 양상을 분석한 동적인 분석이 아니라, 신호전달 체계의 교란 정도의 가능성을 나타내는 정적인 분석이 주를 이루었다. 이번 과제에서는 동적인 실험 분석뿐만 아니라, 이 분석을 바탕으로 in silico model을 완성하고, 이 model을 이용한 simulation을 통해 HBV protein에 의한 동역학적 의미를 밝히고자 노력하였다.

### IV. 연구개발결과

이 연구 과제를 통하여 본 연구팀은 총 10가지의 HBV viral protein과 interaction하는 세포내 protein 발굴하였으며, 또한, 총 197 가지의 HBV 감염에 의해 변화되는 유전자 발굴하였다. 이를 바탕으로 총 3가지의 HBV에 의해 활성화되는 세포내 pathway의 정량적 분석을 하여 in silico simulation을 위한 data를 제공하였다.

### V. 연구개발결과의 활용계획

본 연구 과제를 통해 완성한 model은 HBV 감염시의 NF- $\kappa$ B activation저해를 위한 drug 개발과 기타의 IKK-p의 upstream pathway에 대한 drug effect를 검증하는데 중요한 model로 사용될 수 있다. 즉, 기존의 phamacokinetics 연구의 model로 응용될 수 있는 연구 결과 이다.

## S U M M A R Y

(영 문 요 약 문)

### I. Title

Analysis of protein network in HBV-associated hepatic cells

### II. The Present State of Technical Development

10% of Korean is infected with HBV and 10% of the infected develops hepatocellular carcinoma. Now, systems biology is important theme and this can be applied to drug discovery. Therefore, this approach is important for investigation of HBV.

### III. Result of the Project

Disturbance of cellular signaling caused by HBV affects liver disease. In this project, we attempt quantitative analysis of HBV-related signaling and provides data for in silico model construction.

### IV. Achievement of the Project

From this project, we found HBV interacting proteins (10 proteins), differentially expressed genes by HBV (197 genes), and quantitative analysis of HBV related cellular pathway (3 pathways).

### V. Utilization of the Research Results

The constructed in silico model can be applied for verification of drug effect (NF-kB pathway inhibiting chemicals) in NF-kB pathway.



C O N T E N T S

(영 문 목 차)

The First. Summary of the Project.....	7
The Second. The Present State of Technical Development.....	8
The Third. Result of the Project.....	11
The Fourth. Achievement of the Project.....	19
The Fifth. Utilization of the Research Results.....	22
The Sixth. Information of Technical Development in Foreign Country.....	23
The Seventh. References.....	24

목 차

제 1 장 연구개발과제의 개요.....	7
제 2 장 국내외 기술개발 현황.....	8
제 3 장 연구개발수행 내용 및 결과.....	11
제 4 장 목표달성도 및 관련분야에의 기여도.....	19
제 5 장 연구개발결과의 활용계획.....	22
제 6 장 연구개발과정에서 수집한 해외과학기술정보.....	23
제 7 장 참고문헌.....	24

## 제 1 장 연구개발과제의 개요

### 1절. 기술적 측면

1. B형 간염바이러스는 간염을 유발할 뿐만 아니라 간경변에 이어 간암을 일으키는 주요 원인 바이러스이다 (Feitelson, 1999).
2. 이 연구에서 보이는 IT 기술을 이용한 BT 결과의 분석은 향후 IT기술의 BT 기술 접목에 기여할 수 있다.

### 2절. 경제·산업적 측면

1. 우리나라의 40대 남성에게 가장 많이 일어나는 암 중의 하나인 간암을 일으키는 원인 바이러스 제어 기술은 간암의 발병률을 낮추는 측면에서 사회의 경제적 손실을 줄일 수 있다.
2. 신약개발의 경험이 크게 부족한 현 실정에서 선진국의 대형 제약회사와의 직접 경쟁은 현재 어려운 형편이므로 비선진국형 질병인 B형 간염 치료제 개발을 위한 연구의 국내에서의 수행은 국가경쟁력의 제고에 크게 기여하게 될 것임.

### 3절. 사회·문화적 측면

1. 우리나라의 경우 전 인구의 약 10% (450여만 명)가 HBV에 감염되어 있는 것으로 나타나 있고, B형 간염환자의 약 10%가 간암으로 진행되어 사망하는 것으로 보고 되고 있다. 즉 간암의 주요 원인이 HBV에 의한 감염이다. 감염과 간암에 의한 사회, 경제적 손실이 막대한 실정으로 치료제 개발을 위한 연구가 매우 시급하다.

## 제 2 장 국내외 기술개발 현황

### 1절. 연구사례의 조사

#### 1. 외국의 경우

시스템 바이올로지에 대한 연구는 게놈프로젝트가 발표된 2000년에 하나의 조류로 자리 잡았다. 이미 스탠포드, 캠브리지, 버지니아 공대, MIT, 하버드, 칼텍, UCLA서 20여 개 팀이 이 분야 연구에 뛰어들었고, 2002년부터 'System biology' 를 주제로 국제 학술회의를 개최했다.

#### 가. 미국의 system biology의 선두 주자

ISB (Institute for System Biology) 의 리로이 박사팀

- (1) DNA, mRNA, protein, 기능 protein간의 상호 연결 방식과 신호 전달 체계에 대한 이해를 통해 신약개발이 돌파구를 찾을 수 있을 것이라고 주장
- (2) 질병의 원인이 세포간 신호전달 체계의 이상에서 기인하므로, 세포간 네트워크를 이해하고 이를 조절한다면 획기적인 신약을 개발할 수 있다는 설명
- (3) 특히 암이나 관절염 에이즈와 같은 면역체계 관련 병에서 획기적인 치료법을 제시할 수 있을 것으로 기대

#### 나. 일본의 system biology 선두 주자

소니 컴퓨터 사이언스 연구소의 키타노 박사팀

- (1) 생명현상을 이해하고, 이 연구결과를 컴퓨터 사이언스와 접목해 사이보그와 같은 인공생명체를 구현하는데 초점
- (2) 이를 위해 키타노는 대장균과 초파리 연구를 통해 생명현상을 컴퓨터 시뮬레이션화 하는 작업에 초점
- (3) 휴먼로봇 시그(Sig)와 피노(PINO)도 이 연구의 일환이다.

#### 다 HBV에 대한 system biology

- (1) 현재 system biology 체계를 통한 HBV 연구의 접근이 이루어지지 않음

#### 2. 국내의 경우

##### 가. KAIST의 이상엽 박사팀

- (1) prokaryote에서의 metabolic pathway에 대한 system biology
- (2) in silico 상에서의 metabolic pathway의 설정
- (3) 이에 따른 실험 설계 및 결과 분석

##### 3. 조사연구개발사례에 대한 평가

가. 국외의 경우 개념의 설정이 이제 시작 되는 분야로 아직 개발기에 속함

나. system biology를 이용한 HBV에 대한 연구의 시작이 필요

## 2절. 세부 기술사항의 검토 분석

### 1. 국내·외 기술수준 비교표

	비 교 (%는 선진국대비)
Gene expression analysis tool	거의 모든 (> 99%) 기술이 선진국과 대등함
Proteom analysis tool	장비면을 제외한 기술이 선진국과 대등 (> 90%)
Signal pathway analysis tool	최근의 BT 연구와 더불어 이 분야의 연구 기술력이 상승 (> 80%)
In vitro reconstitution analysis tool	최근의 BT 연구와 더불어 이 분야의 연구 기술력이 상승 (> 80%)
Protein expression & purification system	전통적인 기술은 국내에도 기술이 축적됨 (> 90%)
High through put analysis tool	자동화 기술과 함께 발전 하지만, 장비 분야의 발전은 미흡 (< 80%)

### 2. 공정단위별로 주요 기술사항 및 그 기술수준의 분석평가를 다음 사항에 걸쳐 기술함 가. 외국의 경우

	국 외 (아직 미흡한점이 초점)
Gene expression analysis tool	고집적, 빠른 시간, 낮은 error율을 목표로 자동화에 초점
Proteom analysis tool	분석의 자동화에 초점
Signal pathway analysis tool	system biology를 통한 대량 분석으로의 접근이 초점
In vitro reconstitution analysis tool	Kit화와 정량화에 그 초점
Protein expression & purification system	빠른시간, 낮은 비용, 자동화에 그 초점
High through put analysis tool	극미량 handling 기술에 그 초점

나. 국내의 경우

	국 내
Gene expression analysis tool	Scanning 기술, 분석 software 등에 그 초점
Proteom analysis tool	장비면의 발전이 미흡
Signal pathway analysis tool	system biology의 개발기에 참여
In vitro reconstitution analysis tool	선진국의 기술력의 이용, 국내에서의 bioassay의 kit화는 미흡
Protein expression & purification system	실험 기술면은 선진국과 대등하나 장비면에서의 발전은 미흡
High through put analysis tool	극미량 handling 기술에 그 초점

## 제 3 장 연구개발수행 내용 및 결과

### 1절. 이론적 접근 방법

HBV의 genome으로부터 coding되어 translation된 protein들은 직접 또는 간접적으로 세포내의 protein들과 interaction하여 정상적인 세포의 신호전달 체계를 교란시킨다 (Seeger and Mason, 2000). 이 신호전달 체계의 교란은 HBV에 의한 간질환의 유도에 원인을 제공한다. 지금까지의 연구는 HBV의 특정 단백질에 의해 변화되는 세포내 신호전달 체계에 대한 분석을 시간적 변화에 따른 양상을 분석한 동적인 분석이 아니라, 신호전달 체계의 교란 정도의 가능성을 나타내는 정적인 분석이 주를 이루었다. 이번 과제에서는 동적인 실험 분석뿐만 아니라, 이 분석을 바탕으로 in silico model을 완성하고, 이 model을 이용한 simulation을 통해 HBV protein에 의한 동역학적 의미를 밝히고자 노력하였다. 이 연구 과제에서는 HBV에 의해 영향 받는 pathway 중 세가지 pathway에 접근하였다. 첫째, HBV에 의한 NF- $\kappa$ B activation pathway; 둘째, HBx에 의한 selenoprotein P의 down-regulation에 관여하는 pathway; 셋째, HBV에 의해 변화되는 IFN- $\gamma$  pathway 등이다.

### 2절. 실험적 접근 방법

HBV protein에 의한 세포내 신호전달 체계의 변화를 일으키는 동역학적 의미를 밝히기 위하여, 이번 연구과제에서는 다음과 같은 단계를 거치도록 하였다.

첫째, HBV protein이 직접적으로 영향을 끼치는 protein들을 yeast two hybrid system 및 immunoprecipitation & MALDI-TOF analysis등을 통해 찾아 내었다.

둘째, 위 protein-protein binding에 의해 영향을 끼칠것으로 예상되는 protein expression 변화 분석을 cDNA microarray 및 2D electrophoresis 방법을 이용하여 분석하였다.

셋째, 위 분석에 의해 영향 받는 세포내 신호 전달 체계에 대한 pathway를 설정하기 위하여 우선 문헌정보 분석을 통해 pathway를 가장 최신의 data에 맞도록 설정하였다.

넷째, 설정된 pathway (NF- $\kappa$ B pathway, IFN- $\gamma$  pathway 등) 내 protein들의 초기 상태의 농도 및 activated form의 농도를 문헌에서 찾고, 찾을 수 없을 경우 실험을 통하여 그 값을 얻도록 하였다.

다섯째, 협동 연구기관인 서울대학교 컴퓨터 공학과 팀과 공동으로 in silico simulation을 하고, pathway에 대한 검증을 하였다.

### 3절. 연구내용 및 연구결과

본 연구과제를 통해 1-2년차 연구 기간 동안 본 연구팀은 타겟의 설정 및 타겟이 된 pathway내 protein들을 이용한 in silico simulation을 할 수 있는 정량적 분석을 하였다. 그 연구는 아래와 같이 진행 되었다.

#### 1. HBV protein과 직접 binding하는 protein screen

##### 가. Yeast two hybrid system을 이용한 HBV polymerase binding protein screen

(1) HBV Polymerase는 in vitro system에서 상당히 expression하기 힘들며, expression이 되더라도 conformational state가 상당히 불안정하다. 이로 인하여, heterologous system에서 expression된 protein을 이용하여 binding하는 protein을 찾기는 매우 힘들다. 이와 같은 이유로 HBV

polymerase binding protein은 yeast two hybrid system을 이용하여 탐색하였다. 그 결과 HBV polymerase와 binding하는 Apolipoprotein을 발굴하였다.

#### 나. Affinity chromatography를 이용한 HBV core binding protein screen

(1) Core protein과 binding하는 host protein들을 찾기 위한 affinity chromatography technique를 사용했다. Core protein은 direct covalent linkage를 통해 NHS-activated sepharose에 immobilization하였다. 이 실험에서 immobilized core antigens과 binding한 host protein은 MALDI-TOF MS에 의해 identification되었다. 그 결과 5가지의 candidate protein들을 찾아내었다. 이들 중 하나는 human nucleolar phosphoprotein B23 (nucleophosmin)이었다. 게다가, HepG2에서 cellular B23은 co-immunoprecipitation, sucrose density gradient, cell cycle synchronization 실험을 통해 core antigens과 binding함을 보였다.

#### 다. Mammalian expression system을 이용한 surface protein 및 HBx binding protein screen

(1) Surface protein은 lipoprotein으로 prokaryotic system에서 원래의 conformation을 유지하면서 expression하여 얻기 힘들다. 또한, HBx protein도 inclusion되는 경향이 강하다. 하지만, mammalian cell에서 expression시 다른 protein들에 비해 상대적으로 expression을 잘 할 수 있었다. 이 과정에서, surface protein을 expression하기 위한 pCMV-HA/HBs와 HBx expression 하기 위한 pCMV-HA/HBx를 construction하였다. Expression된 protein의 immunoprecipitation & MALDI-TOF analysis 결과 HBx protein에 binding하는 HBXIP와 TFIIB protein을 찾아내었고, surface protein과 binding하는 Hsp70를 찾아내었다.

#### 2. HBV에 의해 직·간접적으로 영향받는 유전자 screen

가. 본 연구실에서 확립된 primary normal human hepatocyte culture system을 이용한 HBV 감염 전후의 gene expression 변화 분석

(1) Culture cell line에는 HBV가 감염되지 않으므로, 실제 HBV 감염에 따른 분석을 하기 위해서는 primary culture가 필수적이다. 본 연구실에서는 이전 연구과제(생명현상 및 기능연구사업)를 통하여 human hepatocyte를 primary culture하는 기술을 확립하였다. 이번 연구과제에서도 이 system을 활용하여 HBV 감염 전후 변화되는 유전자 발현을 분석하였다. 그 결과 149 가지의 유전자가 의미성 있게 변화됨을 발견하였다.

나. 임상 sample 중 HBsAg(+) 간암 환자와 HBsAg(-) 간암 환자의 liver specimen에서의 2DE technique를 이용한 protein display 및 비교 분석

(1) NBNC-, B-type HCC 환자에서 얻은 tumorous tissue sample에서 total protein들을 two-dimensional gel을 통해 분리하여 1000 ~ 1200개의 spot을 찾았다. 각각 tissue sample로부터 얻은 two-dimensional gel image를 분석하여 protein spot들의 volume percentage를 nontumorous tissue sample과 tumorous tissue sample에서 서로 비교하였다. 이렇게 선택된 protein spot은 MALDI-TOF mass spectrometric analysis를 통해 protein identification되었다. 약 60개 정도의 protein이 clearly identification되었다.



Gene	Accession	Description	Expression
PDE2A	Ns.104257	phosphodiesterase 2B, cGMP-stimulated	1.0
IGFBP3	Ns.17360	insulin-like growth factor-binding protein 3 (IGFBP3)	1.0
PPP1R151	Ns.42357	protein phosphatase 1, regulatory (inhibitor), subunit 11	1.0
TRAF7	Ns.40029	TNF receptor-associated factor 7	1.0
CDP66	Ns.40077	G protein-coupled receptor 66	1.0
CDL	Ns.79262	Homo sapiens G protein beta subunit mRNA, partial cds	1.0
WAS	Ns.2157	Wiskott-Aldrich syndrome	1.0
NRK	Ns.47007	NF-kB inducing kinase	1.0
TBL3	Ns.114416	transducin beta-like 3	1.0
ANKBP	Ns.17009	ankyrin repeat domain-like 6 interacting protein	1.0
OPTOL2	Ns.173381	oligodactylin-like 2	1.0
PPP4D	Ns.150238	protein phosphatase 4-panc4c, 15kD	1.0
SVR64	Ns.17354	stem-specific tyrosine-tyrosinase regulator-like 1	1.0
ANKK5	Ns.10294	ankyrin kb	1.0

Gene	Accession	Description	Expression
TFEB	Ns.36	TFE factor	1.0
AOC	Ns.17360	apoptosis-associated speck-like protein containing a CARD	1.0
TRAF3IP3	Ns.11223	tumor necrosis factor receptor superfamily, member 10b	1.0

Gene	Accession	Description	Expression
TRD2	Ns.2034	T cell receptor delta chain	1.0
CD226	Ns.26521	granulysin m	1.0
TRAF2	Ns.173482	Hypoxia inducible factor 1 c-terminus, clone HRA15	1.0
CD73	Ns.118234	Hypoxia inducible factor 1 region protein-rich protein mRNA	1.0
SEK1A2	Ns.79040	serpincan 1b	1.0
FKBP1a	Ns.732	FK506-binding protein 1A (12kD)	1.0

Gene	Accession	Description	Expression
NOL1	Ns.10243	nucleolar protein 1 (12kD)	1.0
ICTL3	Ns.5016	immunohistochemically inducible 3	1.0
SOX1	Ns.11210	SOX1, LIM domain protein 2	1.0
CDP11	Ns.33841	growth differentiation factor 11	1.0
RAD50	Ns.41567	RAD50, BRCA1, chromosome homology	1.0
MYC	Ns.79078	myc proto-oncogene, protein	1.0
PRK2	Ns.6005	protein kinase 2	1.0
EPHA1	Ns.88620	EphA1	1.0
CDK	Ns.77753	c-src tyrosine kinase	1.0

Gene	Accession	Description	Expression
ZNF173	Ns.1282	zinc finger protein 173	1.0
ZNF212	Ns.109120	zinc finger protein 212	1.0
LINC2	Ns.140277	LINC00002 protein 2	1.0
EPHA	Ns.102661	ephrin type 2 receptor (EPA-2, PJA-1 related)	1.0
MAP3C	Ns.179274	mitral factor of actin-stimulated 3-cyclic cytoplasmic 2	1.0
IRF5	Ns.10738	interferon regulatory factor 5	1.0
ZNF175	Ns.110414	zinc finger protein 175	1.0
SCMH1	Ns.11425	sex comb on midleg domain 1	1.0
TYRO8	Ns.6090	tyrosine kinase receptor, viral factor 3	1.0
CDK1	Ns.115243	cyclin dependent kinase 1 (CDK1, p34)	1.0
TBL2	Ns.273290	tbl2 in mouse protein TBL2	1.0

Gene	Accession	Description	Expression
AIP1C1	Ns.156423	apoptosis-inducing factor 1 complex, gamma polypeptide 1	1.0
LOC110161	Ns.14181	NB1 protein	1.0
TYN	Ns.172084	tyrosinase	1.0
NHL	Ns.40027	nucleolar VCP-like	1.0
CTSL	Ns.125361	Cat and B-lysozyme component 1	1.0
SLC22A1	Ns.32621	solute carrier family 25 (SLC22), acid transporter, member 1	1.0
UNC118	Ns.41728	unc-118 (C. elegans) homolog	1.0
HMP13	Ns.53041	CC-22 protein	1.0
IFITM1P	Ns.4364	IFITM1 protein, isoform 1	1.0
BR3	Ns.40008	brachyury integrator-1	1.0
LOC151227	Ns.29277	mitochondrial inner membrane protein DSCD7	1.0
SEPP1	Ns.12516	serpincan 1, plasma, 1	1.0
CAST	Ns.37867	cathepsin	1.0
VDUP1	Ns.170226	upregulated by 1,25-dihydroxyvitamin D-3	1.0
SLC6P	Ns.1300	solute carrier 6, sodium dependent protein	1.0
LOC151099	Ns.40079	soluble RNA helicase, DEAD domain of DEAD box family	1.0
MC2L1	Ns.126222	MC2L1, cytoskeleton-binding protein	1.0
MAA1846	Ns.7218	similar to acetyl-CoA synthetase	1.0
MGCD287	Ns.109881	CDP-1 protein	1.0
SOX2	Ns.173699	SOX2, HMG box domain, O/ECD variant of DEAD box family	1.0
SLIT	Ns.126081	slit-like protein 1, cytoskeleton-binding protein	1.0
LOC151089	Ns.40073	cytoskeleton-binding protein 1, precursor isoform	1.0
SRB142	Ns.20459	SRB142, membrane protein precursor	1.0
LOC151083	Ns.131919	SRB142, membrane protein precursor	1.0
CDK107	Ns.5542	cyclin dependent kinase 10, isoform 2	1.0
HMA1	Ns.251177	hematopoietic, alpha 1	1.0
TRCC	Ns.75064	ribosome-associated chaperone c	1.0
SRPQ	Ns.100019	signal recognition particle receptor	1.0

Gene	Accession	Description	Expression
AIF1	Ns.74204	apoptosis-inducing factor 1	1.0
MCB1	Ns.74212	mitochondrial chaperone 1	1.0
RAD51	Ns.143728	protein kinase Chk1	1.0
HSP27	Ns.144770	heat shock 70 kDa protein family class B member 2	1.0
CDP7	Ns.27360	cytoskeleton-binding protein 7	1.0
PPP2R2	Ns.151231	protein phosphatase 2, regulatory subunit 2B	1.0
PP2C	Ns.74	protein phosphatase 2C	1.0

Gene	Accession	Description	Expression
MAA181	Ns.4814	membrane-associated, alpha class 1B, member 1	1.0
CLC	Ns.289	chloride channel protein	1.0
MBIP1	Ns.75020	membrane-binding protein 1	1.0
HAP1	Ns.10174	proteinase A	1.0

Gene	Accession	Description	Expression
EST10099	Ns.25657	Homo sapiens EST10099 gene, last exon	1.0
KSP3	Ns.98235	EBT3	1.0
LOC147846	Ns.128412	EBT3	1.0
FLJ12345	Ns.287128	EBT3	1.0
	Ns.100753	EBT3	1.0
	Ns.124640	EBT3	1.0
RA442L1	Ns.15066	EBT3	1.0
	Ns.284285	Homo sapiens pBCR1 mRNA, partial cds	1.0
	Ns.8941	EBT3	1.0
	Ns.270023	EBT3	1.0
FLJ12742	Ns.173153	EBT3	1.0
	Ns.196162	EBT3	1.0
	Ns.22094	EBT3	1.0
	Ns.88237	Homo sapiens clone 1766 and 2395 mRNA sequence	1.0
	Ns.50243	EBT3	1.0
	Ns.49047	EBT3	1.0
	Ns.264399	EBT3	1.0
	Ns.270023	EBT3	1.0
	Ns.108258	EBT3	1.0
	Ns.105622	EBT3	1.0
FLJ22478	Ns.238246	EBT3	1.0
	Ns.37912	EBT3	1.0
	Ns.57834	EBT3	1.0
	Ns.71100	EBT3	1.0
	Ns.35080	EBT3	1.0
	Ns.90266	EBT3	1.0
	Ns.211233	EBT3	1.0
	Ns.10970	EBT3	1.0
	Ns.129148	EBT3	1.0
HMA1754	Ns.8501	EBT3	1.0
	Ns.10521	EBT3	1.0

Gene	Accession	Description	Expression
VRL	Ns.31718	Homo sapiens cDNA FL110348, clone PLAC1094258	1.0
FLJ10449	Ns.81946	hypothetical protein	1.0
FLJ10729	Ns.77553	Homo sapiens cDNA FL10729, clone H7729P264116	1.0
P158	Ns.122575	Homo sapiens HSPC295 mRNA, partial cds	1.0
	Ns.279006	Homo sapiens mRNA: cDNA DKFZp434E99	1.0
HSPC134	Ns.275781	HSPC134 protein	1.0
PRO146	Ns.287116	hypothetical protein PRO146	1.0
	Ns.12377	Homo sapiens clone 2349 and 23755 unknown mRNA, partial	1.0
UPF3A	Ns.273412	Homo sapiens cDNA FL108207, clone HEBM1055175	1.0
KIAA1020	Ns.108445	KIAA1020 protein	1.0
KIAA1973	Ns.105619	KIAA1973 gene product	1.0
DNF2P18	Ns.102750	hypothetical protein DNF2P180E1623	1.0
KIAA1982	Ns.11564	KIAA1982 protein	1.0
KIAA1980	Ns.1852	KIAA1980 protein	1.0
FLJ11305	Ns.1064	hypothetical protein FLJ11305	1.0
PRO1073	Ns.80975	PRO1073 protein	1.0
FLJ29113	Ns.100393	hypothetical protein FLJ29113	1.0
	Ns.172989	Homo sapiens mRNA: cDNA DKFZp434E99	1.0
HSPC047	Ns.279643	HSPC047 protein	1.0
KIAA1512	Ns.172150	KIAA1512 protein	1.0
KIAA1983	Ns.104981	KIAA1983 protein	1.0
ELK2	Ns.161789	KIAA1981 protein	1.0
HSPC126	Ns.181112	HSPC126 protein	1.0
PRO2525	Ns.10804	hypothetical protein PRO2525	1.0
KIAA1984	Ns.115778	KIAA1984 gene product	1.0
SLT	Ns.64609	hypothetical protein	1.0
LOC151083	Ns.211545	hypothetical protein	1.0
FLJ20728	Ns.152061	Homo sapiens cDNA FLJ20728, clone HEPH227	1.0
FLJ10512	Ns.100913	hypothetical protein FLJ10512	1.0
HSD11	Ns.177311	hypothetical protein	1.0

Fold 1/18 1/4 1/2 2 4 16

그림 1. Primary normal human hepatocyte의 HBV 감염 전후의 gene expression 변화분석. Primary normal human hepatocyte를 이용한 HBV 감염 model을 이용하여 HBV의 감염에 의해 변화되는 유전자 149 가지를 찾아내었다. 위 결과는 HBV에 의해 발현이 변화되는 유전자를 그 기능별로 분류하여 나타낸 것이다.

significant variantion을 보이는 60 protein들 중에 7개의 stress-associated protein들과 nucleophosmin, elongation factor 2는 모든 type의 HCC에서 증가하였다. 반대로 CPSASE I과 argininosuccinate synthase는 모든 type의 HCC에서 감소하였다

(2) Hepatitis virus 감염에 의해 differentially expressed protein들의 analysis는 NBNC proteins뿐 아니라 HBV-associated protein들을 알 수 있게 해준다. SOD (Cu-Zn)의 expression은 B-type HCC (5 of 7, 2.5 ± 1.1 ; 그림 4에서 decrease 경향을 보였고 dihydrolipoamide dehydrogenase는 NBNC-type HCC (7 of 7, 2.4 ± 0.5 ; 그림 3)에서 증가하는 경향을 보였다. ribosomal binding protein은 NBNC type HCC에서는 거의 변화가 없지만 B-type HCC (6 of 7, 3.6 ± 2.8 ; Table 2)에서는 clearly decrease하였다.

3. HBV에 의해 영향받는 발굴된 유전자 중 network을 이루는 유전자군 선별 및 정량적 분석

가. In vitro 상에서의 HBV 감염 model을 이용하여 설정된, TNF- $\alpha$  & LT- $\alpha$ ->IKKs->NF- $\kappa$ B activation pathway의 정량적 분석

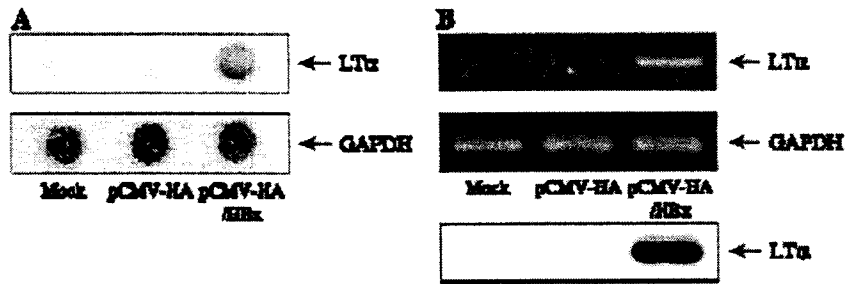


그림 2. HBx에 의한 LT-a의 expression의 증가를 mRNA level에서 dot blot analysis (A)와 RT-PCR (B)을 이용하여 확인하였다. 이 결과는 HBV component 중 HBx가 LT-a의 expression을 증가 시킨다는 것을 나타낸다.

(1) In vitro 상에서의 HBV 감염 model을 이용하여, HBV 감염에 의해 LT- $\alpha$ 가 up-regulation됨을 밝혀내었다. 이는 LT- $\alpha$  -> IKKs -> NF- $\kappa$ B로 이어지는 pathway가 강화될 수 있음을 나타낸다. 우선 이번 연구년도에서는 HBV 감염에 의해 생성된 LT- $\alpha$ 가 autocrine으로써의 감염된 세포에 대한 영향을 정량적으로 분석하였다. 그 방법은 HBV component인 HBx에 의해 expression이 증가된 TNF- $\alpha$ 와 LT- $\alpha$ 에 대한 blocking antibody를 처리하여 어느 정도 영향을 주는 지를 분석하였다. 그 결과 TNF- $\alpha$  및 LT- $\alpha$  는 각각 전체 NF- $\kappa$ B activation의 20 % 정도씩 차지하였다.

나. HBV에 의한 NF- $\kappa$ B pathway내 protein들의 gene up-regulation에 의한 NF- $\kappa$ B activation 강화

(1) In vitro 상에서의 HBV 감염 model을 이용하여 설정된 TNF- $\alpha$  & LT- $\alpha$  -> TNFR -> TRAF2 -> NIK -> IKKs -> NF- $\kappa$ B activation pathway에 참여하는 protein들의 level을 분석하였다. 이 protein들의 level 증가는 NF- $\kappa$ B activation pathway의 positive feedback을 줄 수 있어, HBV 감염에서의 NF- $\kappa$ B activation modeling시 중요한 의미를 가진다. 이번 연구에서는 LT- $\alpha$ , TRAF2, NIK의 expression 증가를 quantitative PCR을 통하여 HBV 감염시에 증가되는지를 분석하였다. 그 결과 이 세 protein들은 모두 2배 이상 유의하게 증가함을 확인하였다. 이 뿐만 아니라, 이 세 protein들을 통해 HBV는 NF- $\kappa$ B를 activation함을 확인하였다. 이 연구에서는 HBV replication을 유도할 수 있는 1.2배의 HBV genome이 들어있는 HBV1.2 plasmid를 사용하여 hepatoma cell line에 transfection하는 model을 사용하였다.

다. HBx에 의한 Selenoprotein P의 down-regulation

(1) In vitro 상에서의 HBV 감염 model을 이용하여 설정된 selenoprotein P의 down regulation을 본 연구팀에서는 이전 과제를 통하여 발견하였다. HBV에 의한 이 selenoprotein P의 down regulation mechanism을 밝히기 위하여 본 연구 팀은 이번 과제에서 selenoprotein P의 regulatory element를 탐색하였다. 실험적 방법은 promoter region의 확보와 이후 minimal promoter region 탐색 및 deletion을 통한 HBV에 의한 regulatory element를 검색하였다. 그 결과 -300 ~ -600 region 사이에 HBV에 의한 selenoprotein down regulation에 참여하는 element가 있음을 확인하였다. 이 연구를 통해 HBV에 의한 oxidative stress 과정에 중요하게 참여하는 protein인

selenoprotein의 regulation mechanism을 modeling 할 수 있는 promotor region의 circuit을 구성할 수 있었다.

라. HBV에 의한 NF- $\kappa$ B activation이 저해 받는 IFN- $\gamma$ 에 의한 pathway

(1) 이 연구의 이론적 접근 방법은 다음과 같다. In vitro 상에서의 HBV 감염 model을 이용하여 설정된 HBV에 의한 NF- $\kappa$ B activation model에서 NIK는 이 activation pathway에 참여함을 본 연구팀은 밝혀내었다. 또한, NIK의 level이 mRNA와 protein level이 in vitro에서의 HBV에 감염

model에서 높아짐을 확인하였다. 실제 HBV 감염된 사람의 간에서도 NIK의 level이 변화되는지를 분석하는 도중 NIK가 HBV 감염된 사람에 특이적으로 nuclear region에 localization되어 있음을 확인하였고, 이것이 in vitro에서 IFN- $\gamma$ 의 처리에 의해 재현될 수 있음을 확인하였다. 최근의 논문에서 의하면, NIK의 nucleo/cytoplasmic shuttling의 저해는 NIK에 의한 NF- $\kappa$ B activation을 저해한다는 논문이 있었다. 이를 바탕으로 HBV 특이적인 NIK nuclear localization이 HBV에 의한 NF- $\kappa$ B activation을 저해할 수 있는지를 분석하였다.

(2) 실험적 접근 방법은 다음과 같다. IFN- $\gamma$ 가 HBV 특이적으로 NF- $\kappa$ B activation을 저해하는지를 electro mobility shift assay (EMSA) 방법과  $\kappa$ B binding site를 가지고 있는 reporter plasmid를 이용한 reporter assay를 이용하여 확인하였다. 또한, NIK nuclear localization signal을 제거한 후 IFN- $\gamma$ 에 의한 NF- $\kappa$ B activation저해를 분석하는 실험을 진행하였다.

(3) 위 실험 결과 우리는 IFN- $\gamma$ 가 HBV에 의한 NF- $\kappa$ B activation을 저해하고 있음을 확인하였다. 이 결과가 NIK의 nuclear localization에 기인함을 밝히기 위한 NIK nuclear localization signal을 제거한 mutant를 이용한 실험에서 NIK nuclear localization이 일어나지 않으면, IFN- $\gamma$ 에 의해 HBV에 의한 NF- $\kappa$ B activation이 저해되지 않음을 확인하였다. 이 결과는 HBV 감염시 activation된 T lymphocyte에서 분비된 IFN- $\gamma$ 가 HBV에 감염된 세포에 특이적으로 NF- $\kappa$ B를 저해하는 것을 뒷받침하는 중요한 결과이다. 이 결과를 통해 in vivo에서 HBV 감염시 면역 시스템에 의한 HBV에 기인한 NF- $\kappa$ B activation을 저해하는 circuit를 작성하였다. 이는 in silico modeling을 들어가기 위

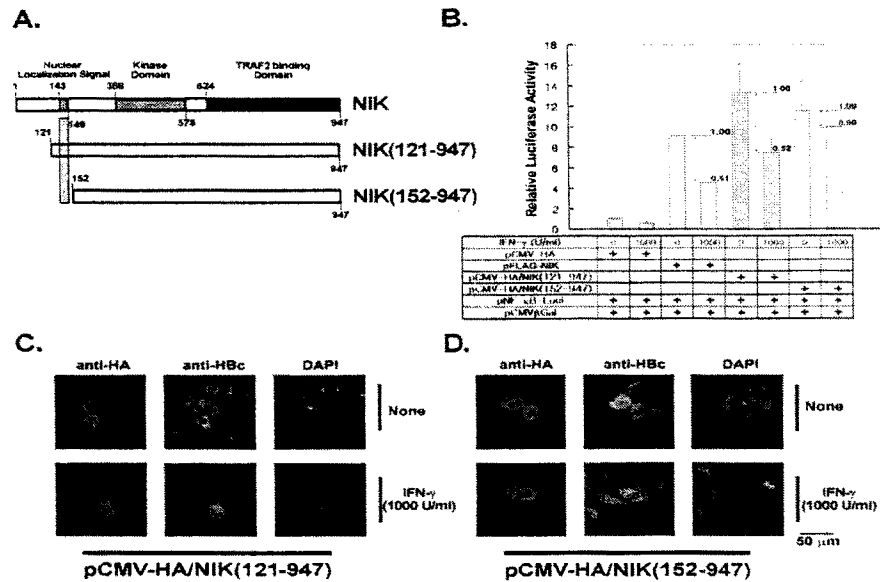


그림 3. IFN- $\gamma$ 에 의한 NIK nuclear localization과 NF- $\kappa$ B activation 저해와의 관계. NIK의 nuclear localization signal을 제거하면, HBV가 replication하고 있는 상황이라도, IFN- $\gamma$ 에 의해 NIK가 nuclear로 이동하지 않는다. 또한, activation된 NF- $\kappa$ B가 IFN- $\gamma$ 에 의해 저해 받지 않는다. 이결과는 IFN- $\gamma$ 에 의한 NF- $\kappa$ B activation 저해는 NIK nuclear localization과 밀접한 관련성을 가짐을 나타낸다.

한 초기 단계의 circuit을 완성한 것이다.

마. HBV감염시 IFN- $\gamma$  pathway에 의한 cell cycle 변화 및 세포 형태 변화 분석

(1) 본 연구과제를 통하여 HBV가 감염되면 IFN- $\gamma$ 에 대한 반응이 달라짐을 확인하였다. 이결과는 HBV에 의하여 IFN- $\gamma$ 에 의한 pathway가 변화됨을 나타내는 것이다. 이는 위 연구결과와 더불어, HBV 특이적으로 cell cycle의 progression을 저해하고, 세포의 cytoskeleton의 reorganization에 의해 세포 모양의 변화를 가져오는 것이다. 이 결과는 이 연구 결과는 IFN- $\gamma$ 의 농도별 실험을 통하여 in silico 상에서의 simulation을 검증할 수 있는 형태의 data를 얻었다. 즉, in silico 상에서 나타나는 simulation pattern을 in vitro 실험으로 검증하기 위한 형태를 가지는 것이다. HBV 특이적으로 cell cycle progression의 저해와 세포의 모양 변화 유도는 특히 세포의 사멸과 밀접한 관련성을 가짐을 tunel assay라는 apoptosis detection 방법을 통하여 검증하였다. 즉, IFN- $\gamma$ 는 HBV 특이적으로 반응하여 감염된 세포의 사멸을 유도할 수 있어 HBV의 제거를 유도할 수 있다는 간접적인 결과를 나타낸다.

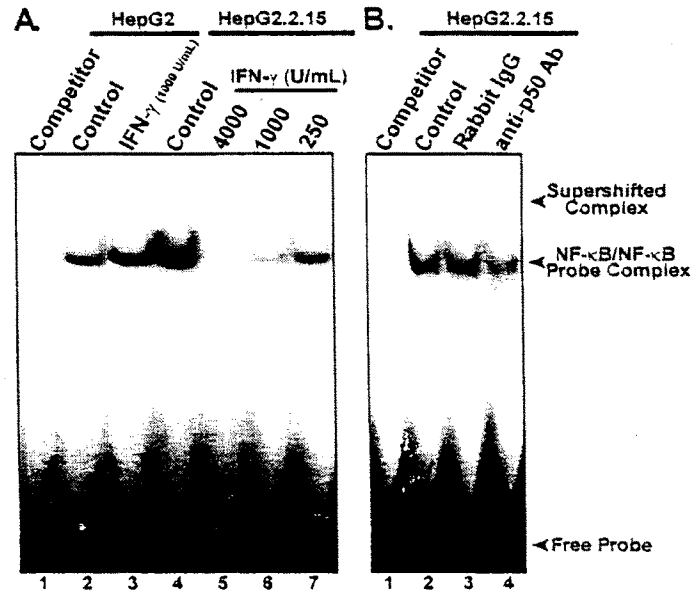


그림 4. IFN- $\gamma$ 에 의한 HBV-mediated NF- $\kappa$ B activation저해. IFN- $\gamma$  처리 후 각 cell에서 nuclear extract를 준비한 후 EMSA 방법을 이용하여 NF- $\kappa$ B activity를 측정하였다. 그 결과 HBV가 replication되는 cell에서만 IFN- $\gamma$ 에 의하여 NF- $\kappa$ B가 저해됨을 확인하였다.

바. HBV 감염에 의한 NF- $\kappa$ B activation시의 state 분석

(1) HBV가 감염되면 HBx, preS2 activator등의 component들에 의한 NF- $\kappa$ B가 activation된다는 보고가 많다. 우리 연구팀의 결과에 의하면, HBV는 TNF- $\alpha$  및 LT- $\alpha$ 의 NF- $\kappa$ B inducing cytokine의 expression을 늘리며, 세포내에서 NF- $\kappa$ B pathway에 참여하는 protein의 level을 증가시켰다. 이는 HBV에 의한 NF- $\kappa$ B activation processes의 성격을 결정할 수 있는 중요한 factor로 작용될 수 있다. 이 분석을 위해 우리 연구팀은 세포의 volume을 측정할 수 있는 방법과 단위 세포 volume내 존재하는 protein 양을 측정할 수 있도록 측정하고자 하는 각 protein의 concentration에 대한 standard를 준비하였다. 준비된 과정을 거쳐 각 protein 특이적인 antibody를 이용하여 immunoblot analysis를 수행하였고, 그 결과 HBV가 없는 HepG2 cell과 HBV 감염 model로 사용되는 HepG2.2.15에서의 각 protein (TNF- $\alpha$ , LT- $\alpha$ , TNFR, TRAF2, NIK, IKK- $\alpha$ , IKK- $\beta$ , NF- $\kappa$ B1, NF- $\kappa$ B2 등)의 initial concentration을 측정할 수 있었다.

사. HBV 감염의한 NF- $\kappa$ B activation root를 찾기 위한 in vitro/in silico 분석

(1) HBV에 의한 NF- $\kappa$ B activation시 어떤 root를 이용하는지를 분석하기 위하여, in silico system에서 보이는 각 protein node의 inhibition curve를 in vitro와 in silico 상에서 분석하였다. 그 결과 대부분의 HBV에 의한 NF- $\kappa$ B activation은 NIK를 위주로 이루어짐을 확인하였다. 각 inhibition curve를 측정하는 방법은 HBV에 replication하고 있는 세포에 각 protein의 dominant negative mutant protein을 도입시켜 inhibition하는 방법을 사용하였다. 또한, 측정법으로는 reporter plasmid를 이용한 NF- $\kappa$ B reporter assay를 이용하였다.

	HepG2	HepG2.2.15
IKK- $\alpha$	0.0242	0.0724
IKK- $\beta$	0.0228	0.0600
NIK	0.0452	0.0671
TRAF2	0.3339	0.7365
TNFR1	0.8937	1.6533
TNF- $\alpha$	0.0000	0.3899

그림 5. HepG2 cell과 HepG2.2.15 cell에서의 각 molecule의 초기농도를 구한 표. 각 초기 농도는 농도가 계산된 protein standard를 heterologous system을 통해 expression & 정량하여 준비한 후 immunoblot analysis를 이용하여 각 protein을 단위 부피당 농도 ( $\mu$ M)를 측정하였다.

아. HBV 감염에 의한 TNF- $\alpha$ 에 의한 NF- $\kappa$ B activation state변화 분석

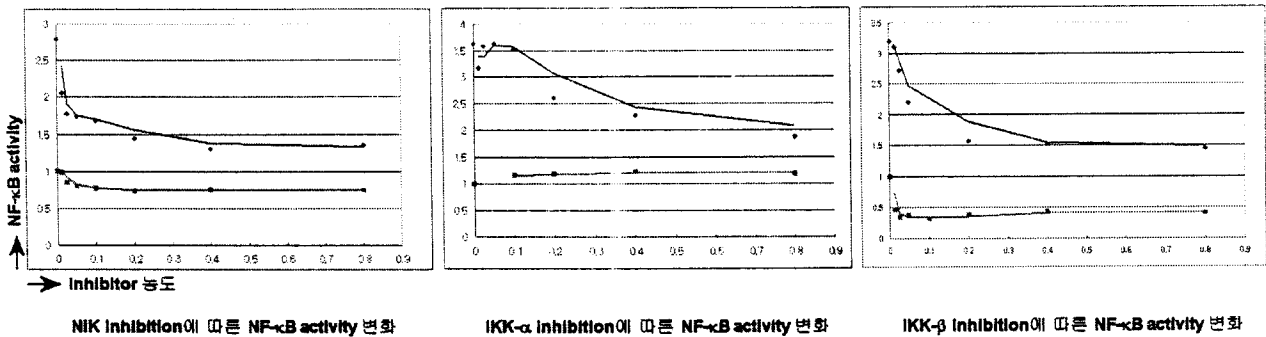


그림 20. HBV에 의한 NF- $\kappa$ B activation root를 분석하기 위하여, HBV가 replication하는 세포(빨간선)와 HBV가 replication하지 않고 있는 세포(파란선)에서 pathway에 참여하는 각 protein별 inhibitor (dominant negative mutant protein)를 사용하여 inhibition curve를 분석하였다. 이 결과 설정된 pathway중 NIK가 HBV에 의한 NF- $\kappa$ B activation시 중요한 root임이 나타내어 졌다.

(1) In vivo에서 HBV 감염시 inflammation을 일으키는 proinflammatory cytokine들 중의 하나가 IFN- $\gamma$ 와 TNF- $\alpha$  이다. 이중 TNF- $\alpha$ 는 FADD를 통한 apoptotic pathway를 activation할 뿐만 아니라, TRAFs 통한 NF- $\kappa$ B pathway를 activation한다. 최근의 연구 결과에 의하면, TNF- $\alpha$ 에 의한 NF- $\kappa$ B activation은 oscillation pattern을 가지며, 이 oscillation pattern을 가지게 되는 원인은 IKK-p의 농도와 negative feedback인 I $\kappa$ Ba의 증가로 보고 있다. 이 oscillation pattern은 in silico 상에서 matlab등을 통해 simulation하는 것으로 그 결과를 뒷받침하고 있다. HBV 감염시 NF- $\kappa$ B activation pathway내 protein들의 level 변화와 activation된 protein 비율이 다름을 우리는 이번 연구과제를 통해 밝혀 내었고, 이는 TNF- $\alpha$ 에 의한 NF- $\kappa$ B activation시 oscillation pattern에 변화를 줌을 확인하였다. 이 연구는 HBV 감염 model로 사용되는 HepG2.2.15 과 control인 HepG2에 TNF- $\alpha$ 를 처리하고 시간별 처리 후 EMSA를 통해 그 oscillation pattern이 달라짐을 확인하였다. 이때, NF- $\kappa$ B probe (NF- $\kappa$ B1 & NF- $\kappa$ B2 complex가 모두 binding)와 NF- $\kappa$ B2 probe (NF- $\kappa$ B2 complex만 binding)를 이용하여, NF- $\kappa$ B1 & NF- $\kappa$ B2 subset을 나누어 비교 분석하였다. 또한, TNF- $\alpha$  signal

시 HBV에 의한 NF-κB activation root인 NIK를 TRAF2를 통해 저해하여, HepG2.2.15에서의 초기 TNF-α에 의한 NF-κB activation시 undershoot 현상이 나타남을 실험적으로 밝혀내었다.

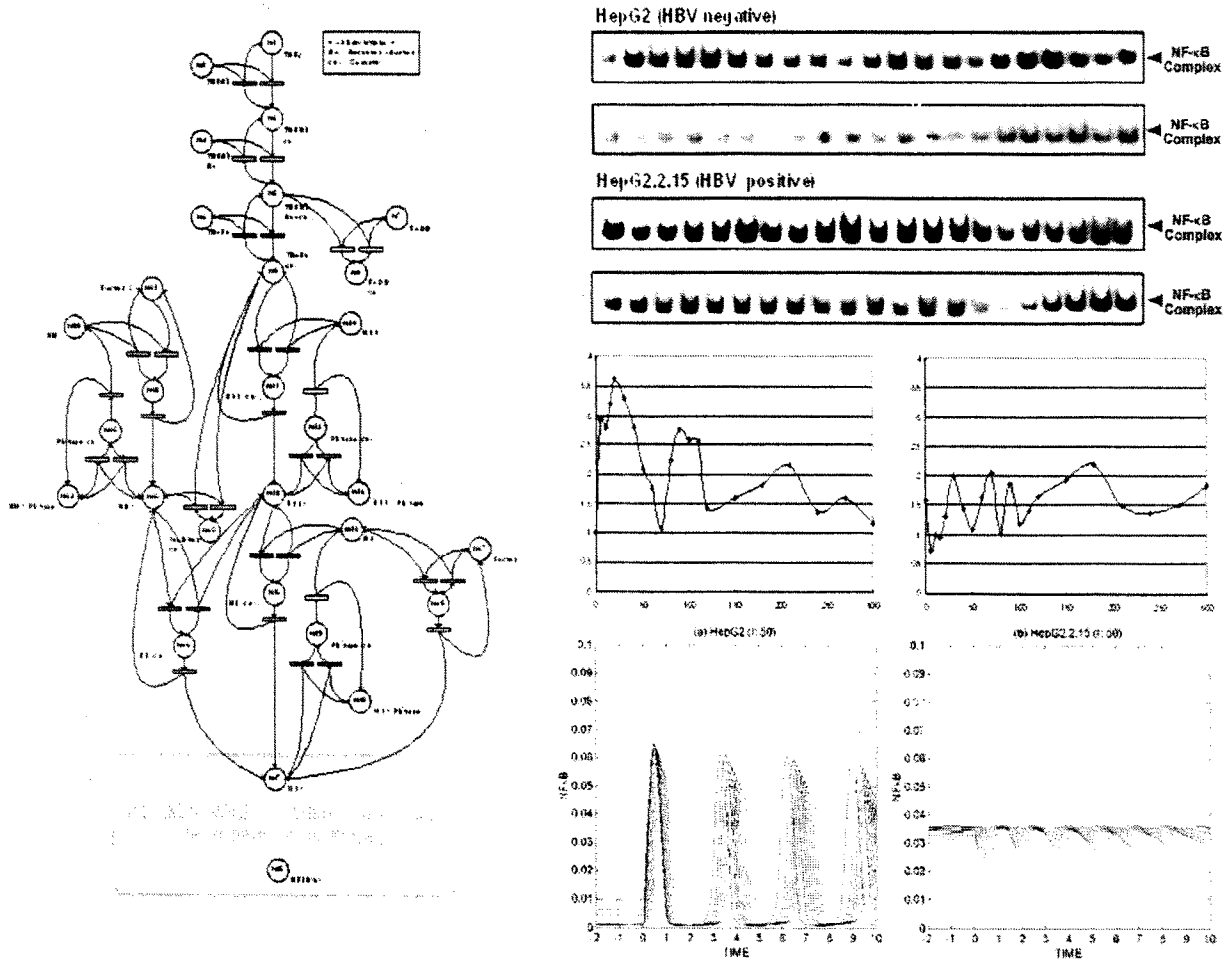


그림 7. 협동 연구기관과의 연구를 통해 얻은 *in vitro/in silico* 실험 결과. 상호작용적인 실험에 의해 *in vitro* 실험만으로 해결할 수 없는 결과를 해석할 수 있었다. (A) TNFα 투여에 의한 NF-κB 반응 EMSA 이미지 (B) EMSA 결과를 정량적으로 표현한 그래프 (C) *in silico* 시뮬레이션 결과. 서로 다른 phosphatase 비율에 대해 시뮬레이션을 진행하였다. (Phosphatase 비율이 적을 수록 진한 선으로 표시)

## 제 4 장 목표달성도 및 관련분야에의 기여도

### 1절. 연구개발의 최종목표

HBV의 viral protein (polymerase, core protein, surface protein, HBx) 과 hepatocyte내의 protein 네트워크의 상호작용을 분석하여 150 가지 이상의 상호작용하는 protein을 발굴한다. 이 protein 이 속한 간세포내 protein 네트워크 연구를 협동 연구기관의 database system을 통해 질병 (간염, 간경변, 간암등을 포함하는 간질환) 유도 기작을 수식적으로 예측할 수 있는 model 2가지 이상을 제시한다. 이 protein들의 이 연구 결과를 토대로 journal of biological chemistry 및 gastroenterology등의 저널에 2편 이상 게재한다.

### 2절. 연차별 연구개발 목표 및 내용

구분	연구개발 목표	연구개발 내용 및 범위
1차년도 (2003)	●HBV viral protein과 상호작용하는 protein 100 가지 이상 발굴	<ul style="list-style-type: none"> <li>●HBV protein (polymerase, core protein, surface protein, HBx protein)들과 결합하는 protein 탐색 (10 가지 이상 발굴)</li> <li>●HBV protein (polymerase, core protein, surface protein, HBx protein)들에 의해 결합이 아닌 방법에 의해 영향 받는 protein 탐색 (150 가지 이상 발굴)</li> </ul>
2차년도 (2004)	●1차년도에 탐색된 protein 군을 선별하여 정량적 분석을 시도 (20 가지 이상 정량적 분석)	<ul style="list-style-type: none"> <li>●협동 연구 기관의 분석을 통해 systematic flow를 완성할 수 있는 타겟 설정 (2 가지 pathway 이상 설정)</li> <li>●설정된 타겟 protein들과 HBV 각 protein들과의 상호작용을 정량적으로 분석 (in vivo system &amp; in vitro system)</li> <li>●본 연구실의 선행 연구인 TNF-b, TRAF2, NIK, SePP와 HBV 각 protein들과의 상호작용관계를 정량적으로 분석 (4 가지 protein 분석)</li> <li>●위 연구 결과를 토대로 HBV에 의해 영향을 받는 또는 영향을 주는 circuit을 3 종류 이상 완성</li> </ul>

### 3절. 계획대비 달성도

번호	세부연구목표	달성내용	달성도(%)
1	○ HBV protein과 직접 결합하는 protein들을 affinity chromatography, I.P. & 2DE 등의 방법으로 발굴	1. Yeast two hybrid system을 이용하여 HBV polymerase에 binding하는 protein 1가지를 발굴하였다. 2. Affinity chromatography 방법을 이용하여 core protein에 binding하는 protein 5가지를 발굴하였다. 3. Hepatoma cell line에서 expression된 surface protein, HBx protein에 binding하는 protein을 immunoprecipitation 방법을 통하여 4가지 발굴하였다.	100
2	○ HBV protein들과 직접 결합이 아닌 방법에 의해 영향 받는 protein들을 microarray, 2DE, phosphoprotein display 등의 방법에 의한 발굴	1. Primary normal human hepatocyte를 이용하여 HBV 감염 전후 변화되는 유전자를 cDNA microarray technique를 이용하여 149가지를 발굴하였다. 2. HBV 관련으로 간암 발달과 관계있는 protein을 2DE방법을 통하여 48 가지를 발굴하였다.	100
3	○ systematic flow를 완성할 수 있는 타겟 설정	1. HBV에 의한 NF- $\kappa$ B activation관련 protein network를 선별하여 NF- $\kappa$ B activation을 반응별로 정량적으로 분석하였다. 2. 세포외 자극에 의한 NF- $\kappa$ B activation변화를 세포내 신호전달 network를 분석하여 규명하였다.	100
4	○ 설정된 타겟 protein들과 HBV protein 들과의 상호작용을 정량적으로 분석 (in vivo & in vitro system을 이용)	1. HBV에 의한 NF- $\kappa$ B activation시 참여하는 protein들(TNF- $\alpha$ , TNFR, TRAF2, NIK, IKK- $\alpha$ , IKK- $\beta$ 등)의 initial concentration과 kinetic pattern을 분석하였다.	100
5	○ TNF- $\beta$ , TRAF2, NIK, SePP와 HBV protein들과의 상호작용을 정량적으로 분석	1. TNF- $\beta$ 에 의한 NF- $\kappa$ B activation을 정량적으로 분석하였다. 2. HBV에 의한 NF- $\kappa$ B activation시 참여하는 TRAF2, NIK의 kinetic pattern을 분석하였다. 3. HBx에 의한 Sepp downregulation에 대한 정량적 분석을 하였다.	100
6	○ HBV와 상호작용하는 protein group를 이용한 circuit 완성	○ HBV와 상호작용하는 protein들을 이용하여, HBV에 의해 상호작용하는 protein group의 circuit을 완성하였다. 1. HBV에 의한 NF- $\kappa$ B activation pathway 2. HBx에 의한 Sepp down-regulation pathway 3. IFN- $\gamma$ 에 의한 HBV에 일어난 NF- $\kappa$ B activation pathway 저해 pathway	100



연차	변경 전 (선정연도 연구계획서)			변경 후			비고
	연구목표	연구내용	연구범위	연구목표	연구내용	연구범위	
1							
2							
3							

## 제 5 장 연구개발결과의 활용계획

### 1절. 추가 연구의 필요성

본 연구과제는 1단계에 속하는 3년 계획으로 짜여진 연구계획을 갖추고 시작되었다. 따라서, 실제 연구가 마무리 되지 않은 상황에서 종료되는 결과가 되었다.

본 연구의 계획은

1년차: in silico modeling을 위한 target 탐색

2년차: 탐색된 target과 관련된 parameter estimation; prototype의 in silico modeling 구축

3년차: in silico modeling의 실험을 통한 검증 및 상호 조정 그리고 적용

이었다.

본 연구팀은 실제 진행 속도보다 빠르게, in silico modeling의 parameter estimation을 시작하였고, 또한 실제의 in vivo 문제를 해결하기 위한 적용단계를 수행하는 과정이었다. 따라서, 2년차에 미쳐 끝내지 못한 제 3년차의 나머지에 해당하는 부분이 추가로 연구 되어야 할 대상이다. 실제 in silico modeling의 결과는 실제 실험과 연동되어 변환되고 수정되어야 하며, 실제 문제를 풀 수 있는 model이어야 하기 때문에 추가 연구가 더 요구되어진다.

### 2절. 타 연구에의 응용

Hoffman model (Hoffmann et al., 2004)에 따른 NF- $\kappa$ B inhibitor들인 drug에 대한 inhibitory effect는 검사는 IKK-p 이후의 pathway에 그 초점이 맞추어져 있었다. 하지만, 본 연구과제를 통해 얻은 NF- $\kappa$ B activation model은 HBV 감염시를 적용했을 뿐만 아니라, IKK-p의 up stream 의 model을 포함하고 있다. 이 model은 HBV 감염시의 NF- $\kappa$ B activation저해를 위한 drug 개발과 기타의 IKK-p의 upstream pathway에 대한 drug effect를 검증하는데 중요한 model로 사용될 수 있다. 즉, 기존의 phamacokinetics 연구의 model로 응용될 수 있는 연구 결과 이다.

### 3절. 기업화 추진방안

본 연구팀에서 제안되어진 NF- $\kappa$ B activation model은 아직 정확한 검증 및 특정 문제에 대한 적용을 할 수 있는지에 대한 검증이 이루어지지 않았다. 실제 제약회사등에서 사용할 수 있는 in silico simulation model을 구축하기 위해서는 본과제의 실험 data와 다른 세부 과제의 in silico model과의 상호 연동한 검증 및 추가 연구가 요구된다.

## 제 6 장 연구개발과정에서 수집한 해외과학기술정보

### 1절. NF- $\kappa$ B activation pathway의 변화

1. 최근의 NF- $\kappa$ B activation pathway에 대한 연구가 활발해 지면서 pathway내 각 subset 별의 특징이 밝혀지고 있다. 특히, RelB/NF- $\kappa$ B2 complex의 역할이 새롭게 규명되면서 지금까지 RelA/NF- $\kappa$ B1 complex의 연구와 더불어 더욱 활발히 NF- $\kappa$ B에 대한 연구가 진행되고 있다 (Bonizzi and Karin, 2004). 최근의 NF- $\kappa$ B activation pathway의 연구는 classical pathway라는 RelA/NF- $\kappa$ B1 activation pathway와 alternative pathway라는 RelB/NF- $\kappa$ B2 activation pathway의 두 가지로 나뉘어지고 있다. 또한, 각각 binding하는 cis-element의 specificity가 상당히 다를 수 있음이 최근 밝혀졌다 (Bonizzi et al., 2004).

2. AUC (area under inhibition curve)를 통한 drug inhibitory effect simulation 분석  
가. in silico simulation을 통해 기존 약물의 시간에 따른 반응성을 예측하는 결과를 최근 선보이고 있다. 이는 simulation을 통해 시간에 따라 inhibition되는 정도를 면적으로 나타내어 특정 약물이 얼마나 target을 inhibition을 잘 하는지를 분석하게 된다. 이는 약물 실험의 비용 및 시간을 크게 절약하는데 도움이 될 것으로 보인다. 이 중 특히 hoffmann (Science 2004)의 IKK-p 이후의 NF- $\kappa$ B pathway simulation이 중요한 모델로써 이용되고 있다.

3. 각종 systems biology 관련 DATA base의 설정

가. systems biology 연구가 활발히 진행되면서 다양한 database가 만들어지고 있고, 각 site는 아직 통일된 분류체계를 갖추지는 못하였지만 다양한 실험에서의 parameter를 추출하여 database화하고 있다.

## 제 7 장 참고문헌

- Seeger, C. and W.S. Mason. 2000. Hepatitis B virus biology. *Microbiol Mol Biol Rev* **64**: 51-68.
- Hoffmann, A., A. Levchenko, M.L. Scott, and D. Baltimore. 2002. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **298**: 1241-1245.
- Cho, K.H., S.Y. Shin, H.W. Lee, and O. Wolkenhauer. 2003. Investigations into the analysis and modeling of the TNF alpha-mediated NF-kappa B-signaling pathway. *Genome Res* **13**: 2413-2422.
- Feitelson, M.A. 1999. Hepatitis B virus in hepatocarcinogenesis. *J Cell Physiol* **181**: 188-202.
- Bonizzi, G. and M. Karin. 2004. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends Immunol* **25**: 280-288
- Bonizzi, G., M. Bebien, D.C. Otero, K.E. Johnson-Vroom, Y. Cao, D. Vu, A.G. Jegga, B.J. Aronow, G. Ghosh, R.C. Rickert, and M. Karin. 2004. Activation of IKKalpha target genes depends on recognition of specific kappaB binding sites by RelB:p52 dimers. *Embo J* **23**: 4202-4210.

주 의

1. 이 보고서는 과학기술부에서 시행한 특정연구개발사업의 연구보고서입니다.
2. 이 보고서 내용을 발표할 때에는 반드시 과학기술부에서 시행한 특정연구개발사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니됩니다.

제 출 문

과학기술부 장관 귀하

본 보고서를 “간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석”과제 (세부과제 “C형 간염바이러스 감염에 의한 간세포 단백질 네트워크 변화 분석”) 의 보고서로 제출합니다.

2005. 5.

주관연구기관명 : 한림대학교

주관연구책임자 : 황 순 봉

연 구 원 : 최 수 호

” : 김 소 연

” : 김 연 주

” : 김 태 광

” : 전 현 정

보고서 초록

과제관리번호	M1-0309-06-0002	해당단계 연구기간	2차년도	단계 구분	2차년도 / 3차년도
연구사업명	중 사업명	국채연구개발사업			
	세부사업명	시스템생물학연구사업			
연구과제명	중 과제명	간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석			
	세부(단위)과제명	C형 간염바이러스 감염에 의한 간세포 단백질 네트워크 변화 분석			
연구책임자	황순봉	해당단계 참여연구원수	총 : 6 명 내부 : 6 명 외부 :     명	해당단계 연구비	정부: 59,000 천원 기업:           천원 계: 59,000 천원
연구기관명 및 소속부서명	한림대학교 (일송생명과학연구소)		참여기업명		
국제공동연구	상대국명 :	상대국연구기관명 :			
위탁연구	연구기관명 :	연구책임자 :			
요약(연구결과를 중심으로 개조식 500자 이내)					보고서 면수
<p>1. HCV viral protein과 상호작용하는 간세포단백질의 탐색</p> <ul style="list-style-type: none"> <li>◆ C형 간염바이러스의 NS5A 단백질이 TGF-β receptor I과 상호 작용함</li> <li>◆ C형간염바이러스의 NS5B 단백질이 IκBα kinase (IKK)와 상호 작용함</li> <li>◆ C형 간염바이러스의 core 단백질이 Smad 단백질과 상호 작용함</li> <li>◆ C형 간염바이러스의 NS5A 단백질이 TRAF2와 상호 작용함</li> </ul> <p>2. HCV viral protein에 의하여 영향을 받는 signaling molecules 의 탐색</p> <ul style="list-style-type: none"> <li>◆ C형 간염바이러스의 NS5A 단백질이 TGF-β receptor I과의 상호작용을 통해 TGF-β dependent signal pathway를 조절함</li> <li>◆ C형 간염바이러스의 replicon cell이 TNF에 의해 유도되는 NF-κB activity는 증가시키고 AP-1 activity는 감소시킴</li> <li>◆ Replicon cell에서 p53의 발현과 TNF에 의한 p53의 nuclear translocation의 감소로 인해 apoptosis가 억제됨</li> </ul> <p>3. C형 간염바이러스의 viral protein과 상호작용하는 세포단백질 변화의 정량적 분석</p> <ul style="list-style-type: none"> <li>◆ C형 간염바이러스 NS5A 단백질이 TGF-beta receptor I과 상호 작용하여Smad2의 phosphorylation을 50% 억제함</li> <li>◆ TGF-β에 의해 증가되는 p21의 발현이 NS5A에 의해 50% 감소됨</li> <li>◆ NS5A가 Smad-DNA complex 형성을 50% 억제함</li> <li>◆ TGF-beta receptor I의 kinase activity에는 NS5A가 전혀 영향을 주지 않음</li> </ul> <p>4. HCV-associated marker 단백질 탐색</p> <ul style="list-style-type: none"> <li>◆ C형 간염바이러스의 control Huh7 cell과 HCV replicon cell에서 total RNA를 추출하여 microarray analysis를 수행함 <ul style="list-style-type: none"> <li>- 2배 이상 up-regulated genes : 840 genes, - 2배 이상 down-regulated genes : 1,896 genes</li> <li>- 4배 이상 fold-changed genes : 142 genes</li> </ul> </li> <li>◆ C형 간염바이러스의 NS5A 단백질에 대한 transgenic mice의 cirrhotic liver 및 control mouse의 liver에서 total RNA를 추출하여 microarray analysis를 수행함 <ul style="list-style-type: none"> <li>- 2배 이상 fold-changed genes : 743 genes, - 4배 이상 fold-changed genes : 206 genes</li> </ul> </li> </ul>					32
색인어 (각 5개 이상)	한 글	C형 간염바이러스, 세포단백질, 중앙괴사인자, 형질전환성장인자, 표지단백질, 정량적 분석			
	영 어	Hepatitis C virus, Cellular protein, Tumor necrosis factor, Transforming growth factor, Marker protein, Quantitative analysis			

## 요 약 문

### I. 제 목

C형 간염바이러스 감염에 의한 간세포 단백질 네트워크 변화 분석

### II. 연구개발의 목적 및 필요성

C형 간염바이러스는 급·만성 간염 및 간암을 유발하는 주된 병원체이다. C형 간염바이러스의 genome은 plus-strand RNA로서 크기가 약 9.6 kb이며 C형 간염바이러스의 만성 감염자는 타인에게 바이러스를 전파할 가능성이 높고, 감염자 자신의 경우도 만성간염, 간경변증, 간세포암과 같은 심각한 합병증을 유발할 위험성이 매우 높아 국민 건강에 심각한 위협성을 가지고 있다. C형 간염의 치료는 Interferon이나 Ribavirin등이 사용되고 있으나 현재까지 C형 간염에 특이적으로 효과가 있는 치료제는 아직 개발되지 못하였다. 한국인의 간암에 의한 사망률이 세계 제 1위라는 심각성을 고려할 때 이에 대한 예방 및 치료기술 개발이 절실히 요구된다.

### III. 연구개발의 내용 및 범위

#### 1차년도 (2003)

- ◎ 간세포에서 C형 간염바이러스의 viral 단백질과 상호 작용하는 간세포 단백질의 탐색
- ◎ C형 간염바이러스의 replicon cell을 이용한 microarray data를 통해 C형 간염바이러스에 의해 영향 받는 세포내 단백질 확인
- ◎ C형 간염바이러스의 NS5A 단백질에 대한 transgenic mice를 이용한 microarray data를 통해 간암과 관련 있는 세포내 단백질 확인
- ◎ C형 간염바이러스의 replicon cell을 이용하여 cytokine에 의해 유도되는 세포내 신호 전달과정에 viral 단백질이 미치는 영향 분석
- ◎ C형 간염바이러스의 NS5A 단백질이 TGF에 의해 유도되는 세포내 신호전달을 억제함

#### 2차년도 (2004)

- ◎ Enzyme kinetics를 기초로 하여 TGF-beta signal transduction에 대한 graphical model 정립
- ◎ Cytokine과 C형 간염바이러스의 viral 단백질을 input으로 하고 HCV-associated marker gene을 output으로 하여 이에 대한 단백질 network를 computer science와 협동으로 modeling함
- ◎ Cytokine과 관련된 C형 간염바이러스가 유발하는 병인성에 대한 mechanism을 computer science와 협력하여 새로운 결과를 창출해냄

### IV. 연구개발결과

- ◆ C형 간염바이러스의 NS5A 단백질과 TGF- $\beta$  receptor I과의 상호 작용
- ◆ C형간염바이러스의 NS5B 단백질과 I $\kappa$ B $\alpha$  kinase (IKK)와의 상호 작용
- ◆ C형 간염바이러스의 core 단백질과 Smad 단백질과의 상호 작용



- ◆ C형 간염바이러스의 NS5A 단백질과 TRAF2와의 상호 작용
- ◆ C형 간염바이러스의 NS5A 단백질이 TGF- $\beta$  receptor I과의 상호작용을 통해 TGF- $\beta$  dependent signal pathway를 조절함
  - TGF- $\beta$  receptor I과 NS5A 단백질이 직접적으로 상호 작용함
  - TGF- $\beta$ 에 의해 유도되는 Smad3와 Smad4와의 complex형성을 NS5A가 억제함
- ◆ C형 간염바이러스의 replicon cell이 TNF에 의해 유도되는 NF- $\kappa$ B activity는 증가시키고 AP-1 activity는 감소시킴
  - TNF에 의해 유도되는 promoter activity가 replicon cell에서 증가함
  - TNF에 의한 IKK의 kinase activity와 NF- $\kappa$ B의 DNA binding activity가 replicon cell에서 증가됨
  - TNF에 의한 JNK activity와 AP-1의 DNA binding activity, ATF-2의 phosphorylation이 replicon cell에서 감소됨
  - Replicon cell에서 p53의 발현과 TNF에 의한 p53의 nuclear translocation이 감소로 인해 apoptosis가 억제됨
- ◆ C형 간염바이러스 NS5A 단백질이 TGF-beta receptor I과 상호 작용하여Smad2의 phosphorylation을 50% 억제함
- ◆ TGF- $\beta$ 에 의해 증가되는 p21의 발현이 NS5A에 의해 50% 감소됨
- ◆ NS5A가 Smad-DNA complex 형성을 50% 억제함
- ◆ TGF-beta receptor I의 kinase activity에는 NS5A가 전혀 영향을 주지 않음
- ◆ C형 간염바이러스의 control Huh7 cell과 replicon cell에서 total RNA를 추출하여 microarray analysis를 수행
  - 2배 이상 up-regulated genes : 840 genes
  - 2배 이상 down-regulated genes : 1,896 genes
  - 4배 이상 fold-changed genes : 142 genes
- ◆ C형 간염바이러스의 NS5A 단백질에 대한 transgenic mice의 cirrhotic liver 및 control mouse의 liver에서 total RNA를 추출하여 microarray analysis를 수행
  - 2배 이상 fold-changed genes : 743 genes
  - 4배 이상 fold-changed genes : 206 genes

## V. 연구개발결과의 활용계획

본 연구팀은 시스템 생물학 연구사업을 통해 전혀 활용된 바 없는 세포내 신호전달과정과 C형 간염바이러스간의 연계성을 찾아내어 좀 더 명확한 바이러스성 간염의 유발 mechanism을 밝히고자 한다. 이러한 연구가 성공적으로 진행된다면 C형 간염바이러스와 유사한 RNA 바이러스에 이러한 결과를 적용시켜 바이러스를 연구하는 학자들에게 암유발 mechanism에 대한 좀더 정확한 기전을 설명할 수 있으며 또한 세포생물학자들에게 있어서는 직접적인 실험을 수행하지 않고도 세포내 신호전달과정을 computer로 예상할 수 있어 타연구에 많은 응용이 이루어질 것으로 생각된다.

## S U M M A R Y

(영 문 요 약 문)

Hepatitis C virus (HCV) is an etiological agent of acute and chronic hepatitis and hepatocellular carcinoma (HCC). Due to the lack of an efficient cell culture system, both pathogenesis and molecular biology of HCV are not fully understood. To date, neither vaccine nor efficient therapy is not available for HCV. HCV is a single-stranded, positive-sense RNA virus of 9.6 kb in length and encodes a single polyprotein precursor of 3,010 amino acid residues. This polyprotein is processed by a combination of hostcellular signal peptidase and viral proteases into 4 structural and 6 nonstructural mature proteins. To understand the mechanisms of HCV-induced HCC, it is necessary to investigate the global changes of the systemic networks in cellular proteins in HCV-infected cells. For this purpose, major concerns are focused on the establishment of cellular protein networks that could be affected by HCV infection. These include the changes of levels in transcription, translation, posttranslational modification, cellular signal transduction, and cytokine networks. In the present study, we identified cellular proteins interacting with HCV proteins using yeast two-hybrid system and further confirmed them in baculovirus and mammalian expression systems. Furthermore, we identified viral and cellular proteins that directly involved in HCC using transgenic animal model. Using HCV replicon system, we also demonstrated that cellular TNF signaling was modulated by viral protein. In addition, we showed that cellular TGF signaling was inhibited through interaction of NS5A and T $\beta$ R1. These results suggest that NS5A may play a major role in chronic hepatitis and HCC. Most importantly, we performed the computational modeling, focusing on the simulation of TGF- $\beta$  signal transduction using parameters obtained from each stage of signaling cascades. In conclusion, we established cellular protein networks in HCV-infected cells and this system may be useful for the control of hepatic diseases by constructing quantitative data base system.

C O N T E N T S

(영 문 목 차)

Chapter 1. Summary of the research project . . . . . 8

Chapter 2. Current status of technical development in korea and foreign countries . . 10

Chapter 3. Contents and results of this project . . . . . 11

Chapter 4. Goal achievement and contribution to related fields . . . . . 22

Chapter 5. Application of the results of current study . . . . . 27

Chapter 6. Technical information obtained from foreign country during the  
course of this project . . . . . 28

Chapter 7. References . . . . . 29

목 차

제 1 장 연구개발과제의 개요 . . . . . 8  
\* 연구개발의 목적, 필요성 및 범위 등을 기술

제 2 장 국내외 기술개발 현황 . . . . . 10  
\* 국내·외 관련분야에 대한 기술개발현황과 연구결과가 국내·외 기술개발현황에서 차지하는 위치 등을 기술

제 3 장 연구개발수행 내용 및 결과 . . . . . 11  
\* 이론적, 실험적 접근방법, 연구내용, 연구결과를 기술

제 4 장 목표달성도 및 관련분야에의 기여도 . . . . . 22  
\* 연도별 연구목표 및 평가착안점에 입각한 연구개발목표의 달성도 및 관련분야의 기술 발전에의 기여도 등을 기술

제 5 장 연구개발결과의 활용계획 . . . . . 27  
\* 추가연구의 필요성, 타연구에의 응용, 기업화 추진방안을 기술  
\* 연구기획사업 등 사업별 특성에 따라 목차는 변경 가능함

제 6 장 연구개발과정에서 수집한 해외과학기술정보 . . . . . 28

제 7 장 참고문헌 . . . . . 29  
\* 보고서 작성시 인용된 모든 참고 문헌을 열거한다

## 제 1 장 연구개발과제의 개요

C형 간염바이러스는 급·만성 간염 및 간암을 유발하는 주된 병원체이다 (1-3). C형 간염 바이러스의 genome이 클로닝 (4)된지 16년이 지난 현재까지 C형 간염바이러스에 대한 *in vitro* 세포배양 기술이 확립되지 못하여 바이러스에 대한 병원성 및 분자생물학적 연구가 지연되고 있는 실정이며, 이로 인해 C형 간염바이러스의 감염을 예방할 수 있는 백신은 물론 치료제 또한 현재까지 개발되지 못한 상황이다. C형 간염바이러스에 감염된 감염자는 전 세계적으로 약 1억 7천만명 정도 되며 현재 우리나라의 경우도 1.5% 이상이 C형 간염바이러스에 감염된 것으로 추정되고 있다.

C형 간염바이러스의 genome은 plus-strand RNA로서 크기가 약 9.6 kb (4-7)이며 3,010개의 아미노산으로 구성된 하나의 polyprotein을 암호화한다. 이 polyprotein은 숙주세포의 signal peptidase와 바이러스의 단백질 분해효소에 의해 4개의 구조단백질과 6개의 비 구조단백질로 분리 된다. C형 간염바이러스의 RNA 복제는 genome을 template으로 하여 상보적인 negative-strand RNA를 만들고 다시 negative-strand가 template으로 되어 genomic RNA를 만들어 낸다. C형 간염바이러스는 6개 이상의 genotype과 30개 이상의 subtype으로 나누어지며, genome의 염기서열이 다양하고 바이러스 외피단백질 (envelope proteins)의 지속적인 돌연변이 (8-11)로 인하여 많은 quasi-species가 생성된다. C형 간염 바이러스의 만성 감염자는 타인에게 바이러스를 전파할 가능성이 높고, 감염자 자신의 경우도 만성간염, 간경변증, 간세포암과 같은 심각한 합병증을 유발할 위험성이 매우 높아 국민 건강에 심각한 위협성을 가지고 있다. C형 간염의 치료는 Interferon이나 Ribavirin등이 사용되고 있으나 일부 genotype에만 효과가 있고 치료를 중지하면 곧바로 재발하는 결점이 있다. 현재까지 C형 간염에 특이적으로 효과가 있는 치료제는 아직 개발되지 못하였다. 한국인의 간암에 의한 사망률이 세계 제 1위라는 심각성을 고려할 때 이에 대한 예방 및 치료 기술 개발이 절실히 요구된다.

System biology는 biology, mathematics, system science, statistical analysis등의 다양한 연구 분야의 연구 방법을 종합적으로 결합시킨 후 systemic한 접근방식을 이용하여 생물학 분야를 좀더 혁신적으로 발전시킬 수 있는 새로운 학문 분야이다. 현재까지 명확히 해석하기 어려운 생물학적 data를 컴퓨터를 통해 하나의 완전한 화상으로 modeling할 수 있다. 이러한 system biology는 cross-disciplinary science, 즉 여러 분야를 통합하여 system biology에 필요한 많고 다양한 실험방식 (DNA sequencer, oligonucleotide or DNA arrays, proteomics 등)을 창출해 냈다. 또한 인터넷을 통해 세계 각 지역에 있는 과학자들과의 즉각적인 대화가 용이해 졌고 많은 양의 데이터를 저장하고 송출할 수 있게 되었다. 특히 microarray를 이용해 얻은 많은 양의 data가 분석되고, 생물의 genome이 해석됨으로써 급속도로 생명공학이 발전하고 있다. System biology에서 biological information (DNA, RNA, protein, protein interactions, biomodules, cells, tissues 등)은 각각의 요소들 (specific genes, proteins)을 가지고 있으며, 그리고 이들 각각은 다른 요소와 연관성을 가지고 있다. 즉, biological information과 그에 해당하는 각각의 요소들 사이의 정보를 전체적인 하나의 system model로써 얻을 수 있다.

간암은 주로 간염바이러스에 의하여 유발된다. 따라서 바이러스성 간암을 제어하기 위하여 바이러스가 숙주세포의 간세포 수용체에 binding한 후 일어나는 모든 일을 이해하여야 된다. 즉, HCV가 간세포에 감염함으로써 발생하는 세포 단백질의 systemic network 변화에 주목해야한다. 이를 위하여 C형간염바이러스가 영향을 미치는 세포단백질의 protein network을 구축하고 이들의 transcription, translation level의 변화는 물론, posttranslational modification, cellular signal transduction, cytokine network의 변화 등을 연구해야 한다. 바이러스는 숙주세포의 protein network을 바이러스에 유리하도록 조절 하기 때문에 감염되지 않은 세포와 감염된 세포 또는 간암간세포와 정상인 간세포의 단백질의 변화를 비교하고 이들 변화를 정량적인 데이터베이스로 구축함으로써 질병 (간 질환)을 제어할 수 있는 시스템을 확립할 수 있다.

## 제 2 장 국내외 기술개발 현황

지난 2000년 genome project가 발표된 후 "System Biology"는 새로운 학문 분야로 자리 잡았으며 2002년부터 국제적인 학술대회 뿐만 아니라 research paper를 전문적으로 출판하는 "Systems Biology"라는 새로운 국제적인 저널이 생겨났다. 이렇듯 현재 system biology에 대한 관심이 국내·외적으로 점차 증가하고 있으며 관련 논문의 발표 수가 계속적으로 증가하고 있고 실험을 수행하는데 필요한 기술 또한 많은 발전을 거듭하고 있는 실정이다. 최근 10년 동안 high-throughput method들을 이용한 실험을 통해 생물학적으로 배경이 다른 요소들의 세포내 발현 정도 차이를 찾아내고 이를 통해 그 기능을 밝히고자 하는 기술들이 genomics, proteomics, bioinformatics, metabolomics등의 이름으로 크나큰 발전을 해왔다. 국외적으로 미국, 유럽, 일본 등에서 system biology에 대한 많은 관심을 기울이고 있다. 현재 이러한 나라들의 system biology에 대한 연구는 여러 분야에서 많은 결과를 산출해 내고 있다. 특히, 미국의 경우 Stanford, Cambridge, Virginia, MIT, Harvard, Caltech, UCLA등 여러 대학에서 20여개 이상의 팀이 이 분야에 대한 연구를 수행하고 있으며, 유럽의 경우 EMBL (The European Molecular Biology Laboratory)을 통해 computational biology, functional genomics, systems biology분야에서 많은 연구를 수행하고 있다. 또한 일본의 경우 Sony Computer Science 연구소와 Kitano 그룹에서 생명현상과 computer science를 접목하여 인공생명체를 구현하고 생명현상을 simulation하는데 중점을 두고 연구를 수행하고 있다. 국내에서는 KAIST의 이상엽 박사팀이 prokaryote에서 metabolic pathway에 대한 system biology를 연구하고 있으며 서울대학교의 조광현 박사팀은 수학적인 접근 방식을 통해 세포내 신호전달과정의 simulation에 대한 연구를 수행하고 있다. 현재 C형 간염바이러스를 이용한 system biology에 대한 연구 결과는 전혀 없는 상태이며 최근 들어 cDNA microarray, proteomics등을 이용한 실험 결과를 통해 C형 간염바이러스에 의해 유발되는 간질환과 관련 있는 유전자 및 단백질을 스크리닝하고 있다. 또한 본 연구팀이 수행하고 있는 TGF-beta signal transduction pathway에 대한 systemic한 접근이 이루어지지 않고 있는 실정이다. 본 연구팀이 수행하고 있는 연구의 결과가 현재 완벽한 상태는 아니지만 현재 여러 parameter들을 이용한 simulation 결과를 분자생물학적인 실험을 통해 확인한 결과 어느 정도 일치하는 결과를 얻을 수 있었다. 이러한 연구 결과는 기술적으로 아직 미흡한 면이 있지만 지속적인 연구를 수행하여 이를 보완한다면 국내뿐만 아니라 국외적으로 이러한 연구를 수행하는 연구자들에게 기초 자료로 사용될 수 있을 것으로 사료된다.

### 제 3 장 연구개발수행 내용 및 결과

#### 1) 이론전, 실험적 접근 방법

- (1) HCV viral protein의 expression vector 제조
- (2) Yeast two-hybrid system
- (3) 염기서열 분석
- (4) Cell culture와 transfection experiment
- (5) Transient transfection과 reporter gene assay
- (6) *In vitro* kinase assay (NF- $\kappa$ B, SAPK/JNK, p38 MAP kinase)
- (7) Apoptosis assay
- (8) DNA fragmentation assay
- (9) TUNEL assay
- (10) Electrophoretic Mobility Shift Assays (EMSAs)
- (11) Immunoblot analysis
- (12) Recombinant baculovirus 제조
- (13) 세포 및 바이러스 배양
- (14) 단백질 정제
- (15) GST binding assay and co-immunoprecipitation
- (16) Northern blot hybridization
- (17) 간접면역형광법
- (18) Electron Microscopy
- (19) Two-dimensional gel eletrophoresis
- (20) Cytokine-related cDNA microarray
- (21) Cell proliferation assay
- (22) *In vivo* phosphorylation experiments
- (23) Metabolic labeling experiments
- (24) Histology, Immunohistochemistry
- (25) RNase protection assay (RPA)
- (26) RT-PCR
- (27) Microinjection, generation of HCV transgenic mice
- (28) Southern blot hybridization
- (29) FACS analysis



## 2) 연구 결과

1. Yeast two-hybrid system, baculovirus expression system, 그리고 mammalian system 에서 사용하기 위해 C형 간염바이러스의 단백질을 발현하는 부위를 expression vector에 클로닝하였다 (Fig. 1).

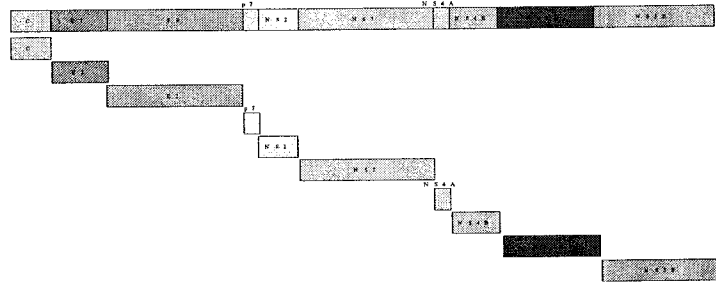


Fig. 1. Schematic diagram of the constructs of HCV viral proteins.

2. C형 간염바이러스의 replicon cell을 이용하여 microarray analysis를 수행한 결과 2배 이상 up-regulated gene 840개, 2배 이상 down-regulated gene 1,896 개, 그리고 4배 이상 fold-changed gene 142개를 확보하였다 (Fig. 2, 3). 이러한 유전자들에서 cytokine-related gene은 약 200여개 정도 되며 이들 중 약 50개정도의 유전자에 본 연구팀이 관심을 가지고 있다 (Table 1, 2).

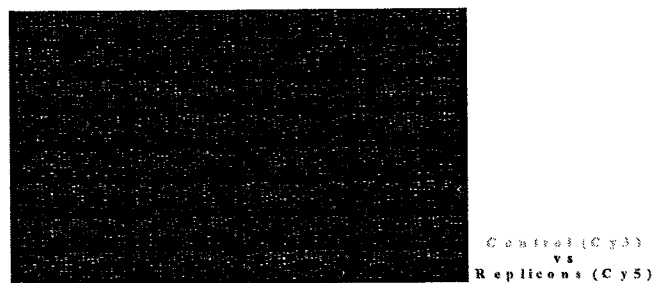


Fig. 2. Scan image of microarray analysis for HCV replicon cells.

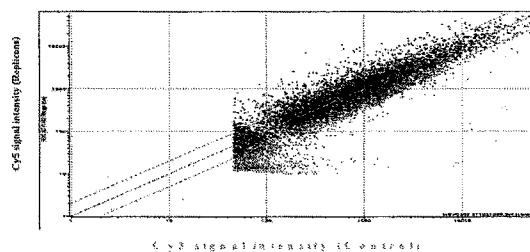


Fig. 3. Scatter plot of microarray analysis for HCV replicon cells.

Table 1. Physiological 2-fold up-regulated cytokine-related genes in HCV replicon cells

Gene description	Fold change	Gene bank #
Interleukin 23	-2.00	AA418747
TNF receptor-associated factor 3	-2.00	AA504259
Rho GTPase activating protein 1	-2.01	AA427934
Activated p21 kinase	-2.18	A1349439
Interferon regulatory factor 1	-2.02	AA478043
Protein tyrosine kinase 7	-2.07	AA453789
Calcium/calmodulin-dependent protein kinase 1	-2.1	H29322
Toll-like receptor 13	-2.12	A1339155
Transforming growth factor-beta receptor III	-2.12	H62473
Janus kinase 3	-2.19	AA490909
Tumor protein p53	-2.22	R39356
MAPK3	-2.22	AA559383
Toll-like receptor 10	-2.31	AA287325
MAPK10	-2.32	A1961583
Rho interacting protein 3	-2.31	AA497027
Caspase 9	-2.38	AA281152
Interferon regulatory factor 6	-2.48	AA043343
CDC42 effector protein 1	-2.51	A1738884
Interleukin 24	-2.51	AA281655
CASP8 and FADD-like apoptosis regulator	-3.03	AA453766
Nitric oxide synthase 3	-3.08	AA284967
Interleukin 2 receptor, beta	-3.32	AA067104
Interleukin 1 receptor, type II	-3.67	H23284
Phosphoinositide-3-kinase, p85- $\alpha$	-3.73	R45961
Interleukin 13	-4.28	A1208823
Transforming growth factor-beta 2	-4.94	AA233739
Interleukin 7	-5.57	AA626701
Death-associated protein kinase 3	-6.16	AA975739
Interferon-gamma	-14.71	AA969594

Table 2. Physiological 2-fold down-regulated cytokine-related genes in HCV replicon cells

Gene description	Fold change	Gene bank #
Protein tyrosine phosphatase	15.98	H24416
Chemokine (C-X-C motif) ligand 4	8.8	AA935273
Ras homolog gene family	5.1	AA443301
Growth factor receptor-bound protein 14	5.09	R14204
Ras homolog gene family 2	4.99	A1649917
Insulin-like growth factor 2	4.64	N54594
Dyckian tyrosine kinase 3 beta	4.37	R93912
TNF- $\alpha$ induced protein 3	4.0	A1963014
Interferon gamma receptor 1	3.76	AA449440
Survivin/Tyrosine kinase 3B	3.47	AA523346
Interferon gamma receptor 2	3.24	AA381497
Chemokine (C-C motif) receptor-like 1	3.12	A1288445
Interleukin 4 signal transducer	2.99	A1349316
TNF receptor-associated factor 5	2.97	AA102634
Protein kinase, cAMP-dependent, gamma 4	3.5	A1824842
MAPK1 interacting protein 1	3.1	AA474988
MAPK5	3.02	W49419
Protein-associated factor 1	3.02	AA486789
SH3 domain CRB2-like endophilin B1	3.3	AA465147
Interferon regulatory factor 5	3.24	N30372
Interferon-stimulated transcription factor 3	3.19	AA293389
CRB3-associated binding protein 2	3.18	AA345573
MAPK9	3.09	AA357286
Activin A receptor	3.08	AA336892

3. NS5A 단백질에 대한 transgenic mice의 liver를 이용해 간의 morphology를 관찰한 결과 liver에서 tumor가 형성됨을 관찰하였다 (Fig. 4). Tumor 조직을 이용해 histological analysis를 수행한 결과 mice의 liver를 section하여 H&E staining한 결과 정상적인 mice와는 달리 간세포 핵의 크기가 다양하게 나타남을 관찰하였고, 또한 glycogen storage foci가 형성되어 있음을 관찰하였다 (Fig. 5). 또한 transgenic mice의 liver를 이용해 microarray analysis를 수행하였다. 그 결과 2배 이상 fold-changed gene 743개, 4배 이상 fold-changed gene 206개를 확보하였고 (Fig. 6), 특히 4배 이상 fold-changed gene들 중 약 23개의 유전자에 관심을 가지고 있는데 이러한 유전자들은 signal transducer, apoptosis regulator, cell cycle regulator, fibrosis, cancer, chaperone 등의 기능을 가지고 있다 (Table 2).

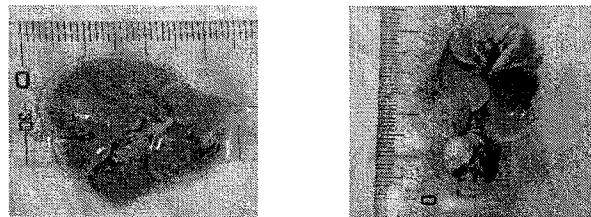


Fig. 4. Gross detection of hepatic tumors in transgenic mouse at 15 months of age (right). Liver

of normal mouse is shown as a control (left).

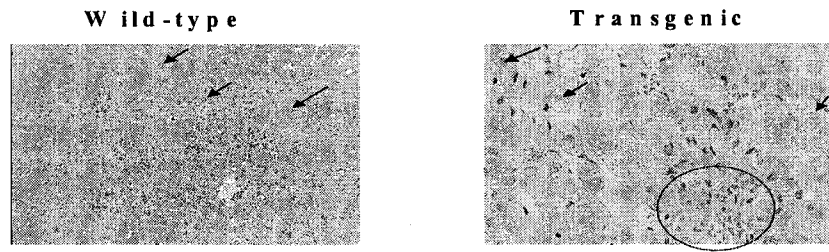


Fig. 5. Pathological changes in the liver of HCV transgenic mouse. Hepatocytes of transgenic mouse are histologically abnormal and are infiltrated with mononuclear cells as shown in red circle.

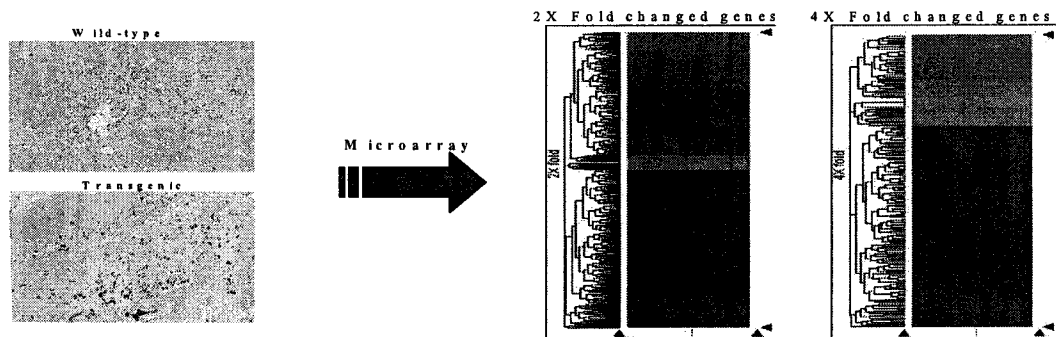


Fig. 6. Differential expression profiles of many genes in cirrhotic liver of transgenic mice. H&E stained liver sections of cirrhotic stages in transgenic mice (Left panels). Microarray data indicate different patterns of gene expression in transgenic mouse.

Table 3. Physiological 4-fold changed genes in cirrhotic liver of HCV NS5A transgenic mice

Gene description	Gene ontology	Fold change	Gene bank #
Interleukin 25	Signal transducer	-5.19	NM_080837
Fibronin 2	Signal transducer	-5.24	NM_010181
FURF, TP53 apoptotic effector	Apoptotic regulator	-6.06	NM_025032
Chemokine (C motif) receptor 1	Signal transducer	-4.43	NM_011978
Contactin 1	Cell adhesion	-4.86	NM_007727
G0/G1 switch gene 2	Cell cycle	-16.01	NM_008059
Vavln 1	Cell growth	-21.9	NM_011704
MAPK6	Signal transducer	6.84	NM_015806
Transforming growth factor-beta receptor 2	Signal transducer	6.04	AK_003861
Interleukin 1 beta	Signal transducer	4.19	NM_008361
Interleukin 10 receptor beta	Signal transducer	4.11	NM_008349
Profilin 1	Chaperone	4.4	NM_011072
Septin 4	Cell cycle	5.1	NM_011129
Aldolase	Immunity	13.9	NM_013459
Matrix metalloproteinase 12	Cancer	20.07	NM_008605
Fibrinogen, type V, alpha 2	Fibrinolysis	19.26	NM_007737
Nidogen 2	Structural protein	17.6	NM_008695
Insulin-like growth factor binding protein 7	Signaling	13.08	NM_008048
Caspase 4	Apoptotic regulator	12.38	NM_007609
Caspase 12	Apoptotic regulator	11.85	NM_009808
Cellular 13	Cell adhesion	41.29	NM_015707
Moesin	Chaperone	45.4	NM_010833
Insulin-like growth factor binding protein 1	Signaling	34.4	NM_008341

4. TNF에 의해 유도되는 NF- $\kappa$ B activation을 reporter gene assay를 통해 검증하였다. 그 결과 C형 간염바이러스의 replicon cell에서 TNF에 의해 유도되는 NF- $\kappa$ B activity가 대조군에 비해 증가함을 관찰하였으며, TNF에 영향 없이도 constitutively하게 activation되어 있음을 확인하였다 (Fig. 7).

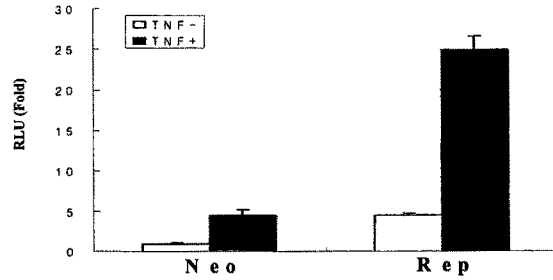


Fig. 7. Activation of TNF- $\alpha$ -induced NF- $\kappa$ B activity in HCV subgenomic replicon cells. HCV replicon cells were co-transfected with NF- $\kappa$ B-Luc and pCH110 reporter plasmids. At 24 h after transfection, cells were either left untreated or treated with human TNF- $\alpha$  (10 ng/ml) for 6 h, and then reporter activity was measured by luminometer.

이러한 결과를 확인하기 위해 NF- $\kappa$ B에 대한 EMSA를 수행한 결과 reporter gene assay와 마찬가지로 replicon cell에 있어서 NF- $\kappa$ B의 DNA binding activity와 IKK $\beta$ 에 대한 kinase activity가 증가됨을 관찰하였다 (Fig. 8A, B).

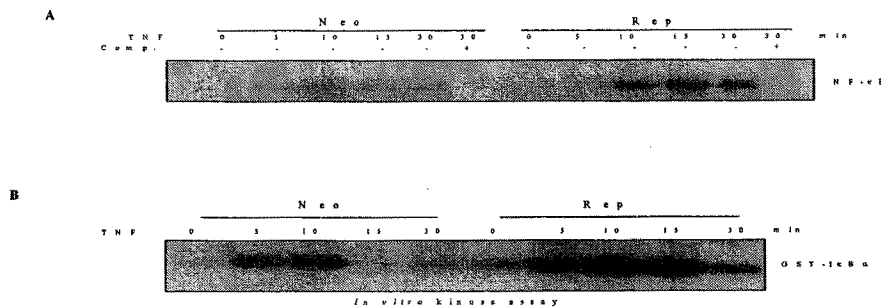


Fig. 8. HCV subgenomic replicon cells activate TNF- $\alpha$ -induced kinase activity of IKK- $\beta$  and NF- $\kappa$ B DNA binding activity. (A) HCV replicon cells were incubated with TNF- $\alpha$  (20 ng/ml) for the indicated times. Nuclear extracts were prepared and assessed for activated NF- $\kappa$ B by EMSA with  $\gamma$ - $^{32}$ P-labeled single-stranded immunoglobulin  $\kappa$ B oligonucleotide. (B) Effects of HCV replicon cells on kinase activity of endogenous IKK- $\beta$  proteins. HCV replicon cells were incubated with TNF- $\alpha$  (20 ng/ml) for the indicated times. After TNF treatment, cell lysates were precipitated with anti-IKK- $\beta$  antibody, and then endogenous IKK- $\beta$  kinase activity was determined using GST-I $\kappa$ B $\alpha$  (1-54aa) as a substrate.

5. TNF에 의해 유도되는 JNK 활성화와 AP-1의 DNA binding activity가 replicon cell에서 감소됨을 관찰하였고 (Fig. 9) ATF-2의 phosphorylation이 감소되어 있음을 확인하였다 (Fig. 10). 이러한 결과들은 C형 간염바이러스가 세포내 단백질의 kinase activity를 조절함으로써 transcription factor의 DNA binding activity에 영향을 미침을 증명하였다.

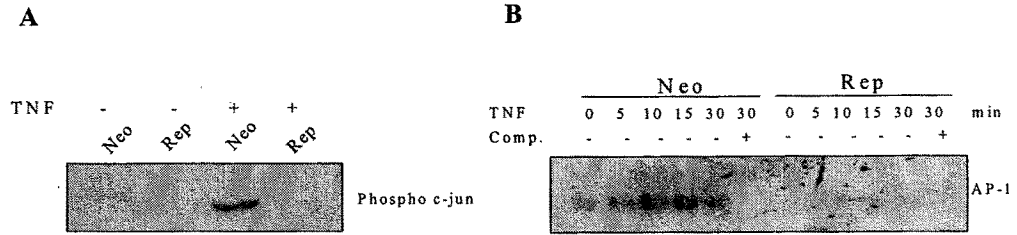


Fig. 9. Inhibition of TNF- $\alpha$ -induced JNK activity and AP-1 DNA binding activity in HCV replicon cells. (A) HCV replicon cells were either left treated or treated with TNF- $\alpha$  (20 ng/ml) for 15 min. Using cell lysates, JNK activity was determined using JNK assay kit. (B) HCV replicon cells were stimulated with TNF- $\alpha$  (20 ng/ml) for the indicated times. Nuclear extracts were assessed for AP-1 DNA binding activity by EMSA with  $\gamma$ - $^{32}$ P-labeled 21-mer double-stranded AP-1 oligonucleotide.

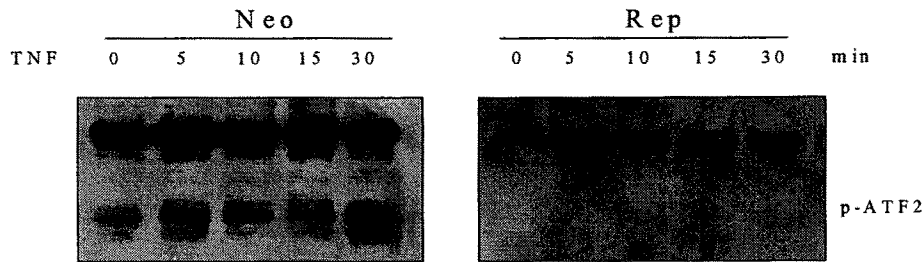


Fig. 10. JNK-mediated ATF-2 phosphorylation is inhibited by HCV viral proteins. HCV replicon cells were stimulated with TNF- $\alpha$  (20 ng/ml) for the indicated times. Nuclear extracts were prepared and immunoblotted with anti-phospho-ATF-2 antibody.

6. Replicon cell에서 p53의 발현이 감소되어 있음을 관찰하였고, 또한 TNF에 의해 유도되는 p53의 nuclear fraction으로의 translocation이 replicon cell에서 억제되어 있음을 관찰하였다 (Fig. 11). 이러한 결과는 C형 간염바이러스가 p53의 발현을 조절하고, 또한 TNF에 의해 유도되는 p53의 nuclear로의 translocation을 억제함으로써 p53에 의한 apoptosis를 억제하여 만성 감염 및 간암으로의 전이에 중요한 역할을 할 것으로 사료된다.

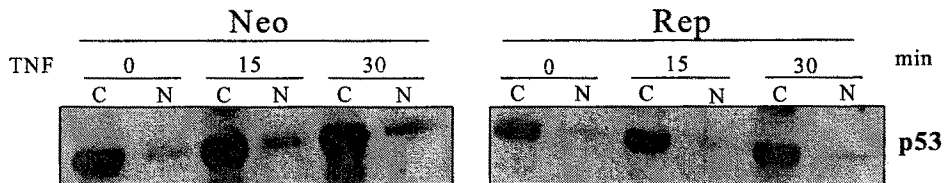


Fig. 11. HCV viral proteins alter the TNF- $\alpha$ -induced subcellular localization of p53. Both neo control and HCV replicon cells were treated with TNF- $\alpha$  (20 ng/ml) for the indicated times and cells were fractionated into cytoplasmic and nuclear fractions. Both extracts were subjected to immunoblot analysis with anti-p53 antibody.

7. TNF/AcD를 함께 처리하였을 경우 대조군에 비해 PARP의 cleavage된 양이 감소한 것

을 관찰하였다 (Fig. 12). 이러한 C형 간염바이러스의 replicon cell은 TNF/AcD에 의해 유도되는 apoptosis에 저항성을 갖게 됨으로써 만성 간염 및 간암으로 전이됨을 제시한다.

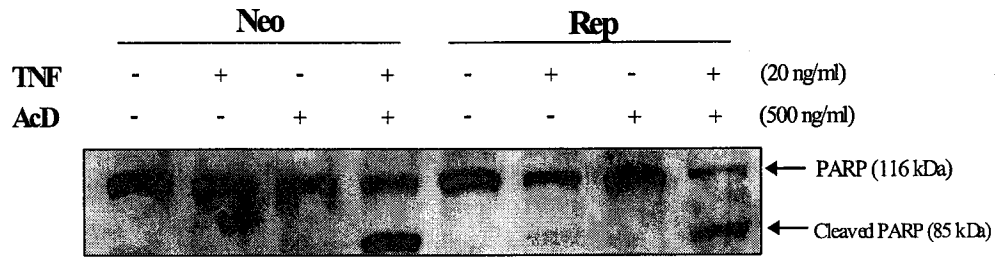


Fig. 12. HCV viral proteins inhibit TNF- $\alpha$  and actinomycin (AcD)-induced apoptosis. HCV replicon cells were treated with TNF- $\alpha$  and AcD for 12 h. Apoptosis was determined by immunoblotting cellular proteins with anti-PARP monoclonal antibody.

8. NS5A 단백질과 TGF- $\beta$  receptor I (T $\beta$ R-I)과의 상호 작용을 확인하였다 (Fig. 13A, B, C). Hepatoma cell line에 plasmid를 co-transfection한 후 각각에 해당하는 항체로 단백질을 staining하였다. 그 결과 hepatoma cell에서 viral 단백질인 NS5A와 세포내 단백질인 T $\beta$ R-1이 cytosol에서 co-localization함을 증명하였다 (Fig. 14).

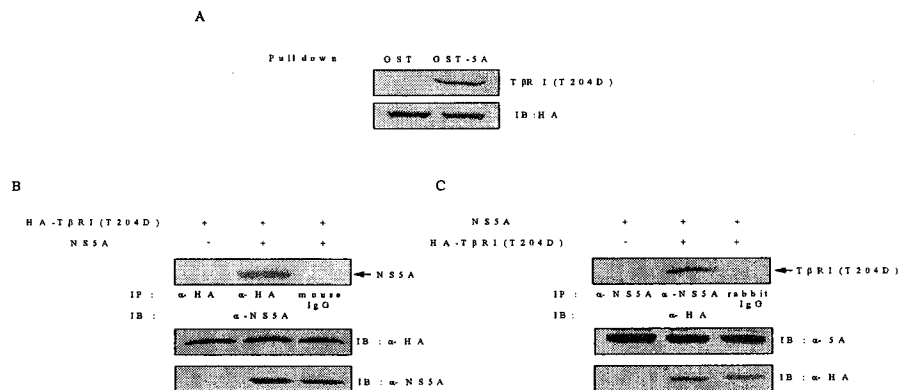


Fig. 13. NS5A protein interacts with Type I TGF- $\beta$  receptor both *in vitro* and *in vivo*. (A) Huh7 cells were transfected with HA-T $\beta$ R-I (T204D) expressing plasmid. At 36 h after transfection, cells were precipitated with either GST or GST-NS5A fusion protein purified from *E. coli*. Bound proteins were detected by western blot analysis using anti-HA monoclonal antibody (top panel). (B) Using recombinant vaccinia virus expressing T7 RNA polymerase, both NS5A and HA-T $\beta$ R-I (T204D) proteins were co-expressed in Cos7 cells. At 12 after transfection, cell lysates were immunoprecipitated with either anti-HA monoclonal antibody or normal mouse IgG, and the immunoprecipitates were analyzed by western blot analysis using anti-NS5A polyclonal antibody (top panel). HA-T $\beta$ R-I (T204D) and NS5A proteins were verified using the same cell lysates by western blotting with either anti-HA monoclonal antibody (middle panel) or anti-NS5A polyclonal antibody (bottom panel). (C) Reciprocally,

cell lysates used in B were immunoprecipitated with either anti-NS5A polyclonal antibody or normal rabbit IgG, and the immunoprecipitates were analyzed by western blot analysis using anti-HA monoclonal antibody (top panel). NS5A and HA-T $\beta$ R-I (T204D) proteins were verified using the same cell lysates by western blotting with either anti-NS5A polyclonal antibody (middle panel) or anti-HA monoclonal antibody (bottom panel).

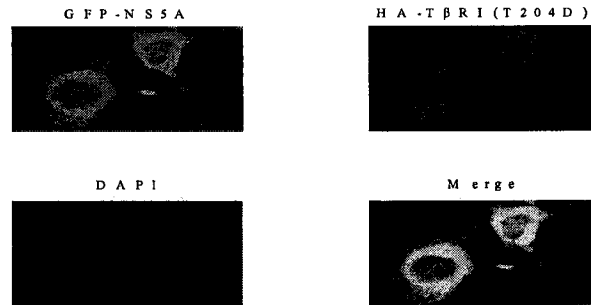


Fig. 14. Co-localization of NS5A and Type I TGF- $\beta$  receptor in Cos7 cells. Both GFP-NS5A and HA-T $\beta$ R-I (T204D) plasmids were co-transfected into Cos7 cells. At 36 h after transfection, cells were fixed in 4% paraformaldehyde, and immunofluorescence assay was performed using an anti-HA monoclonal antibody and tetramethylrhodamine isothiocyanate-conjugated goat anti-mouse IgG to detect T $\beta$ R-I (red). Dual staining showed co-localization of NS5A and HA-T $\beta$ R-I (T204D) as yellow fluorescence in the merge. Cells were counterstained with 4',6'-diamidino-2-phenylindole (DAPI) to label nuclei.

9. TGF- $\beta$  receptor I (T $\beta$ R-I)과 TGF- $\beta$ 에 의해 유도되는 Smad2 (endogenous & exogenous)의 phosphorylation이 NS5A 단백질에 의해 억제됨을 관찰하였으며 (Fig. 15A, B) NS5A 단백질을 transiently transfection한 cell과 stable하게 발현하는 cell을 이용하여 TGF- $\beta$ 를 time-dependent하게 처리한 후 endogenous Smad2의 phosphorylation을 관찰한 결과 Smad2의 phosphorylation이 NS5A에 의해 억제됨을 관찰하였다 (Fig. 16A, B).

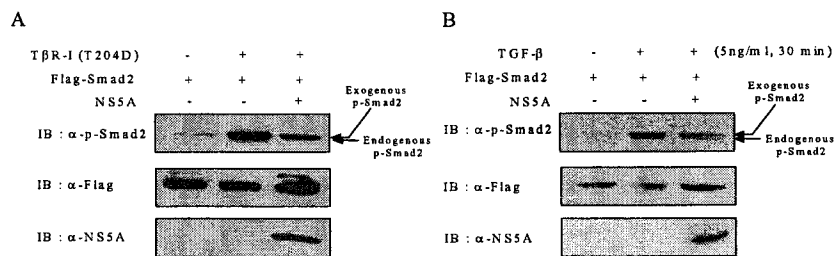


Fig. 15. NS5A protein inhibits TGF- $\beta$ -stimulated phosphorylation of Smad2. (A) Huh7 cells were transfected with constitutively activated HA-T $\beta$ R-I (T204D), Flag-Smad2, and either empty vector or NS5A. Equal amounts of cell lysates were separated by SDS-PAGE, and detected by western blot analysis using anti-phospho-Smad2 antibody (top panel). (B) Huh7 cells were transfected with Flag-Smad2 expressing plasmid either in the absence or presence of NS5A. At 36 h after transfection, cells were either left untreated or treated with TGF- $\beta$  (5 ng/ml) for 30 min, and then detected by western blot analysis using anti-phospho-Smad2

antibody (top panel).

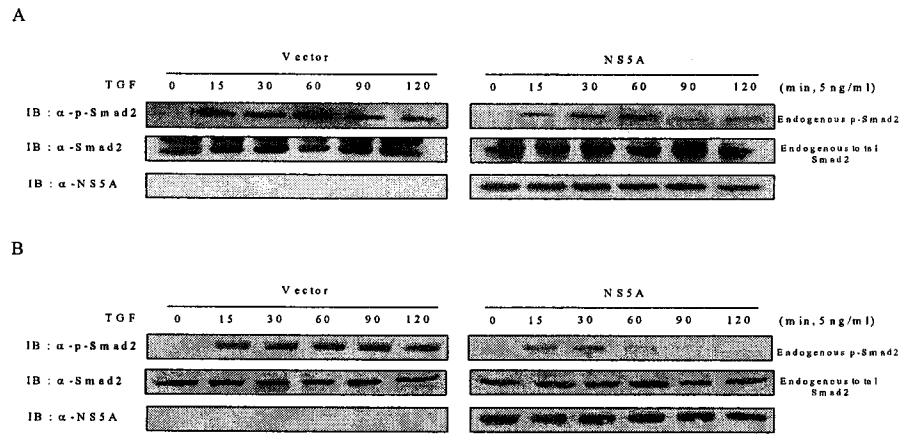


Fig. 16. NS5A protein inhibits TGF- $\beta$ -stimulated phosphorylation of Smad2 in a time-dependent manner in Huh7 cells expressing NS5A transiently (A) and stably (B). At 36 h after transfection, cells were treated with TGF- $\beta$  (5 ng/ml) for the indicated times, and the phosphorylation level of endogenous Smad2 was detected by phospho-Smad2 antibody (top panels).

10. NS5A 단백질이 TGF- $\beta$ 에 의해 유도되는 Smad3와 Smad4의 complex형성을 억제함을 관찰하였으며 (Fig. 17) 또한 Smad-DNA complex 형성도 억제됨을 확인하였다 (Fig. 18). 이러한 결과는 C형 간염바이러스의 NS5A 단백질이 세포내 단백질인 T $\beta$ R-I과 직접적으로 상호 작용함으로써 TGF- $\beta$ 에 의해 유도되는 signal transduction pathway를 억제할 것으로 사료된다.

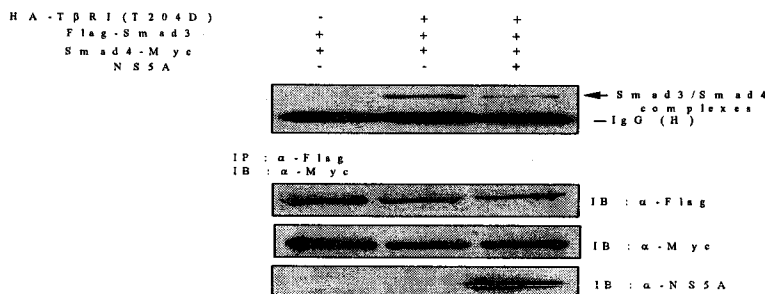


Fig. 17. NS5A protein abrogates T $\beta$ R-I-induced heterodimerization of Smad3 with Smad4. Huh7 cells were co-transfected with Flag-Smad3, Smad4-Myc, constitutively activated HA-T $\beta$ R-I (T204D), and either empty vector or NS5A. Cell lysates were immunoprecipitated with anti-Flag monoclonal antibody, and then complex formation between Smad3 and Smad4 was detected by western blot analysis using anti-Myc monoclonal antibody (top panel).



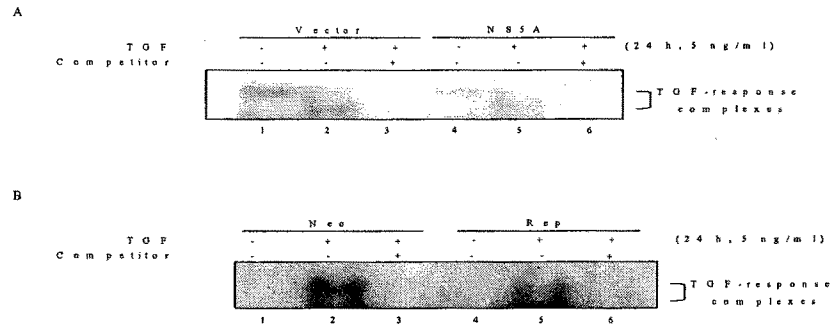


Fig. 18. NS5A inhibits the formation of the Smads-DNA complex. Huh7 cells stably expressing NS5A (A) and HCV subgenomic replicon cells (B) were stimulated with human TGF- $\beta$  for 24 h, and nuclear extracts were prepared and assessed for the formation of activated Smads protein and DNA.

11. TGF- $\beta$ 를 dose-dependent하게 처리한 후 endogenous p21의 발현을 관찰한 결과 NS5A가 발현될 경우 p21의 발현이 감소함을 확인하였다 (Fig. 18B). 이러한 결과는 TGF- $\beta$ 에 의해 유도되는 p21의 발현을 NS5A 단백질이 감소시킴으로써 TGF- $\beta$ 에 의해 유도되는 cell cycle arrest를 억제하여 만성 감염 및 간암으로의 전이를 유도하게 될 것으로 사료된다.

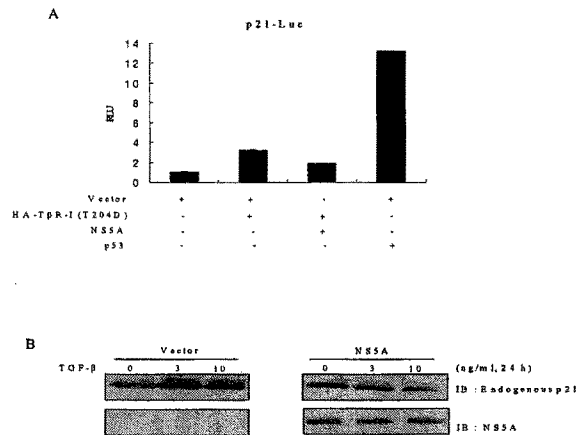


Fig. 19. TGF- $\beta$ -induced p21 expression was decreased by NS5A protein. (A) Huh7 cells were transfected with 2  $\mu$ g of expression plasmids, 0.1  $\mu$ g p21-Luc, and 0.1  $\mu$ g of a reference plasmid pCH110, respectively. At 36 h post-transfection, cells were harvested and luciferase activity was measured. (B) Effects of TGF- $\beta$  on p21 expression. Huh7 cells were transiently transfected with either empty vector or NS5A expression plasmid. At 24 h after transfection, cells were treated with TGF- $\beta$  in a dose-dependent manner. p21 was detected by immunoblotting with anti-p21 antibody.

12. TGF- $\beta$  signal transduction pathway를 computer science와 연계하여 computational model을 세우고자 우선적으로 TGF- $\beta$ 에 대한 graphical model을 enzyme kinetics에 기초하여 modeling하였다 (Fig. 20). 다음으로 각각의 단계별로 선정된 parameter value를 이용하여 signal transduction에 대한 simulation을 수행하였다. 그 결과 몇 가지 value에서

약 24시간에서 수렴하는 정도의 값을 얻을 수 있었다 (Fig. 21). 이러한 결과를 molecular level에서 확인하고자 Huh7 cell line에 TGF- $\beta$ 를 time-dependent하게 처리한 후 Smad-DNA의 binding affinity를 측정된 결과 역시 24시간까지 계속 증가하다가 그 이후로 감소함을 알 수 있었다 (Fig. 22). 이러한 결과는 computer를 이용한 simulation 결과와 실제 세포에서 일어나는 신호전달의 과정이 유사함을 증명하는 것이다.

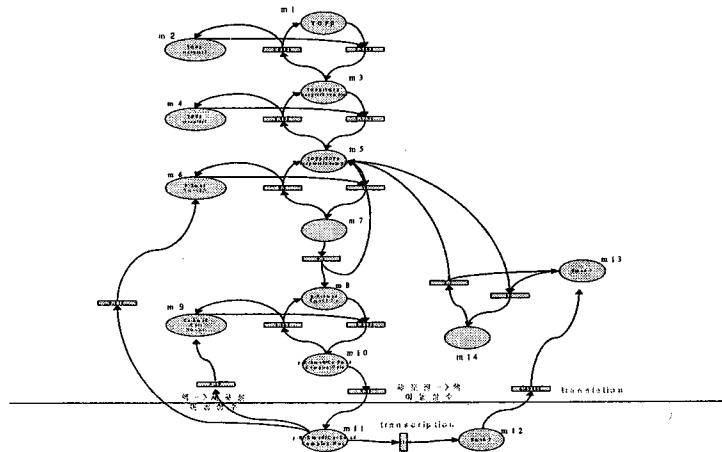


Fig. 20. Graphical model of TGF-beta signal transduction pathway

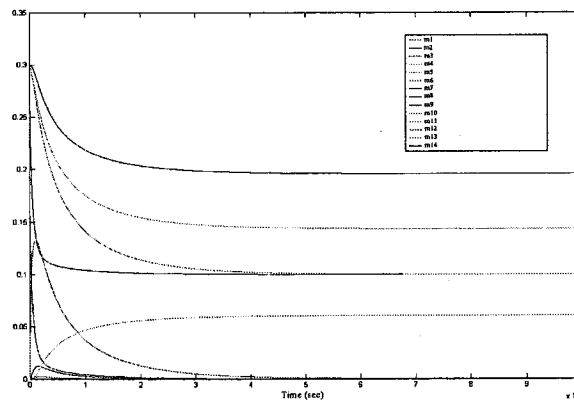


Fig. 21. Simulation result of the TGF-beta signal transduction pathway

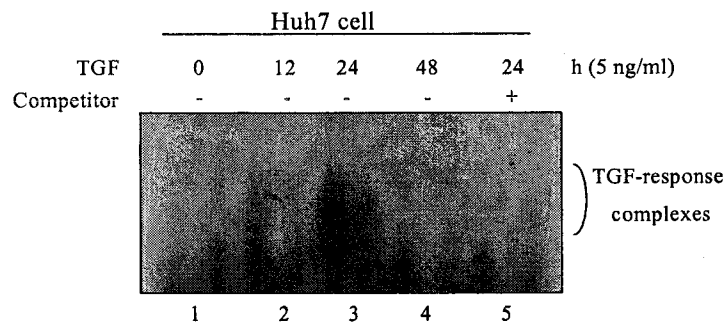


Fig. 22. The formation of Smads-DNA complex in Huh7 cell line for indicated times

## 제 4 장 목표달성도 및 관련분야에의 기여도

1차년도 (2003)

- ◎ 간세포에서 C형 간염바이러스의 viral 단백질과 상호 작용하는 간세포 단백질의 탐색
  - C형 간염바이러스의 viral 단백질을 발현하는 expression vector 제조
  - 각 viral 단백질의 발현 유·무 및 발현 정도를 포유동물세포, 곤충세포, yeast에서 확인
  - Yeast-two hybrid system 및 co-immunoprecipitation assay등을 이용하여 viral 단백질과 상호 작용하는 간세포 단백질 확인
- ◎ C형 간염바이러스의 replicon cell을 이용한 microarray data를 통해 C형 간염바이러스에 의해 영향 받는 세포내 단백질 확인
  - 2배 이상 up-regulated genes :840 genes
  - 2배 이상 down-regulated genes :1,896 genes
  - 4배 이상 fold-changed genes :142 genes
- ◎ C형 간염바이러스의 NS5A 단백질에 대한 transgenic mice를 이용한 microarray data를 통해 간암과 관련 있는 세포내 단백질 확인
  - 2배 이상 fold-changed genes :743 genes
  - 4배 이상 fold-changed genes :206 genes
- ◎ C형 간염바이러스의 replicon cell을 이용하여 cytokine에 의해 유도되는 세포내 신호전달과정에 viral 단백질이 미치는 영향 분석
  - TNF에 의해 유도되는 세포내 신호전달과정을 C형 간염바이러스가 조절함 (transcription factor인 NF-kB의 활성은 증가 시키고 AP-1의 활성은 감소시킴)
  - TGF에 의해 유도되는 신호전달과정이 억제됨
  - TNF에 의한 apoptosis를 억제함
  - Tumor suppressor인 p53의 발현이 낮고 TNF에 의한 핵내로의 translocation이 감소함
- ◎ C형 간염바이러스의 NS5A 단백질이 TGF에 의해 유도되는 세포내 신호전달을 억제함
  - NS5A 단백질이 TGF-beta receptor I과 상호 작용함
  - Smad 단백질의 phosphorylation이 억제됨
  - Smad 단백질의 핵내로의 translocation이 억제됨
  - Smad-DNA complex 형성이 감소함
  - TGF에 의한 CDK inhibitor인 p21의 발현이 감소함

2차년도 (2004)

- ◎ Enzyme kinetics를 기초로 하여 TGF-beta signal transduction에 대한 graphical model 정립
  - 여러 가지 parameter value 산출
  - 산출한 parameter value를 이용하여 신호전달과정에 대한 simulation 시도
  - 분자 생물학적인 실험을 통해 simulation과 실제 실험 결과와 유사함을 확인
- ◎ Cytokine과 C형 간염바이러스의 viral 단백질을 input으로 하고 HCV-associated marker gene을 output으로 하여 이에 대한 단백질 network를 computer science와 협동으로 modeling함
- ◎ Cytokine과 관련된 C형 간염바이러스가 유발하는 병인성에 대한 mechanism을 computer science와 협력하여 새로운 결과를 창출해냄

- Computer science를 통한 dynamic한 simulation 결과를 통해 cytokine에 의해 유발되는 signal transduction의 작용을 확인할 수 있음
- C형 간염바이러스의 viral protein이 cytokine에 의해 유발되는 signal transduction에 미치는 영향을 simulation을 통해 예측함

## 1) 연구개발의 최종목표

- o HCV viral protein과 상호작용 하는 cellular protein의 발굴 및 특성 규명
- o HCV viral protein에 의하여 유도되는 cellular protein의 변화 및 수식의 정량화
- o 변화된 단백질의 정량을 수학적으로 data base 화함
- o data base화된 단백질의 network 구축
- o data base system을 통한 질병 유발 기전 model 구축
- o 시뮬레이션 및 실험을 통한 검증
- o HCV 감염에 의한 간질환 제어 시스템 확립

궁극적인 목표는 C형간염바이러스의 viral protein에 의하여 조절되는 세포단백질의 정량적인 변화를 수학적인 데이터베이스로 입력하여 단백질네트워크를 구축하고 간암 유발 기전을 model 化 함으로써 치료방법을 제시하고자 함.

(2) 연차별 연구개발 목표 및 내용

구분	연구개발 목표	연구개발 내용 및 범위
1차년도 (2003)	<p>I. C형 간염바이러스의 viral protein 과 상호작용하는 단백질의 발굴 (10가지 이상 발굴)</p> <p>II. HCV protein 및 replicon을 이용한 cytokine network 연구 (5가지 이상 발굴)</p>	<p>I-1. Viral protein의 expression vector 및 yeast two-hybrid screening vector construction</p> <ol style="list-style-type: none"> <li>1) HCV core protein</li> <li>2) HCV E1/E2 envelope protein</li> <li>3) HCV NS2 protein</li> <li>4) HCV NS3 protein</li> <li>5) HCV NS4A/4B protein</li> <li>6) HCV NS5A/5B protein</li> </ol> <p>I-2. Yeast two-hybrid screening 수행 Liver cDNA library와 HCV viral protein과의 protein-protein interaction 에 의한 간세포 단백질 스크리닝</p> <p>II-1. HCV viral proteins 과 replicon의 TNF alpha 및 TGF beta signaling 조절연구</p> <ol style="list-style-type: none"> <li>1) NF-kappaB 활성화도 변화</li> <li>2) JNK 활성화도 변화</li> <li>3) apoptosis 영향</li> </ol>
2차년도 (2004)	<p>I. C형 간염바이러스의 viral protein 과 상호작용하는 세포단백질 변화의 정량적 분석 (10가지 이상 정량 분석)</p> <p>II. HCV-associated marker 단백질 탐색 (50가지 이상 발굴)</p>	<p>I-1. 스크리닝된 세포 단백질의 특성규명</p> <p>I-2. viral protein에 의한 세포단백질의 발현 및 활성화도 변화</p> <p>II-1. HCC 임상 sample을 cDNA microarrays에 의한 marker 유전자 탐색</p> <p>II-2. HCC 임상 sample에서 세포 단백질 expression level 변화에 의한 marker 단백질 탐색</p>

### (3) 계획대비 달성도

번호	세부연구목표	달성내용	달성도(%)
1	○HCV viral protein과 상호작용하는 간세포단백질의 탐색	<ul style="list-style-type: none"> <li>◆ C형 간염바이러스의 NS5A 단백질과 TGF-β receptor I과의 상호 작용</li> <li>◆ C형간염바이러스의 NS5B 단백질과 IκBα kinase (IKK)와의 상호 작용</li> <li>◆ C형 간염바이러스의 core 단백질과 Smad 단백질과의 상호 작용</li> <li>◆ C형 간염바이러스의 NS5A 단백질과 TRAF2와의 상호 작용</li> </ul>	100%
2	○HCV viral protein에 의하여 영향을 받는 signaling molecules 의 탐색	<ul style="list-style-type: none"> <li>◆ C형 간염바이러스의 NS5A 단백질이 TGF-β receptor I과의 상호작용을 통해 TGF-β dependent signal pathway를 조절함</li> <li>- TGF-β receptor I과 NS5A 단백질이 직접적으로 상호 작용함</li> <li>- TGF-β에 의해 유도되는 Smad3와 Smad4와의 complex형성을 NS5A가 억제함</li> <li>◆ C형 간염바이러스의 replicon cell이 TNF에 의해 유도되는 NF-κB activity는 증가시키고 AP-1 activity는 감소시킴</li> <li>- TNF에 의해 유도되는 promoter activity가 replicon cell에서 증가함</li> <li>- TNF에 의한 IKK의 kinase activity와 NF-κB의 DNA binding activity가 replicon cell에서 증가됨</li> <li>- TNF에 의한 JNK activity와 AP-1의 DNA binding activity, ATF-2의 phosphorylation이 replicon cell에서 감소됨</li> <li>- Replicon cell에서 p53의 발현과 TNF에 의한 p53의 nuclear translocation이 감소로 인해 apoptosis가 억제됨</li> </ul>	100%
3	○C형 간염바이러스의 viral protein과 상호작용하는 세포단백질 변화의 정량적 분석	<ul style="list-style-type: none"> <li>◆ C형 간염바이러스 NS5A 단백질이 TGF-beta receptor I과 상호 작용하여 Smad2의 phosphorylation을 50% 억제함</li> <li>◆ TGF-β에 의해 증가되는 p21의 발현이 NS5A에 의해 50% 감소됨</li> <li>◆ NS5A가 Smad-DNA complex 형성을 50% 억제함</li> <li>◆ TGF-beta receptor I의 kinase activity에는 NS5A가 전혀 영향을 주지 않음</li> </ul>	100%

4	<p>○ HCV-associated marker 단백질 탐색</p>	<p>◆ C형 간염바이러스의 control Huh7 cell과 replicon cell에서 total RNA를 추출하여 microarray analysis를 수행</p> <ul style="list-style-type: none"> <li>- 2배 이상 up-regulated genes : 840 genes</li> <li>- 2배 이상 down-regulated genes : 1,896 genes</li> <li>- 4배 이상 fold-changed genes : 142 genes</li> </ul> <p>◆ C형 간염바이러스의 NS5A 단백질에 대한 transgenic mice의 cirrhotic liver 및 control mouse의 liver에서 total RNA를 추출하여 microarray analysis를 수행</p> <ul style="list-style-type: none"> <li>- 2배 이상 fold-changed genes : 743 genes</li> <li>- 4배 이상 fold-changed genes : 206 genes</li> </ul>	100%
---	---------------------------------------	---	------

## 제 5 장 연구개발결과의 활용계획

### 1) 추가연구의 필요성

본 연구의 개발주기는 아직까지 기초단계에 머무르고 있는 실정이다. 또한 simulation에 사용한 paramete 또한 많은 부분에서 미흡한 상태이다. 따라서 좀 더 정확한 simulation 과정을 통해 직접적인 실험을 거치지 않고 computer를 이용한 simulation을 통해 세포내에서 발생하는 신호전달과정의 정확한 예측과 이러한 과정에 C형 간염바이러스가 미치는 영향과 이를 통한 간암 유발 mechanism을 설명하기 위해서는 앞으로 몇 년간의 추가 연구가 필요할 것으로 사료된다.

### 2) 타연구에의 응용

본 연구팀은 시스템 생물학 연구사업을 통해 전혀 활용된 바 없는 세포내 신호전달과정과 C형 간염바이러스간의 연계성을 찾아내어 좀 더 명확한 바이러스성 간암의 유발 mechanism을 밝히고자 한다. 이러한 연구가 성공적으로 진행된다면 C형 간염바이러스와 유사한 RNA 바이러스에 이러한 결과를 적용시켜 바이러스를 연구하는 학자들에게 암유발 mechanism에 대한 좀더 정확한 기전을 설명할 수 있으며 또한 세포생물학자들에게 있어서는 직접적인 실험을 수행하지 않고도 세포내 신호전달과정을 computer로 예상할 수 있어 타연구에 많은 응용이 이루어질 것으로 생각된다.

### 3) 기업화 추진방안

현재의 실험 결과로는 산업체가 참여할 시기는 아니지만 추가연구의 수행을 통해 많은 data가 축적된다면 향후 산업체가 참여하여 공동 연구를 수행하고 systemic한 접근 방식을 통해 바이러스성 간염 및 간암의 발병 경로가 예측 가능하며 이를 통해 간염 및 간암에 대한 제어가 가능함으로 치료후보물질이 도출될 수 있으리라 사료된다.



## 제 6 장 연구개발과정에서 수집한 해외과학기술정보

현재 TGF-beta signal transduction과 C형 간염바이러스에 대한 systemic 접근은 전혀 이루어지지 않은 상태이다. 본 연구팀이 연구를 수행하는 과정에서 수집한 해외과학기술정보로는 우선 NF-kB와 MAPK pathway, Ras와 epidermal growth factor (EGF)가 관련된 세포내 신호전달과정에서 systems biology에 대한 연구 결과와 Kitano 그룹에서 발표한 초파리에서의 Smad 신호전달과정, 그리고 *C. elegans*에서의 TGF 신호전달과정의 systematic interaction mapping과 genetic perturbation analysis등에 대한 정보를 수집하여 이를 토대로 본 연구와 관련 있는 parameter value를 선정하고 이러한 값들을 적용하여 simulation을 수행하였다.

## 제 7 장 참고문헌

1. Aach, R. D., Stevens, C. E., Hollinger, F. B., Mosley, J. W., Peterson, D. A., Taylor, P. E., Johnson, R. G., Barbosa, L. H., and Nemo, G. J. 1991. Hepatitis C virus infection in post-transfusion hepatitis: an analysis with first- and second-generation assay. *N. Engl. J. Med.* 325:1325-1329.
2. Alter, M. J., Margolis, H. S., Krawczynski, K., Judson, F. N., Mares, A., Alexander, W. J., Hu, P.-Y., Miller, J. K., Gerber, M. A., Sampliner, R. E. et al. 1992. The natural history of community-acquired hepatitis C in the United States. *N. Engl. J. Med.* 327:1899-1905.
3. Tremolada, F., Casarin, C., Tagger, A., Ribero, M. L., Realdi, G., Alberti, A., and Ruol, A. 1991. Antibody to hepatitis C virus in post-transfusion hepatitis. *Ann. Intern. Med.* 114:277-281.
4. Choo, Q.-L., Kuo, G., Weiner, A. J., Bradley, L. R. D. W., and Houghton, M. 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244:359-362.
5. Kato, N., M. Hijikata, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi, T. Sugimura, and K. Shimotohno. 1990. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc. Natl. Acad. Sci. USA* 87:9524-9528.
6. Inchauspe, G., S. Zebedee, D.-H. Lee, M. Sugitani, M. Nasoff, and A. M. Prince. 1991. Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. *Proc. Natl. Acad. Sci. USA* 88:10292-10296.
7. Takamizawa, A., C. Mori, I. Fuke, S. Manabe, S. Murakami, J. Fujita, E. Onishi, T. Andoh, I. Yoshida, and H. Okayama. 1991. Structure and organization of the hepatitis C virus genome isolated from human carriers. *J. Virol.* 65:1105-1113.
8. Farci, P., Alter, H. J., Govindarajan, S., Wong, D. C., Engle, R., Lesniewski, R. R., Mushahwar, I. K., Desai, S. M., Miller, R. H., Ogata, N. and Purcell, R. H. 1992. Lack of protective immunity against reinfection with hepatitis C virus. *Science* 258:135-140.
9. Shimizu, Y. K., Hijikata, M., Iwamoto, A., Alter, H. J., Purcell, R. H. and Yoshikura,

- H. 1994. Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. *J. Virol.* 68:1494-1500.
10. Kao, J. H., Chen, P. J., Wang, J. T., Yang, P. M., Lai, M. Y., Wang, T. H. and Chen, D. S. 1996. Superinfection by homotypic virus in hepatitis C virus carriers: studies on patients with post-transfusion hepatitis. *J. Med. Virol.* 50:303-308.
11. Wyatt, C. A., Andrus, L., Brotman, B., Huang, F., Lee, D. H. and Prince, A. M. 1998. Immunity in chimpanzees chronically infected with hepatitis C virus: role of minor quasispecies in reinfection. *J. Virol.* 72:1725-1730.

본문 작성 요령

1. 본문의 순서는 장, 절, 1, 가, (1), (가), ①, ㉠, 등으로 하고,
  - 장은 17 포인트 고딕계열
  - 절은 15 포인트 명조계열
  - 본문은 11 포인트 명조계열로 한다.단, 본문의 내용중 중요부문은 고딕계열을 사용할 수 있다.
2. 장은 원칙적으로 페이지를 바꾸어 시작한다.
3. 본문은 11 포인트 횡으로 작성한다.
4. 페이지 번호는 하단 중앙 끝에 11 포인트로 한다.
5. 각주는 해당 페이지 하단에 8포인트 활자로 표기하며, 본문과 구분토록 한다.
6. 페이지수는 편집순서 2의 제출문부터 시작한다.  
단, 삽입물이 있을 때는 그 삽입물의 크기에 불문하고 1면을 한 페이지로 하여 일련번호를 붙인다.
7. 한글, 한문, 영문을 혼용한다.
8. 뒷면지에 주의문을 넣는다.
9. 참고문헌(reference) 인용의 경우 본문중에 사용처를 반드시 표시한다.

주 의

1. 이 보고서는 과학기술부에서 시행한 특정연구개발사업의 연구보고서입니다.
2. 이 보고서 내용을 발표할 때에는 반드시 과학기술부에서 시행한 특정연구개발사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니됩니다.

제 출 문

과학기술부 장관 귀하

본 보고서를 “간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축과 systematic 분석”과제 (세부과제 “간염바이러스 감염에 의한 간세포 단백질 네트워크 데이터베이스 구축 및 모델링”) 의 보고서로 제출합니다.

2005 . 5 . 31

주관연구기관명 : 서울대학교

주관연구책임자 : 박 근 수

연 구 원 : 김 진 욱

” : 김 이 용

” : 노 강 호

” : 이 선 호

” : 이 태 형

” : 최 승 결

보고서 초록

과제관리번호	M1-0309-06-0003	해당단계 연구기간	2003년 6월 23일 2005년 3월 31일	단계 구분	(해당단계) / (총단계) 1 / 1
연구사업명	중 사업명	특정연구개발사업			
	세부사업명	시스템생물학연구사업			
연구과제명	중 과제명	간염 바이러스의 간암유발 기작 규명을 위한 간세포 단백질 network 구축 과 systematic 분석			
	세부(단위)과제명	간염바이러스 감염에 의한 간세포 단백질 네트워크 데이터베이스 구축 및 모델링			
연구책임자	박 근 수	해당단계 참여연구원수	총 : 7명 내부 : 1명 외부 : 6명	해당단계 연구비	정부: 94,000 천원 기업: 천원 계: 94,000 천원
연구기관명 및 소속부서명	서울대학교 컴퓨터공학부		참여기업명		
국제공동연구	상대국명 :		상대국연구기관명 :		
위탁연구	연구기관명 :		연구책임자 :		
요약(연구결과를 중심으로 개조식 500자 이내)					보고서 면수
<p>43</p> <ul style="list-style-type: none"> <li>● HBV, HCV 관련 간세포 네트워크의 <i>in silico</i> 모델 개발 및 <i>in vitro/in vivo</i> 실험 데이터를 통한 모델 검증은 통해, 간염 바이러스가 간세포 단백질과 상호 작용을 통하여 간질환 및 간암을 유발하는 동역학적 메커니즘을 규명하고자 다음과 같은 연구를 진행하였다.</li> <li>● 국내외 연구 결과를 조사, 분석하여 모델링 및 파라미터 추정, 모델의 동역학 분석, 파라미터 데이터베이스 활용 기법 등을 개발</li> <li>● <i>in vitro</i> 실험 결과를 분석하여, HBV, HCV 관련 신호전달경로를 표현하는 ODE 모델을 개발하고, 이를 이용하여 신호전달경로의 동역학 특성을 분석</li> <li>● 개발된 ODE 모델을 이용하여, target molecules의 억제 효과 및 바이러스에 의한 신호전달경로의 변화 등을 시뮬레이션하고, 이를 <i>in vitro</i> 실험 결과와 비교, 검증</li> <li>● 효율적인 패턴매칭 알고리즘, 효율적인 서열정보의 정렬 기법, 인덱스 자료구조 등 생물정보학 기초기술 관련 연구</li> </ul>					
색인어 (각 5개 이상)	한 글	간세포, 간염 바이러스, 단백질 네트워크, 데이터베이스, 모델링			
	영 어	Hepatocyte, Hepatitis virus, Protein network, <i>in silico</i> modeling, database			

## 요 약 문

### I. 제 목

간염바이러스 감염에 의한 간세포 단백질 네트워크 데이터베이스 구축 및 모델링

### II. 연구개발의 목적 및 필요성

#### 가. 연구개발의 목적

본 연구는 간세포 단백질 네트워크 상에서 간염 바이러스가 간세포 단백질과 상호 작용을 통하여 간질환 및 간암을 유발하는 동역학적 메커니즘을 규명하고자 다음을 목적으로 한다.

- HBV, HCV 관련 간세포 네트워크의 *in silico* 모델 개발
- 협력연구기관의 *in vitro/in vivo* 실험 데이터를 통한 모델 검증
- HBV, HCV 관련 데이터베이스 설계 및 구축

#### 나. 연구개발의 필요성

우리나라 전체 인구의 약 5~8%가 간염 바이러스에 감염되어 있으며, 이들 중 상당수가 간암으로 진행하여 사망하는 것으로 보고되고 있다. 이렇듯 국민 보건에 큰 위협이자 가정과 사회의 안녕을 해치고 있는 간염 및 간암을 예방하고 이를 치료하기 위해, 간염 바이러스의 간질환 및 간암 유발 기작을 규명하는 연구가 매우 시급하다.

간염 바이러스의 간세포 내 간 질환 유도 메커니즘을 이해하기 위해서, 간세포 네트워크 및 간염 바이러스 관련 데이터베이스 구축 및 *in silico* 모델 시스템 개발이 필요하다. 본 연구를 통해 개발된 HBV/HCV 관련 간세포 신호전달경로 *in silico* 모델과 데이터베이스, 생물정보학 관련 기초 기술은 향후 간염의 간질환 유도 작용에 대한 동역학적 메커니즘 규명 및 효과적 치료 기술 개발에 폭넓게



활용될 것으로 기대된다.

### III. 연구개발의 내용 및 범위

#### 가. Genetic network 모델 연구

국내외 연구 결과를 조사, 분석하여 신호전달경로를 표현하는 적합한 모델링 기법을 선정한다.

#### 나. HBV, HCV 네트워크 모델 구성 및 검증

*in vitro* 실험 결과를 분석하여, HBV, HCV 관련 신호전달경로를 표현하는 ODE 모델을 개발하고, 이를 이용하여 신호전달경로의 동역학 특성을 분석한다. 또한 개발된 ODE 모델을 이용하여, 간염 바이러스에 의한 신호전달경로의 변화 등을 시뮬레이션하고, 이를 *in vitro* 실험 결과와 비교, 검증한다.

#### 다. 관련 데이터베이스 설계

Systems biology 관련 데이터 표현의 표준인 SBML을 조사, 분석하고, HBV, HCV 관련 네트워크를 SBML로 표현한다. 또한 Pathway 데이터베이스를 분석하기 위한 프로그램을 개발한다.

#### 라. 생물정보학 기초 기술 연구

효율적인 패턴매칭 알고리즘, 효율적인 서열정보의 정렬 기법, 인덱스 자료구조 등 생물정보학 기초기술 관련 알고리즘을 개발한다.

### IV. 연구개발결과

#### 가. Genetic network 모델 연구

본격적인 간세포 네트워크 모델링에 앞서, 적합한 모델링 방법을 선정하기 위해 genetic network를 표현하기 위해 제안된 여러 가지 모델들을 조사, 분류하여 비교하였다. Genetic network를 표현하는 Boolean model, Linear model, Differential model 중 Differential model이 간세포 신호전달경로 모델링에 적합한 것으로 나타났다.

#### 나. HBV, HCV 네트워크 모델 구성 및 검증

신호전달경로의 ODE 모델에 관한 국내외 연구 결과를 조사 분석하여 모델링

및 파라미터 추정, 모델의 동역학 분석, 파라미터 데이터베이스 활용 기법 등을 개발하였다. 이를 통해 제1, 2세부과제 연구기관에서 진행된 실험결과를 분석하여, **HBV, HCV 관련 신호전달경로를 표현하는 ODE 모델을 4종 이상** 개발하였다. 개발된 ODE 모델은 간세포 신호전달경로의 동역학 특성을 분석하는데 활용되었다. 특히 target molecules의 억제 효과 및 간염 바이러스에 의한 신호전달경로의 변화 등을 시뮬레이션하고 *in vivo* 실험 결과를 비교하여 이를 검증하였다. 연구 결과는 국내 학술대회 및 국제 저널에 발표하였다. (논문 1편 발표, 1편 심사 중)

#### 다. 관련 데이터베이스 설계

○ SBML(Systems Biology Markup Language)은 생물정보학 관련 프로그램/데이터베이스 상에서 biochemical network 모델을 표현하는 표준 방법을 제공한다. 제1세부과제와 제2세부과제의 HBV, HCV 네트워크를 design 한 후, SBML tool인 Cell-designer를 이용하여 SBML 모델을 구성하였다.

○ KEGG PATHWAY 데이터베이스에서 제공하는 pathway 그래프를, EST data와 융합하여, 그래프 상의 각 노드에 해당하는 단백질들의 발현 패턴을 tissue source별로 선택, 비교·분석할 수 있는 소프트웨어 도구를 개발하였다.

○ 그래프 모델과 ODE 모델의 인터페이스 역할을 하는 ODE 모델 명세 언어를 개발하였다.

#### 라. 생물정보학 기초 기술 연구

아래와 같은 연구를 통해 대용량 생물정보처리 데이터베이스에서 활용될 수 있는 효율적인 알고리즘을 개발하였다.

- 효율적인 패턴매칭 알고리즘의 개발
- 효율적인 서열정보의 정렬 기법의 개발
- 인덱스 자료 구조에 관련된 연구

#### V. 연구개발결과의 활용계획

본 연구를 통해 구축된 간세포 신호전달경로 및 HBV/HCV 관련 *in silico* 모델은 향후 간염의 간질환 유도 작용에 대한 동역학적 메커니즘 규명 및 효과적 치료 방법 개발에 폭넓게 활용될 것으로 기대된다. 본 연구개발결과를 활용하여 다음과 같은 추가 연구가 진행될 수 있다.

- *in silico* 모델 구조의 정립
- *in silico* 상 주요 파라미터 선별 기법 개발
- *in silico* 모델의 동역학 특성 분석 기법 개발

## S U M M A R Y

Title: Protein Network Database Construction & Modeling of Hepatocyte

Throughout the *in silico* simulation study as well as *in vitro* experiments on the protein network in hepatocyte cells, we investigated into the dynamic mechanism of liver disease development induced by hepatitis virus.

- By analyzing *in vitro* experimental data, we developed ODE models of HBV/HCV-related signaling network in hepatocyte cells. With developed models, we conducted computational simulation and analysis to find out dynamical properties on the signaling network.
- Based on the developed model, inhibitory effect of target molecules on the signaling network was analyzed. We further constructed ODE models of HBV-infected cells, which were utilized to find out the relationship between NF- $\kappa$ B activation and liver diseases induced by HBV.
- We developed efficient algorithms for pattern matching, sequence alignment and index data structure, which are practically useful in the area of bioinformatics.

It is expected that the dynamic mechanism of HBV/HCV mediated liver disease development is to be thoroughly discovered through utilizing the developed models and the results can be applied for prevention of liver disease and advanced treatment for hepatitis.

C O N T E N T S

Chapter 1 Introduction .....	9
Chapter 2 The State of the Art .....	10
Chapter 3 Results of Research & Development .....	11
Section 1 Genetic Network Models .....	11
Section 2 HBV/HCV Network Modeling and Verification .....	13
Section 3 Database Design .....	23
Section 4 Bioinformatics Algorithms .....	29
Chapter 4 Accomplishments and Contributions .....	37
Chapter 5 Suggestion for Utilization .....	40
Chapter 6 Technical Information .....	41
Chapter 7 References .....	42

목 차

제 1 장 연구개발과제의 개요 .....	9
제 2 장 국내외 기술개발 현황 .....	10
제 3 장 연구개발수행 내용 및 결과 .....	11
제 1 절 Genetic network 모델 연구 .....	11
제 2 절 HBV, HCV 네트워크 모델 및 검증 .....	13
제 3 절 관련 데이터베이스 설계 .....	23
제 4 절 생물정보학 기초 기술 연구 .....	29
제 4 장 목표달성도 및 관련분야에의 기여도 .....	37
제 5 장 연구개발결과의 활용계획 .....	40
제 6 장 연구개발과정에서 수집한 해외과학기술정보 .....	41
제 7 장 참고문헌 .....	42

## 제 1 장 연구개발과제의 개요

본 연구 과제는 간세포 단백질 네트워크 상에서 간염 바이러스가 간세포 단백질과 상호 작용을 통하여 간질환 및 간암을 유발하는 동역학적 메커니즘을 규명하고자 다음을 목표로 진행되었다.

- HBV, HCV 관련 간세포 네트워크의 *in silico* 모델 개발
- 협력연구기관의 *in vitro/in vivo* 실험 데이터를 통한 모델 검증
- HBV, HCV 관련 데이터베이스 설계 및 구축

이를 위해 먼저 기존 genetic network model 조사 및 검토가 이루어졌으며, 이를 바탕으로 HBV, HCV 관련 네트워크에 대한 ODE 모델을 개발하고 수치해석 프로그램을 이용한 시뮬레이션을 진행하였다. 또한 개발된 모델을 확장하여, 약물에 의한 억제 효과 및 바이러스에 의한 모델의 동역학적 변화를 시뮬레이션 하였고, 협력연구기관의 *in vitro* 실험을 통한 비교 및 검증을 수행하였다. 한편, 시스템생물학 관련 데이터베이스 표준인 SBML 언어에 대한 조사를 통해, HBV, HCV 관련 네트워크의 SBML 모델을 구현하였으며, EST 정보를 이용한 pathway 데이터베이스 검색 프로그램을 개발하였다. 또한 대용량 생물정보 데이터베이스에 응용될 수 있는 효율적인 정보 저장 및 처리 알고리즘을 개발하였다.

본 연구 개발의 필요성 및 기대 효과는 다음과 같다.

### ■ 기술적 측면

간염 바이러스의 간세포 내 간 질환 유도 메커니즘을 이해하기 위해서, 간세포 네트워크 및 간염 바이러스 관련 데이터베이스 구축 및 *in silico* 모델 시스템 개발이 필요하다. 본 연구를 통해 개발된 HBV/HCV 관련 간세포 신호전달경로 *in silico* 모델과 데이터베이스, 생물정보학 관련 기초 기술은 향후 간염의 간질환 유도 작용에 대한 동역학적 메커니즘 규명 및 효과적 치료 기술 개발에 폭넓게 활용될 것으로 기대된다.

### ■ 경제 산업적 측면

전 세계적으로 간염 및 간암 치료제 시장은 약 7조원 이상이 될 것으로 추정되고 있다. HBV 및 HCV의 작용 기작을 시스템적으로 규명하여 이를 간염 치료제 개발에 활용한다면, 제약 관련 산업의 발전 및 수출 증대 효과 뿐만 아니라, 간염 및 간암으로 인한 국가적 비용 절감 및 노동력 손실 회복 등 국가 경제·산업 경쟁력 강화에 크게 기여할 것으로 기대된다.

### ■ 사회 문화적 측면

2003년 통계청 자료에 따르면 가정 및 사회의 중심축인 40대 남자의 사망원인 중 상당수가 간암 및 간질환(2위)인 것으로 나타나있다. 이렇게 국민 보건에 큰 위협이자 가정과 사회의 안녕을 해치고 있는 간염 및 간암을 예방하고 이를 치료하기 위해, 간염 바이러스의 간질환 및 간암 유발 기작을 규명하는 연구가 매우 시급하다.

## 제 2 장 국내외 기술개발 현황

Systems Biology 연구는 국내외적으로 아직 개발기에 속해있다. 특히 전체 기술 분야를 주도할 수 있는 이렇다 할 구체적인 방법론이 개발되어 있지 않아 대상과 방법에 대한 연구의 여지가 많이 남아있다 하겠다. 이로 인해 생명 과학 분야뿐만 아니라 컴퓨터 과학, 제어 공학, 물리학 등 다양한 학문 분야의 관심을 불러 일으켜 학제간(interdisciplinary) 학문 연구를 촉진시키고 있다.

생물정보 데이터베이스와 관련하여 외국의 경우 기존에 밝혀진 생명 현상과 관련된 지식을 통합한 데이터베이스를 구축하는 작업이 진행되고 있다. 스위스 바이오인포매틱스 연구소의 단백질 데이터베이스인 ExPASy (Expert Protein Analysis System)나, 일본 교토대학교의 KEGG (Kyoto Encyclopedia of Genes and Genomes) 등이 그 대표적인 예이다.

한편, 2000년대에 들어 포스트 지놈 시대 생명 과학 연구가 본격화됨에 따라 생명 현상에 대한 *in silico* 모델 구축을 통해, 축적된 *in vitro/in vivo* 실험 데이터를 통합하고, 생명 현상에 내재된 동역학적 특성을 밝히려는 움직임도 활발히 진행되고 있다. *in silico* 모델링은 궁극적으로, 세포 내 물질들의 상호 작용 경로를 수학적 모델로 기술하여, 세포의 상태 변화 및 물질 간 상호 작용이 가지는 동적 특성을 *in silico* 상에서 시뮬레이션하는 E-Cell을 구축하고자 한다. 이를 위하여 미국, 일본, 독일 등 선진 각국에서는 다양한 학문 분야를 기반으로 한 연구자들의 상호 교류 및 협력 연구가 활발히 시도되고 있다.

특히 신호전달경로의 모델링과 관련한 연구로는, 경로를 이루는 주요 구조인 kinase cascade, receptor system 및 transcription factor 조절 시스템에 관한 연구가 다각적으로 진행되고 있다. 그러나 MAPK cascade 및 EGFR 시스템 등 일부 시스템에만 제한적으로 모델링이 이루어지고 있으며, 이를 종합한 전체 신호전달경로를 모델링하는 연구는 아직 연구 초기 단계에 있다.

이와는 별도로 wet 실험의 target을 설정하기 위해, sensitivity analysis, Monte Carlo 시뮬레이션 등 공학적 분석 방법을 적용하여 *in silico* 모델 상에서 중요한 파라미터 및 molecule을 선별하는 연구도 진행되고 있다. 이러한 연구들은 향후 dry 시뮬레이션을 통한 drug target 발굴 등에 적극 활용될 전망이다. 그러나 아직 특정 질병, 질환과 관련된 신호전달경로의 *in silico* 모델 개발은 미진하다.

선진 각국의 활발한 연구 동향과 비교해 볼 때 국내 *in silico* 모델 연구는 일부 분야를 제외하고는 아직 초기 단계에 머무르고 있다. 특히 wet 실험실과 dry 실험실 간의 실질적인 연구 협력 체계 구축이 미비한 형편이다.

본 연구에서는 인간 간세포에서 HBV, HCV와 관계된 신호전달경로를 모델링하고, 이에 대한 *in silico* 시뮬레이션 및 *in vitro* 검증을 시도하였다. 이는 국내외적으로도 선행 연구를 찾아보기 힘든 것으로, *in silico* 모델링을 통한 질병의 작용 기작에 대한 동역학적 원리 규명 뿐만 아니라, 치료 목적으로의 응용 가능성을 열었다는 점에서 큰 의미를 지니고 있다.

또한, 본 과제를 수행하기 위해 wet/dry 실험실 간 구축해온 긴밀한 연구 협력 관계는 아직 걸음마 단계에 있는 국내 Systems Biology 연구 분야에서, wet/dry 연구 간의 성공적인 협력 사례를 제시할 수 있으리라 기대된다.

# 제 3 장 연구개발수행 내용 및 결과

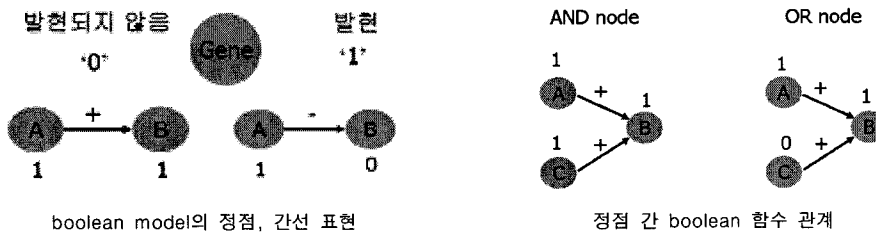
## 제 1 절 Genetic network 모델 연구

Genetic network는 생명 현상을 나타내는 유전자, 단백질 등의 컴포넌트와 이들 간의 상호 작용의 집합을 말한다. 실제 생명 현상이 일어나는 과정을 규명하기 위해서는, 실험으로부터 얻은 데이터를 이용하여 이 현상에 참여하는 컴포넌트를 찾아내고, 이들 간의 상호 작용을 발견하는 과정이 필요하다.

본 연구에서는 본격적인 간세포 네트워크 모델링에 앞서, 적합한 모델링 방법을 선정하기 위해 genetic network를 표현하기 위해 제안된 여러 가지 모델들을 조사, 분류하여 보았다. Genetic network 모델은 컴포넌트 간의 상호 작용을 표현하는 방식에 따라 다음과 같이 분류할 수 있다.

### 1. Boolean Model

유전자 또는 단백질의 발현 상태를 0/1로 나타내는 정점(node)으로 나타내고, 정점 간의 유도/억제 관계를 +/- 간선과 boolean 함수(AND,OR)로 표현한 모델이다. 이 때 AND 함수는 입력으로 들어오는 간선에 연결된 정점이 모두 발현될 때만 출력 정점이 발현되는 관계를, OR 함수는 입력 정점 중 하나만 발현이 되어도 출력 정점이 발현되는 관계를 나타낸다 (아래 그림 참고).

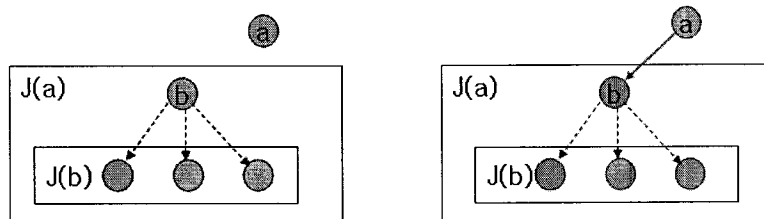


### ■ Boolean model 추론을 위한 휴리스틱 알고리즘

특정 유전자들에 대해 선택적으로 overexpression 또는 disruption을 수행한 후, microarray 등 대규모 gene expression 측정 기술을 이용하여 얻은 데이터로부터 genetic network를 추론하는 몇 가지 휴리스틱 알고리즘들이 존재한다.

그 중 한가지 방법으로 다음 휴리스틱을 이용하여 genetic network를 추론할 수 있다.

한 정점  $x$ 의 값이 0일 때와 1일 때, 서로 다른 값을 갖는 정점들의 집합을  $J(x)$ 라고 하자. Genetic network내의 모든 정점에 대해서  $J(x)$ 를 구한 다음,  $J(b) \cup \{b\} = J(a)$ 인 두 정점  $a, b$ 에 대해서  $a \rightarrow b$ 인 간선을 추가한다. 이와 같은 방법으로 genetic network를 구할 수 있다. 이때 필요한 실험의 회수는 network 내의 정점의 개수의 두 배임이 증명되어 있다.

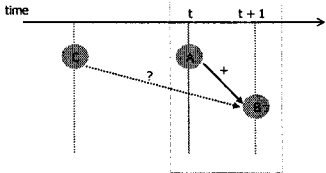
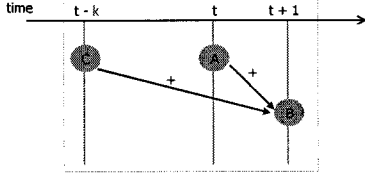
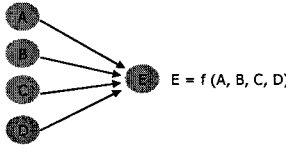
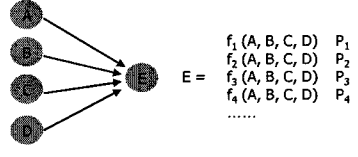


실험 데이터로부터 boolean model을 추론하는 휴리스틱 알고리즘



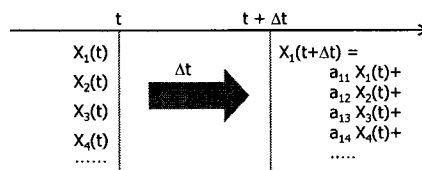
## ■ Boolean model의 확장

다음은 boolean model의 표현 능력을 확장한 모델이다.

구분	내용
Temporal Boolean Model	<p>실제 생명 현상에서 발생하는 유도/억제 작용의 시간 delay를 반영한 모델. 일반 boolean model과 달리, 컴포넌트의 상태를 결정하기 위해 바로 직전 시점에서의 컴포넌트의 상태 뿐 아니라, 그 이전 시점에서의 컴포넌트의 상태까지 고려하도록 제한된 모델이다. 구체적으로는 아래 그림에서와 같이, <math>t+1</math> 시점에서의 컴포넌트의 상태는 <math>t, t-1, \dots, t-k</math> 시점에서의 컴포넌트의 시점들에 의해 결정된다(아래 그림 참고).</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>일반 boolean model</p> </div> <div style="text-align: center;">  <p>Temporal boolean model</p> </div> </div>
Probabilistic Boolean Model	<p>제한된 데이터와 오류로 인하여 발생할 수 있는 불확실성을 확률 함수의 형태로 반영한 모델. 일반 boolean model과 달리, 한 컴포넌트의 상태를 결정하는 함수를 하나만 두지 않고 여러 함수를 둔 후, 각 함수 별로 이 함수를 선택할 확률을 계산하고, 이에 의거하여 확률적으로 함수를 선택하여 컴포넌트의 상태를 결정한다. 아래 그림의 예에서 컴포넌트 E의 상태는 컴포넌트 A, B, C, D에 대해 정의된 함수 <math>f_1, f_2, f_3, f_4 \dots</math> 중, 각각의 함수에 연관된 확률 <math>P_1, P_2, P_3, P_4 \dots</math> 에 의거, 하나의 함수를 선정하여 결정된다.</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>일반 boolean model</p> </div> <div style="text-align: center;">  <p>Probabilistic boolean model</p> </div> </div>

## 2. Linear Model

유전자의 발현 여부를 발현율과 같은 정량적인 수치로 표현하는 모델. 특정 유전자의 발현율은 다른 유전자들의 발현율의 linear combination으로 표현된다.



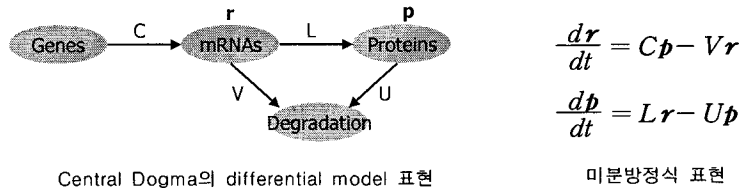
Linear model

즉, 시각  $t$  에서 유전자 1의 발현율  $X_1(t)$ 에 대해,  $\Delta t$  시간이 흐른 뒤의 발현율  $X_1(t+\Delta t)$ 는 시각  $t$  에서 유전자 1, 2, 3, 4, ... 의 발현율  $X_1(t), X_2(t), X_3(t), X_4(t), \dots$ 의 linear combination 인  $X_1(t+\Delta t) = a_{11}X_1(t) + a_{12}X_2(t) + a_{13}X_3(t) + a_{14}X_4(t) \dots$ 로 표현된다. 실험을 통하여 linear combination의 계수  $a_{ij}$ 를 구함으로써, 표현하고자 하는 생명 현상에 대한 linear model을 얻게 된다.

### 3. Differential Model

시간에 따른 유전자 및 단백질의 농도 변화를 미분 방정식의 형태로 정량화한 모델. Genetic network의 각 컴포넌트 간의 관계를 미분 방정식으로 표현하고, 이 미분 방정식을 풀어 컴포넌트 간 관계가 시간에 따라 변화하는 동역학을 표현한다.

아래 그림은 유전자에 담긴 정보가 mRNA를 통해서 단백질로 표현되는 Central Dogma를 differential model로 나타낸 것이다. 생성된 mRNA와 단백질의 양을 각각  $r$ 과  $p$ 라는 벡터로, 유전자에서 mRNA로 가는 관계를 행렬  $C$ , mRNA에서 단백질로 가는 관계를 행렬  $L$ 이 표현한다고 하자. 또한 mRNA와 단백질은 각각 행렬  $V$ 와  $U$ 로 표현되는 관계에 의해 분해된다고 하면, 이 genetic network은 아래와 같은 미분 방정식으로 정량화할 수 있다.



위 미분 방정식을 풀고, 실험 데이터를 통하여 행렬  $C$ ,  $V$ ,  $L$ ,  $U$ 의 값을 결정하면 시간의 변화에 따른  $r$ 과  $p$ 의 함수를 얻을 수 있다.

## 제 2 절 HBV, HCV 네트워크 모델 및 검증

### 1. ODE 모델

일반적으로 시스템 생물학 연구는 생명 현상에 대한 수학적 모델을 설정하고, 이 모델에서 발견되는 구조적, 동역학적 특성을 실험을 통해 검증하는 방식으로 이루어진다. 여러 가지 수학적 모델 가운데 상미분 방정식(Ordinary Differential Equation, ODE) 모델은 세포 내 여러 가지 생화학 반응을 모델링 하는데 광범위하게 사용되고 있다. 이 모델은 생명 현상에 관여하는 물질들의 시간에 따른 농도 변화를 미분 방정식을 통해 정량적으로 나타내고 있다. 특히, 이 모델을 이용하면 생화학 반응의 동역학적 특성을 이해하는데, 수학적으로 잘 정의된 분석 방법과 수치 해석 패키지를 효과적으로 활용할 수 있다.

신호전달 경로 상에서 발생하는 생화학 반응은 기본적으로 효소 반응과 같은 단백질 간의 상호작용의 형태로 나타난다. 이러한 기본적인 생화학 반응들은 다음과 같은 형태로 나타낼 수 있다. 아래와 같이 단백질 간의 상호작용을 모델링 하기 위해, 반응 상수( $k_f, k_b$ ) 및 반응 물질의 농도 ( $[A]$ ,  $[B]$ )로 이루어진 미분 방정식으로 ODE 모델을 구성한다.

구분	반응식	ODE 모델
합성반응	$A + B \xrightleftharpoons[k_b]{k_f} AB$	$\begin{aligned} d[A]/dt &= -k_f[A][B] + k_b[AB] \\ d[B]/dt &= -k_f[A][B] + k_b[AB] \\ d[AB]/dt &= k_f[A][B] - k_b[AB] \end{aligned}$
치환반응	$A + B \xrightleftharpoons[k_b]{k_f} C + D$	$\begin{aligned} d[A]/dt &= -k_f[A][B] + k_b[C][D] \\ d[B]/dt &= -k_f[A][B] + k_b[C][D] \\ d[C]/dt &= k_f[A][B] - k_b[C][D] \\ d[D]/dt &= k_f[A][B] - k_b[C][D] \end{aligned}$
효소반응	$E + S \xrightleftharpoons[k_2]{k_1} ES \xrightarrow{k_3} E + P$	$\begin{aligned} d[E]/dt &= -k_1[E][S] + (k_2 + k_3)[ES] \\ d[S]/dt &= -k_1[E][S] + k_2[ES] \\ d[ES]/dt &= k_1[E][S] - (k_2 + k_3)[ES] \\ d[P]/dt &= k_3[ES] \end{aligned}$ <p>※ Michaelis-Menten 효소 반응 모델이 모델에서 효소-기질 복합체(ES)에서 효소(E)와 반응산물(P)이 분리되는 과정은 비가역적임을 가정한다.</p>

위와 같이 생화학 반응 모델을 미분방정식으로 나타내고, 실험과 시뮬레이션을 통해 관련 파라미터를 구하면, 이 단위 모델들을 통해 보다 큰 반응 네트워크 모델 시스템을 구축할 수 있다. 또한, 이 정량적 모델에 대한 수학적 해석을 통해 동역학 분석 및 반응 예측 등을 수행하여 생물학적으로 의미 있는 발견을 기대할 수 있다.

본 연구에서는 위와 같은 ODE 모델을 이용하여 HBV/HCV 감염으로 인한 간세포 생화학 반응의 변화를 모델하고, 이 모델의 동적 특성을 실험과 시뮬레이션을 통하여 검증하는 방법을 제시하였다.

## 2. HBV 관련 신호전달경로 ODE 모델 구성 및 검증

### 가. NF-κB 신호전달경로

NF-κB 신호전달경로는 세포의 사멸 및 생장을 조절하는 중요한 신호전달경로이다. NF-κB는 TNFα 등 proinflammatory cytokine에 의해 주로 activation 되는데, 이러한 세포 외부 신호가 전달되어, IKK가 phosphorylation되면, NF-κB와 binding된 IκB의 ubiquitination을 유도함으로써 NF-κB가 핵으로 이동한다는 메커니즘이 밝혀져 있다. NF-κB는 여러 질병과 치료 작용에 관련되어 있는데, 제1세부과제팀의 연구 결과 특히 HBV 감염에 의해 그 activation이 증가하는 것으로 밝혀졌다. 본 연구에서는 NF-κB 신호전달경로의 *in silico* 모델을 구성하고, 제 1세부과제팀의 wet 실험 결과를 바탕으로 HBV가 NF-κB activation 신호전달경로에 주는 동역학적 변화

에 대한 모델링 및 시뮬레이션을 진행하였다.

## 나. ODE 모델의 구성

NF- $\kappa$ B activation과 관련하여, IKK에 의해 조절되는 NF- $\kappa$ B 신호전달 모듈이 Hoffmann 등에 의해 ODE로 모델링되었다[6]. 이 후 IKK downstream에서 NF- $\kappa$ B activity가 보이는 다양한 동적 특성을 *in silico* 모델을 이용해 규명하려는 연구가 활발히 진행되고 있다[8]. 한편, NF- $\kappa$ B를 activation시키는 주요 cytokine인 TNF $\alpha$ 로부터 시작되는 신호전달경로 모델을 구성하려는 연구도 진행 중에 있다[3, 4]. 그러나 아직까지 *in silico/in vitro* 실험 협력을 통한 모델 검증 및 특정 질병과 관련한 신호전달경로의 변화 모델링 등의 연구는 미흡한 실정이다.

본 연구에서는 먼저, 기존 연구를 통해 부분적으로 구성된 NF- $\kappa$ B 신호전달경로 모델을 종합하여 전체 신호전달경로 *in silico* 모델을 구성하고 이에 대한 시뮬레이션을 진행하였다.

*in silico* 모델의 설계 및 분석을 용이하게 하기 위하여, 전체 시스템을 기능 및 세포 내 위치에 따라 나누는 모듈화 원칙[5]을 적용, 전체 신호전달경로 모델을 아래와 같이 3개의 서브시스템으로 구성하였다.

### ■ TNF $\alpha$ receptor 시스템

세포 외부에서 전달되는 cytokine TNF $\alpha$ 가 세포막에 닿아 TNFR1 receptor 및 TRADD, RIP 등의 receptor adaptor 단백질과 차례로 complex를 이루는 과정을 ODE 모델로 구성하였다. 최종적으로 TRAFs 단백질이 complex에 참여하면, 이를 통해 다음 단계인 kinase cascade가 activation 된다. FADD 단백질은 세포 사멸과 관련된 또 다른 신호전달경로를 매개하는 물질로 본 모델에서는 receptor complex에 binding되는 부분만 모델 되었다. Receptor 시스템과 관련하여 EGFR receptor 시스템 연구[2] 등 선행 연구의 파라미터들을 참고하였다.

### ■ Kinase cascade 시스템

Kinase cascade는 일련의 kinase 단백질들이 연쇄적으로 다음 단계 protein을 phosphorylation을 통해 activation시키는 신호전달 구조이다. 지금까지 대표적인 kinase cascade인 MAPK cascade에 대한 연구가 활발히 진행되어 왔으며, amplification, switching behavior, ultrasensitivity 등 kinase cascade 구조가 갖는 동적 특성들도 많이 연구되고 있다[1, 7]. 본 연구에서는 1단계 TNF $\alpha$  receptor 시스템에서 형성된 TRAFs complex를 입력 kinase로 IKKK(RIPK, NIK 등), IKK가 차례로 activation되는 모델을 구성하였다.

또한, kinase cascade 모델에는 kinase를 deactivation시키는 phosphatase도 포함되어 있다. 본 연구에서는 이 phosphatase와 kinase의 상대적 비율에 따라 kinase cascade의 동적 특성이 결정됨을 *in silico* 시뮬레이션을 통해 확인하였다(그림3).

### ■ NF- $\kappa$ B 조절 시스템

2단계 kinase cascade에서 activation된 IKK는 NF- $\kappa$ B와 binding된 I $\kappa$ B isoforms를 phosphorylation함으로써 I $\kappa$ B의 ubiquitination 및 proteolysis를 유도한다. I $\kappa$ B의 감소는 세포질 내에 묶여 있던 transcription factor NF- $\kappa$ B가 핵으로 이동하는 것을 허용함으로써 DNA 상에서 NF- $\kappa$ B와 binding하는 여러 가지 유전자들을 transcription한다. 이 유전자들 중 핵 내에 있는 NF- $\kappa$ B를 다시 세포질로 이동시키는 I $\kappa$ Ba가 포함되어 있어, NF- $\kappa$ B를 조절하는 핵심적인 역할을 담당하고 있다. 이러한 negative feedback loop 시스템은 NF- $\kappa$ B 조절 시스템의 동역학 모델링에 있어 중요한 연구 주제로 주목받아 왔다. 본 모델에서는 Hoffmann의 NF- $\kappa$ B 조절 모델

중 NF- $\kappa$ B 동역학적 변화에 결정적인 역할을 담당하는 I $\kappa$ Ba와 관계된 구조와 파라미터를 사용하였다.

위와 같이 서브시스템 모델을 구성하고, 각각의 동역학 특성을 분석한 후, 이를 종합하여 28개의 단백질과 39개의 파라미터로 구성된 전체 신호전달경로 모델을 (그림 1)와 같이 구성하였다.

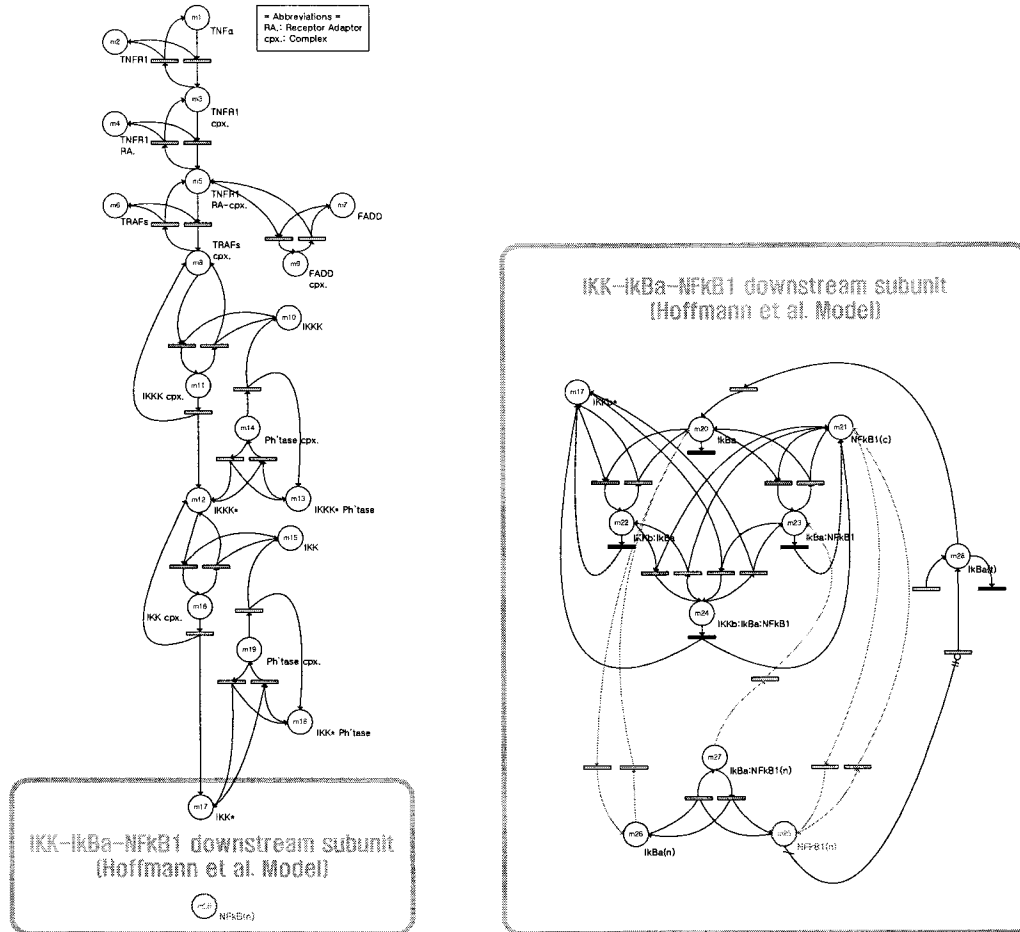


그림 1 NF- $\kappa$ B 조절 시스템(우)과 전체 NF- $\kappa$ B 신호전달경로 모델 구성(좌)

#### 다. 초기 농도 및 파라미터 결정

신호전달경로 ODE 모델은 기본적으로 반응 상수와 초기 조건에 매우 민감한 영향을 받는 비선형 시스템이다. 따라서 간세포 내에서 일어나는 생화학 반응을 정확히 모델링하기 위해서는 실제 *in vitro* 실험을 통한 간세포 특성적인 파라미터 및 주요 단백질의 초기 농도 측정이 필수적이다. 제1세부과제팀의 *in vitro* 실험을 통하여 다음과 같은 모델 파라미터 값들을 결정하였다.

구분	TNF $\alpha$	TRAF2	NIK	IKK $\alpha$	IKK $\beta$
HepG2	0.00039	0.33389	0.04522	0.02417	0.02278
HepG2.2.15	0.00039	0.73653	0.06714	0.07243	0.05999

표 11 *in vitro* 실험을 통한 모델 초기 조건 측정 (단위  $\mu$ M)

기타 모델 파라미터는 *in vitro* 실험값과 literature를 참고하여 결정하였다.

## 라. *in silico* 시뮬레이션

### ■ 서브시스템 시뮬레이션 결과 분석

전체 신호전달경로에 대한 시뮬레이션에 앞서 모듈별로 *in silico* 모델을 검증하고, 특성을 파악하기 위한 시뮬레이션을 진행하였다.

먼저, kinase cascade 상의 입력 kinase, IKKK 및 IKK의 농도와 이에 대한 phosphatase 비율을 변화시키면서 IKKK\*(phosphorylated IKKK, active form)과 IKK\*(phosphorylated IKK, active form)의 시간에 따른 농도 변화(그림 2B)와 평형 상태 농도(그림 2C)를 각각 2D, 3D plot으로 분석하였다.

시뮬레이션 결과 kinase의 농도는 출력 IKK\*(phosphorylated IKK) 농도의 차수(order)를 결정함을 알 수 있다. 패널 B는 세로축을 따라 kinase의 양이 10배씩 증가됨에 따라 출력 IKK\*(E2\*로 표시)도 비슷한 차수로 증가하는 경향을 나타내고 있다.

또한 phosphatase의 상대적 비율이 높아질수록 출력 IKK\*의 농도는 낮아지며 평형 상태에 이르는 시간도 늘어남을 볼 수 있다. B에서 가로축 방향으로 phosphatase의 비율이 늘어날수록 IKK\*의 농도가 평형 상태에 이르는 시간이 늘어남과 동시에, 농도 역시 낮아지는 경향이 발견된다. 그림 2C에서는 phosphatase의 비율 및 kinase의 농도를 x, y축으로 놓고, 평형 상태에서의 IKK\* 출력 농도를 z축 방향의 높이로 나타내고 있다. 등고선 그래프를 통해 살펴보듯이

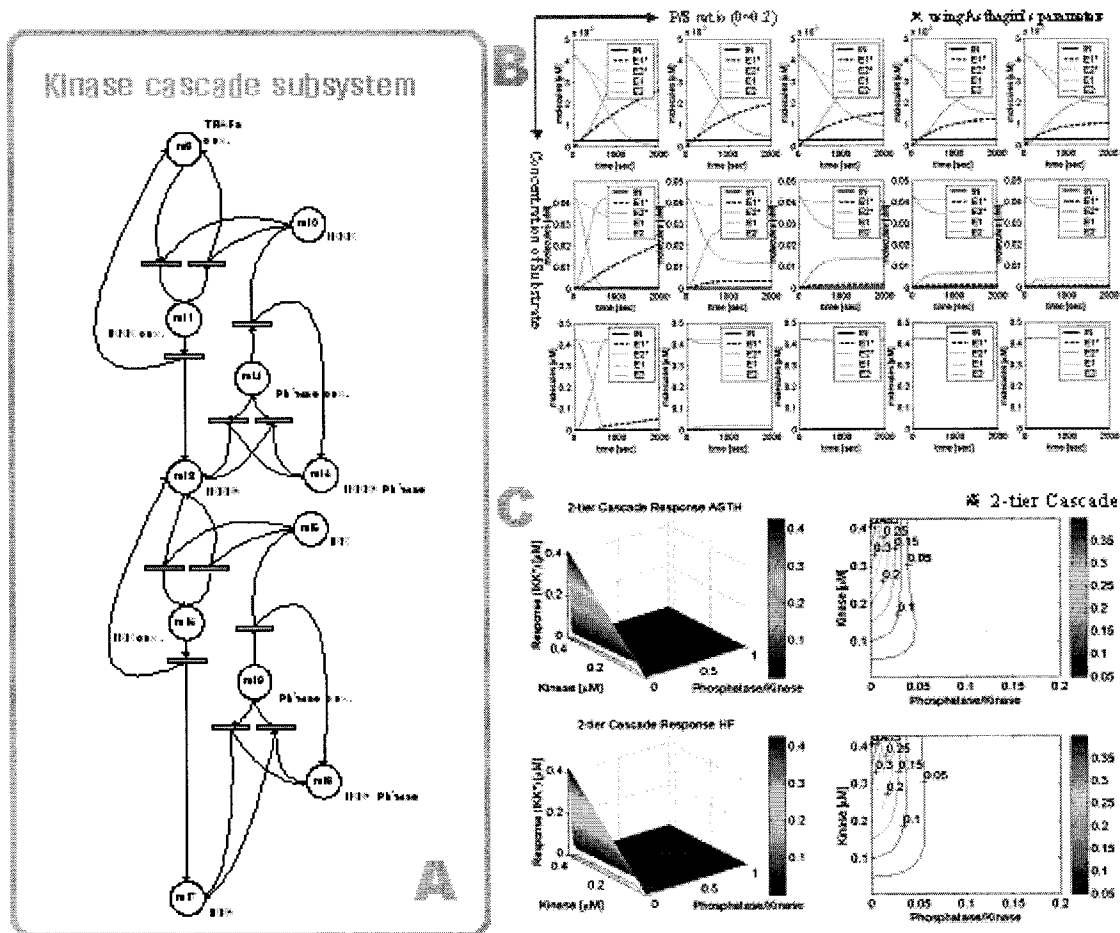


그림 2 Kinase cascade 모델 및 *in silico* 시뮬레이션 분석 (A) Kinase cascade ODE 모델 (B) Time plot 분석 (C) Kinase 농도 및 Phosphatase 비율에 따른 출력 3D 분석

phosphatase의 비율이 낮을수록, 그리고 kinase의 농도가 높을수록 IKK\*의 농도가 높아지며, 등고선의 기울기를 고려해 볼 때, 이는 phosphatase의 비율 변화에 더욱 민감함을 알 수 있다.

NF- $\kappa$ B 조절 서브시스템에 대해서도 입력 IKK\*에 따른 출력 NF- $\kappa$ B(n) (핵으로 이동한 free NF- $\kappa$ B)의 동적 특성 변화를 시뮬레이션을 통해 살펴보았다. Hoffmann의 연구와 같이 NF- $\kappa$ B 조절 서브시스템에서는 I $\kappa$ B $\alpha$ 과 매개하는 강한 negative feedback loop로 인하여 NF- $\kappa$ B(n) 농도의 oscillation 현상이 관측된다. 특히 NF- $\kappa$ B(n)은 입력 IKK\*의 농도 및 지속 시간에 따라 서로 다른 주기와 진폭을 가지는 oscillation 출력을 보임을 확인하였다(그림 3).

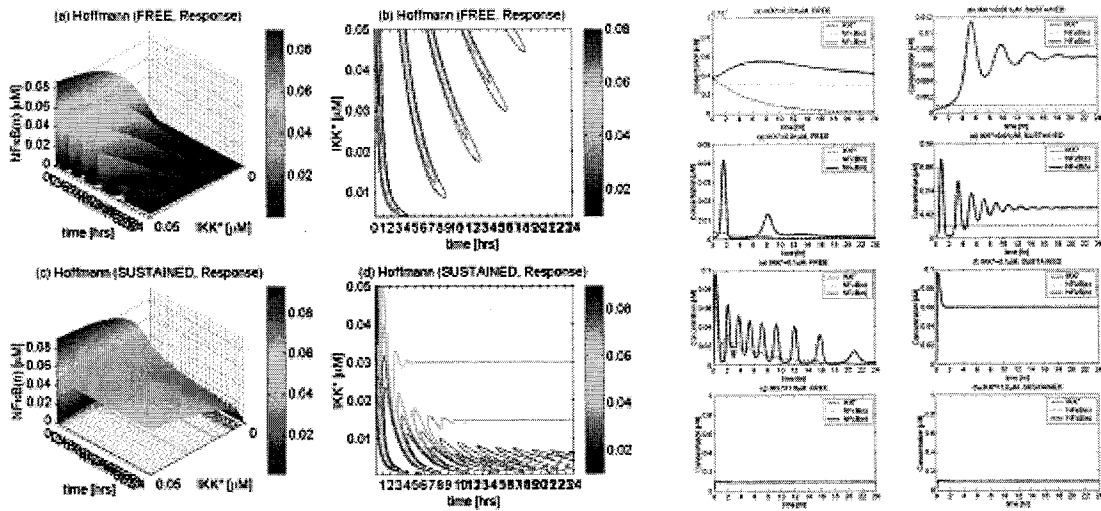


그림 3 NF- $\kappa$ B 조절 서브시스템 동역학 분석

Transcription factor NF- $\kappa$ B(n)가 갖는 다양한 출력 패턴은 유전자의 선택적 발현과 관련이 있다고 추정된다. 즉, NF- $\kappa$ B의 binding 시간 및 강도에 따라 서로 다른 발현율을 갖는 유전자들을 조절하는 메커니즘이 바로 negative feedback loop를 통해 구현되는 것이다.

서브시스템별 시뮬레이션 분석 결과를 종합하여 볼 때, 세포 외부에서 가해지는 TNF $\alpha$  신호에 대하여, receptor 서브시스템 및 kinase cascade 서브시스템의 구성 조합이 NF- $\kappa$ B 조절 시스템에 서로 다른 입력 IKK\* 입력 패턴을 유도하여 선택적으로 유전자 발현을 조절함을 유추해 볼 수 있다. 특히 kinase cascade 시뮬레이션에서 살펴보았듯이 kinase 농도 및 phosphatase 비율 등 모델 파라미터 변화가 서브시스템의 출력 패턴에 영향을 준다는 사실과, *in vitro* 실험을 통해 HBV 감염이 신호전달경로 구성 단백질의 농도에 변화를 준다는 사실은 바이러스에 의한 신호전달경로의 교란이 NF- $\kappa$ B activity의 현저한 변화를 일으킬 수 있음을 시사해주고 있다. 이에 앞서 구성한 신호전달경로 모델을 기반으로 HBV 영향에 관한 모델링 및 시뮬레이션을 진행하였다. 또한 이와는 별도로 신호전달경로 상의 주요 단백질을 억제하였을 때 NF- $\kappa$ B의 activation 변화를 측정하는 시뮬레이션도 진행하였다.

### ■ inhibition 효과 시뮬레이션

신호전달경로에 대한 *in silico* 모델은 생명 현상의 단순한 재현뿐만 아니라, 필요에 따라 실험자가 임의로 구성한 환경 하에서 신호전달경로에 발생하는 변화에 대한 예측을 가능하게 한다. 예를 들어 약물이나 바이러스 또는 독성 물질에 의해 단백질이 inhibition 또는 activation되었을 때 신호전달경로에 나타나는 동적 특성 변화를 *in silico* 시뮬레이션을 통해 어느 정도 예측해 볼 수

있다. 이는 특히 신약 개발 시, 최적의 투약 효과를 낼 수 있는 target 단백질 후보군을 *in silico* 시뮬레이션을 통해 미리 선정하여, 엄청난 비용과 시간이 드는 *in vitro/in vivo* 실험을 줄일 수 있게 해준다는 점에서 경제적으로 큰 의미를 가진다.

본 연구에서는 NF-κB 신호전달경로를 inhibition하는 최적의 단백질 노드를 찾기 위하여 inhibition 효과 시뮬레이션을 시행하였다. 이를 위해 앞서 구성한 기본 신호전달경로 모델을 바탕으로 kinase cascade의 주요 단백질 TRAFs complex와 IKKK, IKK에 대하여 dominant negative competitor type의 inhibitor를 모델링 하였다. (그림 4 붉은색 표시 부분)

특정 단백질을 inhibition 하여 얻은 NF-κB inhibition 효과를 측정하기 위해 Area Under Inhibition Curve(AUIC)를 다음과 같이 정의한다.

$$AUIC = \left[ 1 - \frac{\text{mean}(NF\kappa B_n)}{\text{(steady state } NF\kappa B_n \text{ in the absence of inhibition)}} \right] \times 100 (\%)$$

즉, AUIC는 특정 단백질을 일정 시간 inhibition 하였을 때, 그 시간 동안 정상 상태보다 감소하게 되는 NF-κB activity의 비율을 나타낸다.

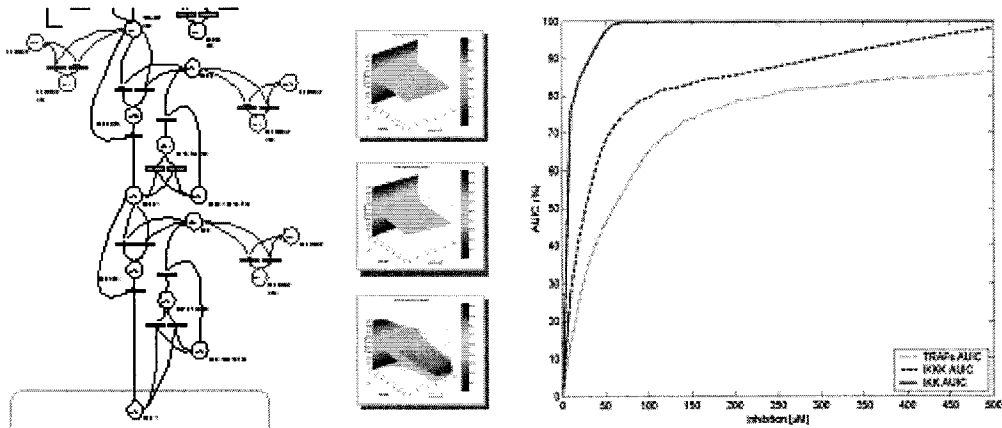


그림 4 신호전달경로 상 단백질의 inhibition 모델(좌) 및 효과 측정AUIC 분석(우)

시뮬레이션 결과 IKK를 inhibition하였을 때, NF-κB가 최대로 inhibition되는 것을 볼 수 있다. 특히 IKK inhibition은 IKKK나 TRAFs에 비하여 상대적으로 빠른 시간 내에 100% inhibition에 도달함을 그래프를 통해 살펴볼 수 있다. 이는 치료 목적으로 NF-κB activation을 감소시킬 필요가 있을 때, IKK를 inhibition하는 molecule을 약물로 설계하는 것이 효과적임을 시사해 준다.



## ■ HBV 감염 효과 시뮬레이션

HBV 감염으로 인한 NF- $\kappa$ B activation의 증가를 *in silico* 상에서 모델링하기 위하여 기본 모델을 다음과 같이 확장하였다.

- 제1세부과제팀의 *in vitro* 실험 결과를 통하여, HBV가 간세포 내 phosphorylated IKK를 증가시킴을 확인하였다. 여기에 HBV가 NIK를 통하여 NF- $\kappa$ B activity가 현저히 증가시킨다는 보고를 바탕으로, 모델에 바이러스로부터 기인하는, NIK를 통한 IKK activation 경로를 추가하였다. 또한 literature를 근거로 TRAFs complex를 통한 NIK 억제 작용도 모델링 하였다.
- 제1세부과제팀의 HBV 감염 세포에 대한 TNF $\alpha$  반응 실험은 HepG2 및 HepG2.2.15 세포를 대상으로 이루어졌다. 정상 간암 세포인 HepG2 세포와 감염 세포인 HepG2.2.15 세포를 정확히 모델링하기 위하여, *in vitro* 실험에 근거, HepG2 및 HepG2.2.15의 미세한 IKK activation 경로를 추가하였다.
- *in vitro* 실험을 통해 밝혀진 HepG2와 HepG2.2.15의 서로 다른 초기 농도 구성을 모델에 적용하였다.

즉, 간암으로 기인한 미세 IKK activation은 임의의 단백질 X로 인한 IKK phosphorylation으로, HBV로 기인한 NIK 및 IKK activation은 임의의 단백질 Y로 인한 NIK phosphorylation과 phosphorylated NIK로 인한 IKK phosphorylation으로 모델링하였다. (그림 5 붉은색 부분)

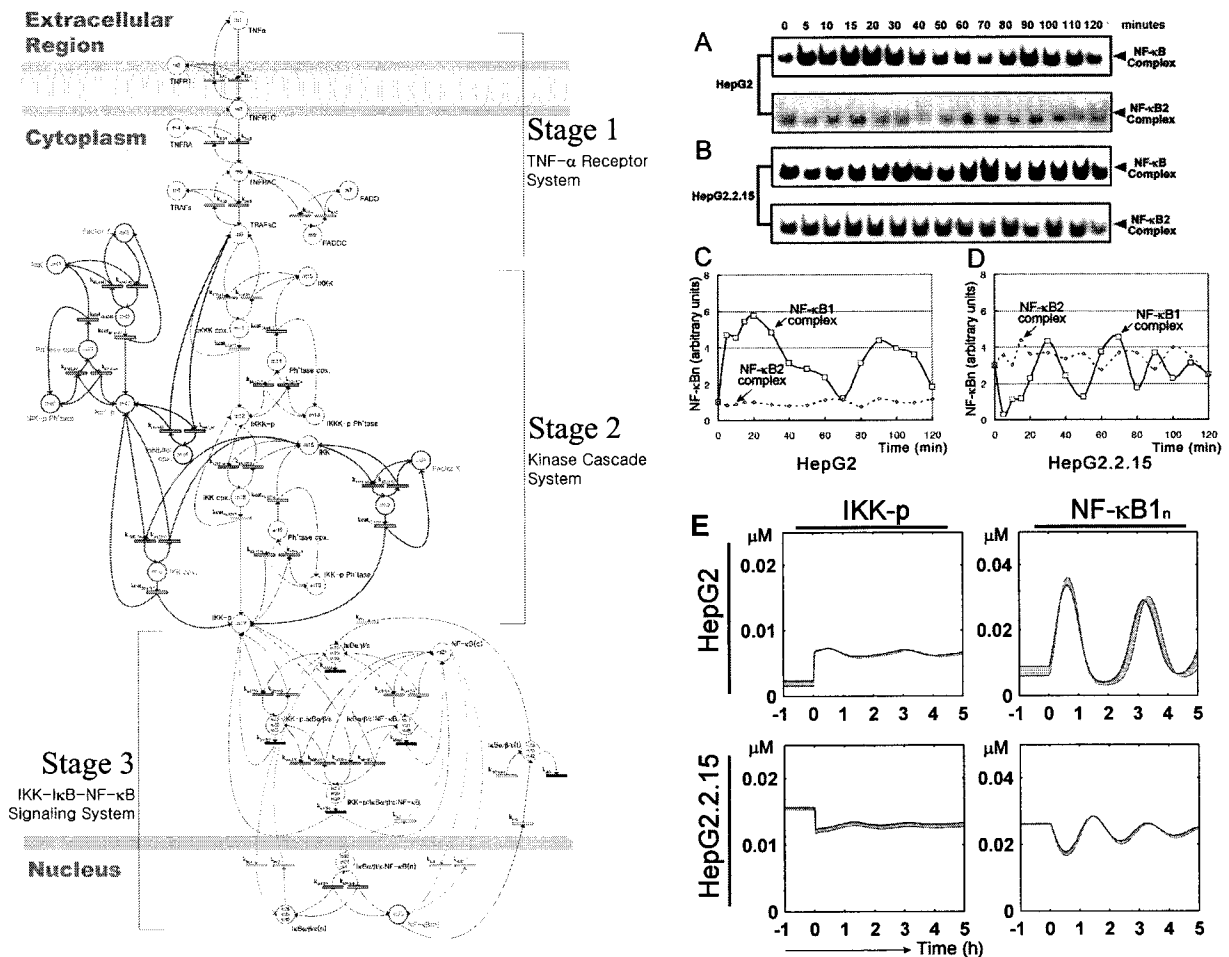


그림 5 HepG2 및 HepG2.2.15 모델과 *in vitro/in silico* 실험 (A,B) TNF $\alpha$  투여에 의한 NF- $\kappa$ B 반응 EMSA 이미지 (C,D) EMSA 결과를 정량적으로 표현한 그래프 (E) *in silico* 시뮬레이션 결과

*in vitro* 실험은 TNF $\alpha$  투여 후 2시간 동안 세포 내 NF- $\kappa$ B1 activity를 일정 시간 간격으로 EMSA 분석하는 방식으로 이루어졌다(그림 5A, B). 그 결과 정상 세포인 HepG2에서는 NF- $\kappa$ B1 activity가 진폭 및 주기가 큰 뚜렷한 oscillation을 띄는 형태로 나타났다(그림 5C). 반면에 HepG2.2.15에서는 진폭 및 주기가 적은 oscillation이 발생함을 볼 수 있다. 특히 TNF $\alpha$  투여 초기 NF- $\kappa$ B1 activity가 오히려 감소하는 undershoot 현상이 관측됨을 살펴볼 수 있다(그림 5D).

*in silico* 시뮬레이션을 통해서도 이와 유사한 패턴의 NF- $\kappa$ B activity가 나타났다(그림 5E). HepG2의 경우 크고 뚜렷한 진폭 및 주기를 갖는 oscillation 형태의 출력이 보이며, 반면에 HepG2.2.15의 경우 작고 좁은 진폭과 주기를 갖는 oscillation 출력이 나타나고 있다. 이는 NF- $\kappa$ B 조절 시스템에 들어오는 입력 phosphorylated IKK가 HepG2.2.15 세포에서는 HBV의 영향으로 정상 세포보다 많이 activation 되어 NF- $\kappa$ B 조절 시스템의 출력 형태가 변화될 수 있음을 시사한다.

또한, HBV 감염 세포인 HepG2.2.15에서 특징적으로 발견되는 undershoot 현상은 TNF $\alpha$  투여 초기 신호전달경로 상에 있는 TRAFs complex의 증가가 바이러스에 의해 활성화된 NIK를 억제하고, 이로 인해 IKK activation이 감소하는 동적 변화에 기인하는 것으로 보인다. 이는 *in silico* 모델을 통하여, 간염 세포의 신호전달경로의 동역학 특성 변화에 대한 의미 있는 가설을 설정하고, 시뮬레이션과 *in vitro* 실험을 통하여 이를 검증할 수 있음을 보여주는 hypothesis driven research의 좋은 사례라 할 수 있다.

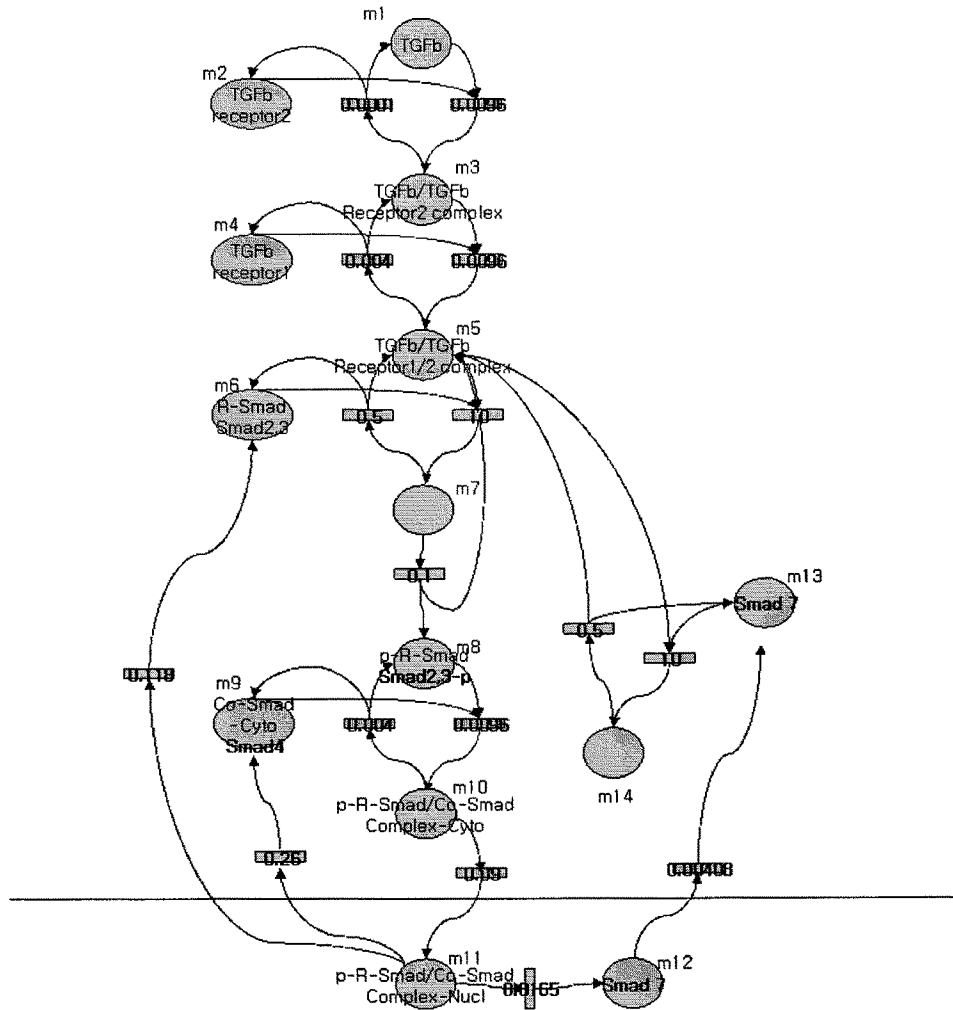
### 3. HCV 관련 TGF $\beta$ 신호전달경로에 대한 ODE 모델 표현

#### 가. TGF $\beta$ 신호 전달 경로

TGF $\beta$ 는 receptor 1, 2와 상호 작용하여 TGF $\beta$ /TGF $\beta$  receptor 1/2 complex를 이루며 이것이 R-Smad인 Smad2, 3를 phosphorylation시킨다. 다음에는 이것이 Co-Smad-Cyto Smad 4와 반응하여 p-R-Smad/Co-Smad Complex-Cyto를 이루게 된다. p-R-Smad/Co-Smad Complex-Cyto이 세포질에서 핵 속으로 이동하여 p-R-Smad/Co-Smad complex-nucleus가 된다. 다음에는 p-R-Smad/Co-Smad complex-nucleus가 target 유전자의 promoter binding하여 target 유전자의 transcription/translation이 이루어지고, R-Smad와 Co-Smad는 다시 세포질로 translocation 되는 과정을 거친다. 이렇게 하여 생겨난 Smad 7이 Smad 2,3의 competitive inhibitor로 작용하게 된다. 이와 같은 과정으로 거쳐 TGF $\beta$  신호 전달 경로가 완성된다.

#### 나. ODE 모델의 구성

앞에서 말한 HBV에서 모델을 구성했던 것과 비슷한 방법으로 HCV 네트워크의 전체 신호전달 경로 *in silico* 모델을 완성하였다. 그 다음 이 모델을 MATLAB을 이용하여 시뮬레이션 작업을 수행하고, 그 결과를 제 2세부과제 팀과 연계하여 필요한 실험을 진행하고 그 결과를 반영하여 더 나은 모델을 만드는 작업을 수행하였다.



#### 다. 초기 농도 및 파라미터 결정

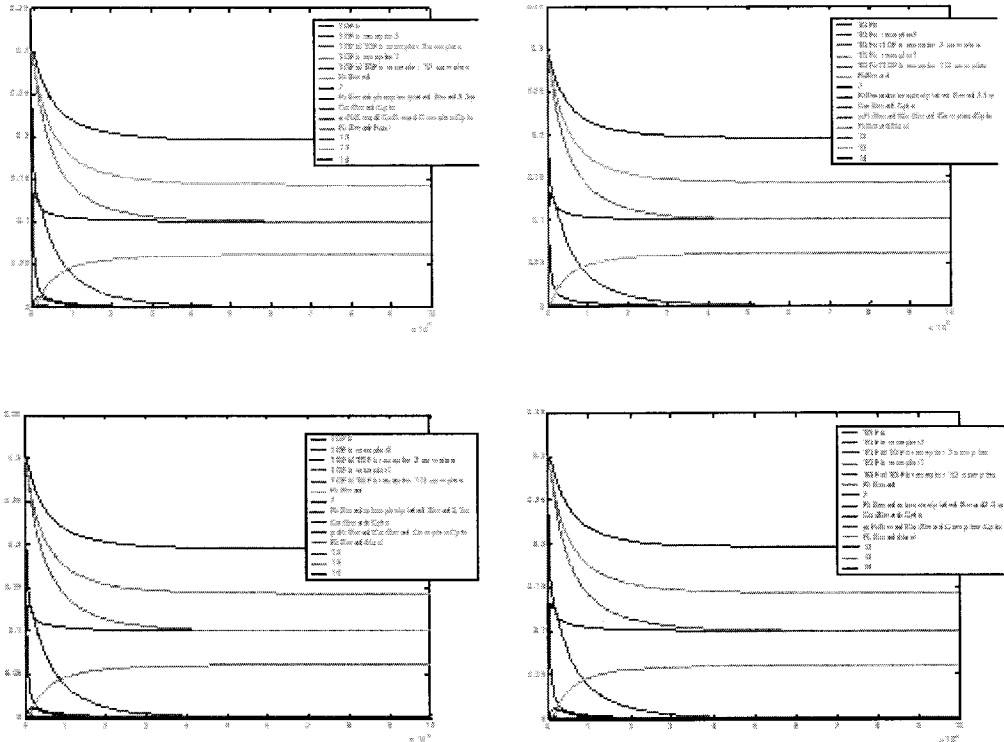
위와 같은 모델은 반응 상수와 초기 농도에 따라 서로 다른 시뮬레이션 결과를 보인다. 따라서 적절한 파라미터 값과 초기 값을 정하는 것이 매우 중요하다. 본 세부연구과제에서는 제2세부과제 팀과의 협력을 통하여 다음과 같은 초기 농도 값을 정하였다. 그 외에 모델에서 사용된 파라미터 값들은 literature를 참고하였다.

TGFβ	TβR-2	TβR-1	R-Smads	Co-Smad
0.2	0.3	0.3	0.3	0.3

*in vitro* 실험을 통한 모델 초기 조건 측정 (단위 *nM*)

#### 라. *in silico* 시뮬레이션

위의 모델을 MATLAB으로 구성하여 실험한 결과 아래와 같은 결과를 얻을 수 있었다.



HCV 관련 신호전달경로 모델 *in silico* 시뮬레이션 결과

위의 결과는 모델에 사용된 파라미터 값을 다양하게 변화시켜 가면서 실험한 결과 그림들이다. 파라미터 값이 바뀔 때마다 시뮬레이션 결과가 바뀌지만, 대부분의 경우 모든 물질이 평형 상태에 이르는 것으로 알아볼 수 있었다. 현재 제2세부과제팀과의 긴밀한 연계를 통하여 위의 모델을 개선해나가는 작업을 진행하고 있으며, HBV 모델에서 진행했던 것과 유사한 방향으로 실험을 진행될 것이다.

### 제 3 절 관련 데이터베이스 설계

#### 1. SBML

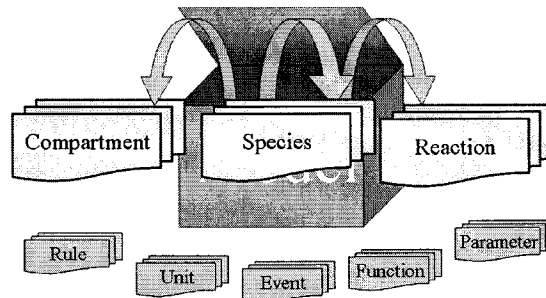
현재 HBV, HCV 모델과 같은 생물학 네트워크 모델을 표현할 수 있는 다양한 프로그램이 존재한다. 이 때 각 프로그램에서 네트워크 모델을 표현하는데 서로 다른 방법을 이용한다면 한 프로그램의 결과물을 다른 프로그램에서는 사용할 수 없을 것이다. 따라서 생물학 모델을 표현하는 표준 방법이 필요하게 되었다. 이와 같은 목적으로 개발 된 것이 SBML이다. SBML은 Systems Biology Markup Language의 약자로서, biochemical network 모델을 표현하는 표준 방법을 제공한다.

SBML은 다음과 같은 장점을 가지고 있다.

- XML을 기반으로 만들어져 있어서 확장성이 우수하다.
- Biochemical network을 표현하는 표준을 제시함으로써, 다양한 소프트웨어의 결과물을 서로

다른 소프트웨어에서 사용하는 것이 쉽게 이뤄질 수 있다.

SBML은 다음 그림과 같은 여러 구성 요소로 이루어져 있다.

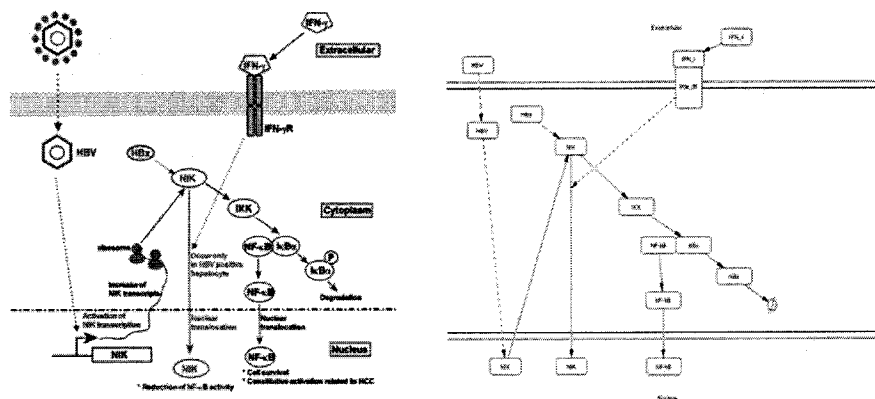


- Unit : SBML에서 사용되는 각종 원소들의 단위 (몰, mg 등)
- Compartment : species가 위치하는 볼륨을 나타냄
- Species : network 내의 proton, calcium와 같은 간단한 ion, glucose, ATP과 같은 간단한 molecule 그리고 RNA, protein과 같은 큰 molecule을 표현함.
- Parameters : network을 표현하는 데 필요한 값들을 선언
- Rules : network 내의 species들은 상태가 변할 수 있다. 이때 제약 조건이 있다면 이것을 선언할 수 있는데, 이런 제약조건을 Rule이라고 한다.
- Reaction : species들끼리 영향을 주고 받는 관계를 나타낸다. 즉 어떤 species의 양이 변하거나 compartment를 옮기거나 species의 상태가 바뀌는 관계를 표현한다.
- 이외 Function, Event와 같은 요소가 존재한다.

SBML은 현재 계속 개선되고 있다. 그리고 50가지가 넘는 다양한 소프트웨어에서 biochemical network을 표현하는데 SBML을 사용하고 있으며, 앞으로는 더 많은 소프트웨어에서 SBML이 사용될 것으로 예상된다.

본 연구에서는 제1세부과제와 제2세부과제의 HBV, HCV 네트워크를 design 한 후, SBML tool인 Cell-designer를 이용하여 그림을 그려보았다.

다음은 제1세부과제의 network model과 그것을 SBML로 표현한 그림이다.



제2세부 과제의 HCV 네트워크 그림과 SBML로 표현한 그림

위와 같이 그림을 그리면, Cell-designer와 같은 소프트웨어에서는 다음과 같은 SBML을 생성하여 준다.

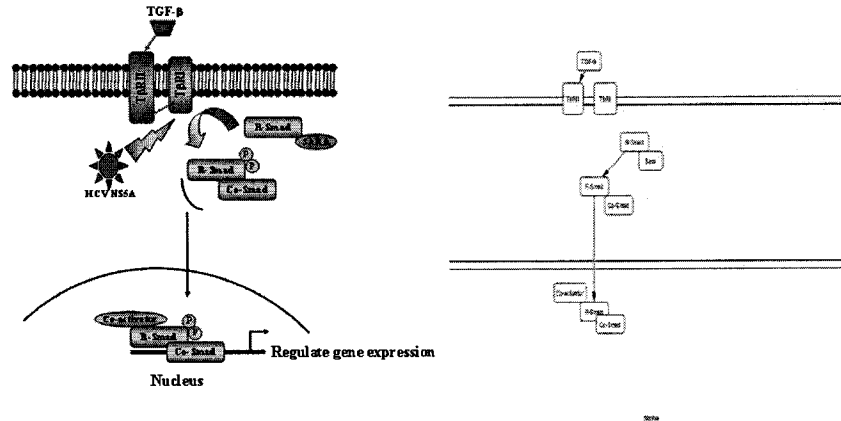
```

<?xml version="1.0" encoding="UTF-8" ?> <!-- Created by SBML API 2.0(a15.2) -->
<sbml xmlns="http://www.sbml.org/sbml/level1" level="1" version="2" xmlns:html="http://www.w3.org/1999/xhtml"
xmlns:celldesigner="http://www.sbml.org/2001/ns/celldesigner">
<model>
<annotation>
<celldesigner:modelVersion>2.2</celldesigner:modelVersion>
<celldesigner:modelDisplay sizeX="1000" sizeY="700" />
<celldesigner:listOfCompartmentAliases>
<celldesigner:compartmentAlias id="ca3" compartment="c3">
<celldesigner:class>SQUARE_CLOSEUP_SOUTH</celldesigner:class>
<celldesigner:point x="416.0" y="124.0" />
</celldesigner:compartmentAlias>
<celldesigner:compartmentAlias id="ca4" compartment="c4">
<celldesigner:class>SQUARE_CLOSEUP_NORTH</celldesigner:class>
<celldesigner:point x="516.0" y="569.0" />
</celldesigner:compartmentAlias>
</celldesigner:listOfCompartmentAliases>
<celldesigner:listOfSpeciesAliases>
<celldesigner:speciesAlias id="a1" species="s1">
<celldesigner:activity>inactive</celldesigner:activity>
<celldesigner:bounds x="145.0" y="155.0" w="70.0" h="30.0" />
</celldesigner:speciesAlias>
<celldesigner:speciesAlias id="a12" species="s14">

```

HBV 네트워크 모델의 SBML 명세

다음은 제2세부과제의 결과를 제1세부과제와 마찬가지로 SBML로 표현한 그림이다.



제2세 부과제의 HCV 네트워크 그림과 SBML로 표현한 그림

```

<?xml version="1.0" encoding="UTF-8" ?> <!-- Created by SBML API 2.0(a15.2) -->
<sbml xmlns="http://www.sbml.org/sbml/level1" level="1" version="2" xmlns:html="http://www.w3.org/1999/xhtml"
xmlns:celldesigner="http://www.sbml.org/2001/ns/celldesigner">
<model>
<annotation>
<celldesigner:modelVersion>2.2</celldesigner:modelVersion>
<celldesigner:modelDisplay sizeX="1000" sizeY="700" />
<celldesigner:listOfCompartmentAliases>
<celldesigner:compartmentAlias id="ca3" compartment="c3">
<celldesigner:class>SQUARE_CLOSEUP_SOUTH</celldesigner:class>
<celldesigner:point x="468.0" y="191.0" />
</celldesigner:compartmentAlias>
<celldesigner:compartmentAlias id="ca4" compartment="c4">
<celldesigner:class>SQUARE_CLOSEUP_NORTH</celldesigner:class>
<celldesigner:point x="428.0" y="439.0" />
</celldesigner:compartmentAlias>
</celldesigner:listOfCompartmentAliases>
<celldesigner:listOfSpeciesAliases>
<celldesigner:speciesAlias id="a18" species="s17" compartmentAlias="ca3">
<celldesigner:activity>inactive</celldesigner:activity>
<celldesigner:bounds x="283.0" y="156.0" w="50.0" h="50.0" />
</celldesigner:speciesAlias>
<celldesigner:activity>inactive</celldesigner:activity>

```

HCV 네트워크 모델의 SBML 명세

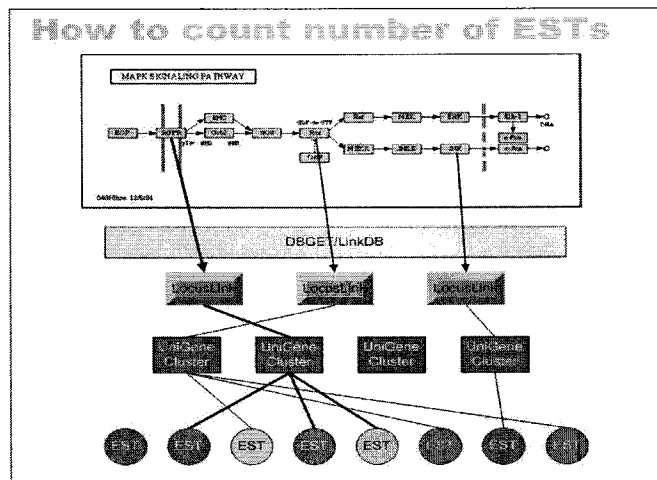
## 2. EST 정보를 이용한 패스웨이 검색 툴 개발

### 가. 연구 목적 및 내용

본 연구에서는 KEGG PATHWAY 데이터베이스에서 제공하는 pathway 그래프를, EST data와 융합하여, 그래프 상의 각 노드에 해당하는 단백질들의 발현 패턴을 tissue source별로 선택, 비교·분석할 수 있는 소프트웨어 도구를 개발하였다.

KEGG PATHWAY 데이터베이스에는 현재까지 알려진 metabolic pathway, regulatory pathway, molecular complex 등이 그래프 형태로 저장되어 있다. 현재의 데이터베이스는 생물체(organism) 별로 pathway 그래프를 제공하며, 노드에 해당하는 단백질에 관련된 외부 데이터베이스 링크를 제공하고 있다.

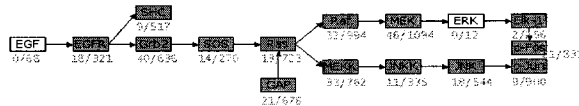
본 연구에서는 각 단백질 노드에 관련된 외부 데이터베이스를 자동으로 탐색하여, 해당 단백질과 관련된 EST sequence를 조사함으로써, 각 단백질의 발현 패턴을 EST의 개수로 정량화하여 보여줄 수 있도록 하였다. 특히, 특정 tissue source를 선택하여 EST를 계수하고, 이를 다른 tissue source의 결과와 비교함으로써, tissue source에서의 pathway 그래프 특성을 살펴볼 수 있다. 다음 그림은 이 과정을 진행하기 위해 데이터베이스를 검색하는 과정을 나타내고 있다.



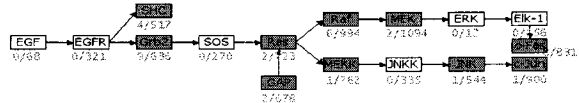
KEGG PATHWAY 데이터베이스 내 MAPK signalling pathway 상의 단백질들이 정상 간 세포(liver)와 간암 세포(hepatocellular carcinoma)에서 어떤 발현 패턴을 보이는지, 위의 그림과 같이 EST 개수를 검색한 결과는 다음과 같다. 아래 그래프에서 노드가 가지는 숫자는 총 EST 개수와 지정된 tissue source로 검색된 EST 개수를 나타내고 있다

#### MAPK SIGNALING PATHWAY (Comparison)

##### Liver (or Liver and Spleen)



##### Hepatocellular carcinoma



## 나. 관련 database

EST(Expressed Sequence Tag)는 오른쪽 그림과 같이 mRNA를 역전사하여 얻어지는 cDNA 클론 분석을 통해 얻어진 짧은 서열이다.

세포 안의 mRNA가 단백질로 발현되어지는 유전자를 나타내기 때문에 cDNA 역시 세포 안에서 실제 발현되는 유전자의 서열을 나타낸다. EST는 인간, 생쥐를 포함한 여러 생물체의 다양한 세포 조직에서 얻어진 cDNA 클론을 자동화된 기계로 읽어서 서열을 결정한 것으로, 다소 부정확할 수 있지만 기계로 빠르고 값싸게 얻을 수 있으며, 대량으로 데이터베이스에 축적될 수 있다는 점에서, 여러 가지 연구에 응용되고 있다.

현재 NCBI에서 관리하는 dbEST 데이터베이스에는 600여종의 생물체(organism)에서 서열화한 20,685,791개의 EST가 공개되어 있다. (dbEST release 040904, 2004년 4월 9일 현재)

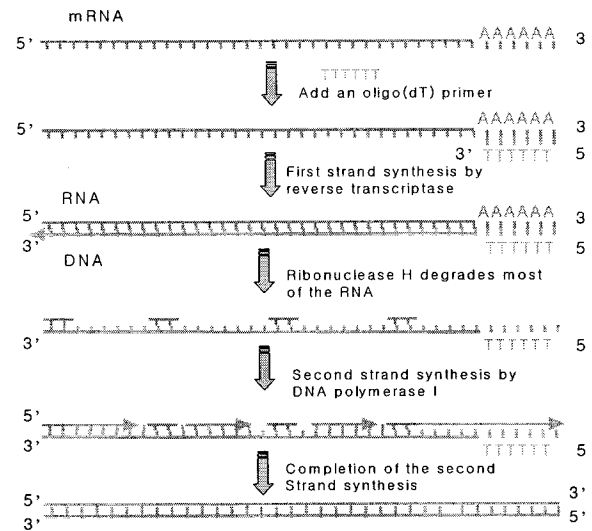
UniGene은 인간, 생쥐, 쥐, 암소 등으로부터 유래한 GenBank의 염기 서열을 유전자를 기준으로 서로 중복되지 않도록 클러스터화한 데이터베이스이다. UniGene에는 EST 서열에 대해서도 클러스터를 만들어 두고 있어서 새로운 유전자의 발굴과 유전자 지도 작성 등의 연구에 편의를 제공하고 있다. UniGene은 유전자를 기준으로 작성되기 때문에 염기 서열의 획득 뿐만 아니라 유전자의 발현 정보와 유전자 지도 정보 등을 손쉽게 얻을 수 있다는 점에서 매우 유용하다.

LocusLink 데이터베이스는 특정 유전자 위치에 대한 curated sequence와 정보, 링크를 제공하는 데이터베이스로, 본 연구에서는 KEGG PATHWAY 데이터베이스와 UniGene 데이터베이스를 연계하는 관계 데이터베이스로 사용되었다.

KEGG PATHWAY 데이터베이스는 Kyoto 대학의 GenomeNet 프로젝트의 일환으로, Metabolism, Genetic Information Processing, Environmental Information Processing, Cellular Process, Human Diseases 관련 pathway에 대한 종합적 지식 데이터베이스를 갖추고 있다. 또한 DBGET/LinkDB를 통하여, 자체 데이터베이스 및 외부 데이터베이스에 대한 검색/링크 기능을 제공하고 있다.

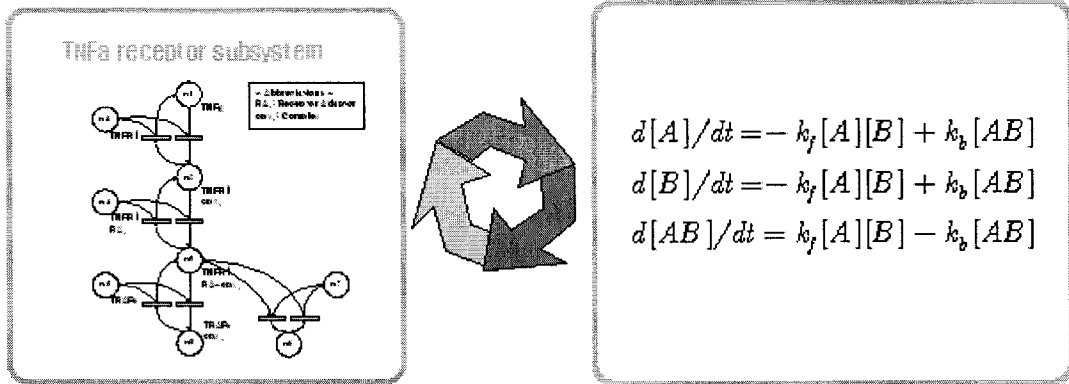
## 3. ODE 모델 명세 언어

생명 현상을 시스템적으로 이해하기 위해서는 생명 현상에 대한 수학적 모델링이 필수적이다. 이러한 수학적 모델링에는 상미분 방정식 (Ordinary Differential Equation, ODE)이 많이 사용된다. 생명 현상은 다음과 같이 그래프 형태로 표현가능하다. ODE 모델에 대한 그래프 표현(아래 그림)은 신호전달 경로 상의 단백질 반응 구조를 직관적으로 설계하는데 유용하다. 그러나 시뮬레이션을 위해서는 이를 컴퓨터가 이해할 수 있는 방정식 표현으로 변환해야만 한다. 그래프에 비해 모델에 대한 방정식 표현은 모델의 확장 및 수정이 어렵고, 반응 물질들 간의 관계를 직관적으로 이해하기 불가능하다는 단점을 가지고 있다. 이러한 차이로 인해, 모델의 크기가 커지고 단백질





간의 반응 구조가 복잡해질수록 그래프에서 방정식으로의 변환은 점점 더 곤란하게 된다.



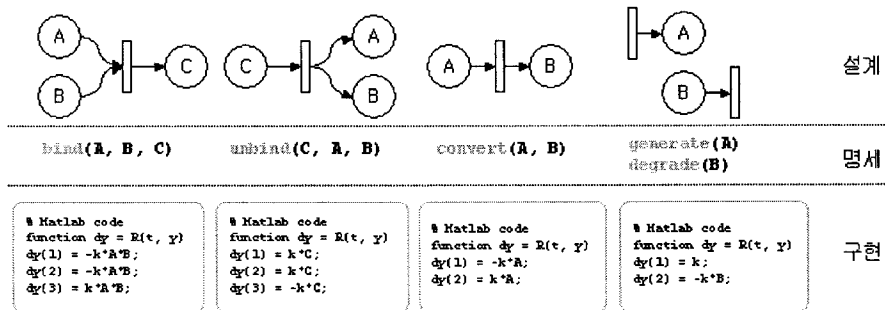
ODE 모델의 그래프(좌) 및 방정식 표현(우)

객체지향 패러다임에서 "추상화"를 통해 설계와 구현을 분리하여, 유연성과 확장성을 얻듯이 ODE 모델 명세 언어라는 추상화 단계를 추가함으로써 그래프 표현(설계)과 방정식 표현(구현)을 분리하여, 위와 같은 문제점을 극복 할 수 있다. 본 논문에서 제안하는 ODE 모델 명세 언어는 다음과 같은 특성을 지닌다.

- ① 단백질 간의 반응 구조를 다음과 같은 관계 연산으로 나타냄으로써 직관적으로 이해하기 쉽고, 수정이 용이하다.

관계연산자: *relation(Entity1, Entity2, ...)*

- ② 각 관계 연산마다 해당하는 반응 방정식 모델을 설정할 수 있도록 한다.



ODE 모델 명세 언어

위의 그림은 ODE 모델 명세 언어가 적용되는 예를 보여준다. 앞서 ODE 모델 그래프 표현에서 나타나는 모든 그래프 구조는 위와 같은 명세 언어로 표현될 수 있다.

모델 명세 언어를 MATLAB 언어로 번역하는 자동화된 프로그램을 도입한다면, ODE 모델을 보다 빠르고 이해와 수정이 용이하도록 구성할 수 있을 것이다

## 제 4 절 생물정보학 기초 기술 연구

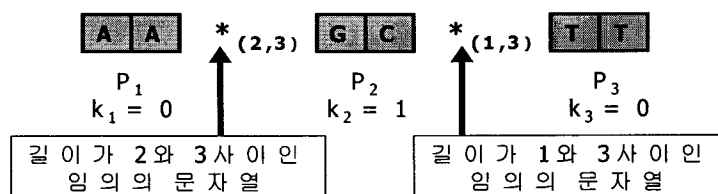
본 연구에서는 아래와 같은 세 가지 분야에서 다양한 연구 결과를 얻었다.

- 효율적인 패턴매칭 알고리즘의 개발
- 효율적인 서열정보의 정렬 기법의 개발
- 인덱스 자료 구조에 관련된 연구

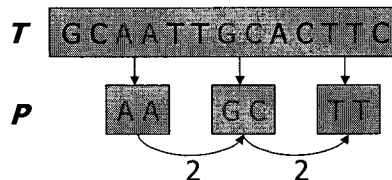
### 1. Finding approximate occurrence of a pattern that contains gaps

본 연구에서 푸는 문제는 아래의 예와 같이, 고정된 범위 내에서 길이가 변할 수 있는 갭(gap)을 가지는 패턴을 텍스트에서 찾는 문제이다.

$$\text{Pattern } P = AA^{*(2,3)}GC^{*(1,3)}TT$$



즉, 아래 그림에서와 같이 텍스트  $T = \text{GCAATTGCACTTC}$ 에 위와 같은 패턴  $P$ 가 나타남을 볼 수 있다.



이 문제는 여러 가지 응용을 가지고 있으나, 주된 응용은 바이오인포매틱스에서 단백질의 기능을 알아낼 때이다. 새로운 단백질의 역할은 기존에 이미 역할이 알려진 단백질과 서열의 유사성을 찾아봄으로 추론할 수 있다. 이때 전체 서열과 비교 대신, 같은 기능을 갖는 단백질 서열들로부터 추출한 모티프(motif)라고 불리는 짧은 중요한 서열과의 비교가 필요하다. 모티프는 위에서 보인 것과 같이 길이의 범위가 고정된 갭을 가진 패턴으로 표현되며, 최근에는 모티프 정보를 저장하는 PROSITE 등과 같은 데이터베이스가 축적되고 있으며 사람들의 관심이 높아지고 있다.

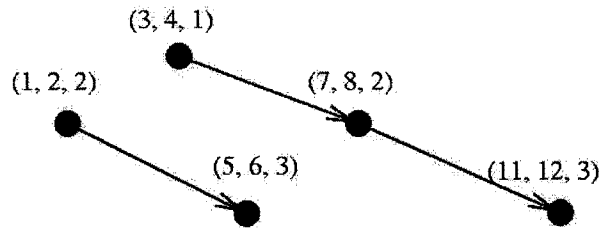
실제 실험에서는 데이터가 오류를 포함할 수 있기 때문에, 한계치 이내의 오차를 허용하여, 텍스트에서 패턴과 근사한 부분을 찾는 것이 중요하게 된다. 본 연구에서는 텍스트에서 패턴이 정확하게 일치하는 부분을 찾는 것 뿐 아니라, 이와 같이 근사하게 일치하는 부분을 찾는 방법을 제시하였다.

이를 위하여 본 연구에서는 갭을 포함하는 패턴을 찾는 문제를 그래프로 바꾸어 풀 수 있음을 보였다. 각각의 부분 패턴이 나온 곳을 그래프에서 정점으로 나타내고, 인접한 부분 패턴들을 나타내는 두 정점 간의 거리가 이들 사이에 있는 갭의 범위를 만족할 경우 두 정점을 간선으로 연결한다. 그 예는 아래 그림과 같다.

1 2 3 4 5 6 7 8 9 10 11 12 13

---

G C A A T T G C A C T T C

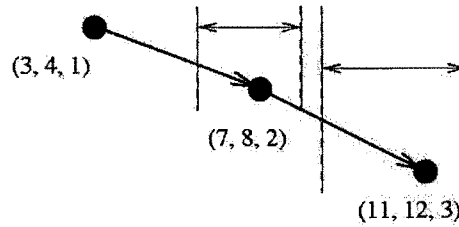


이 과정에서, 첫 번째 부분 패턴을 나타내는 정점부터 시작하여 마지막 부분 패턴을 나타내는 정점까지 이어진 간선이 있다면 텍스트가 전체 패턴을 포함함을 알 수 있다. 이 과정을 보다 효율적으로 풀기 위해서는 다음과 같은 과정을 거친다. 먼저 첫 번째 부분 패턴만을 텍스트에서 찾는다. 찾은 첫 번째 부분 패턴들에서, 겹을 고려하였을 때 두 번째 부분 패턴이 나올 수 있는 영역을 찾고, 이 영역에서 두 번째 부분 패턴들을 찾는다. 이 과정을 계속 반복하여 마지막 부분 패턴을 찾게 되면, 전체 패턴을 찾은 것이 된다.

1 2 3 4 5 6 7 8 9 10 11 12 13

---

G C A A T T G C A C T T C



오류를 감안한 근사적인 패턴을 찾는 경우도 위의 방법과 비슷하게 진행한다. 텍스트에서 정확하게 일치하는 부분 패턴을 찾는 대신, 동적 계획법(dynamic programming)을 이용하여 차례대로 각각의 부분 패턴마다 허용된 범위 이내의 오차를 가지고 출현하는 위치를 찾는다. 그 후 다음 위치에 있는 부분 패턴이 나올 수 있는 범위를 계산하고, 이 범위 내에서 다음 부분 패턴들을 계속 찾아나간다. 이 과정에서 최악의 경우 시간 복잡도는 패턴의 길이와 텍스트의 길이의 곱에 비례한다.

## 2. Linear-time search in suffix arrays

생명정보학에서는 많은 서열 정보를 다루게 된다. 이들 서열 정보에 관련된 문제를 효율적으로 풀기 위해서 인덱스 자료 구조가 필요하게 된다. 대표적인 인덱스 자료 구조로는 접미사 트리와 접미사 배열이 있다. 접미사 트리는 주어진 서열의 모든 접미사를 저장하는 트리로서 생명정보학 전반에서 매우 유용하게 쓰인다.

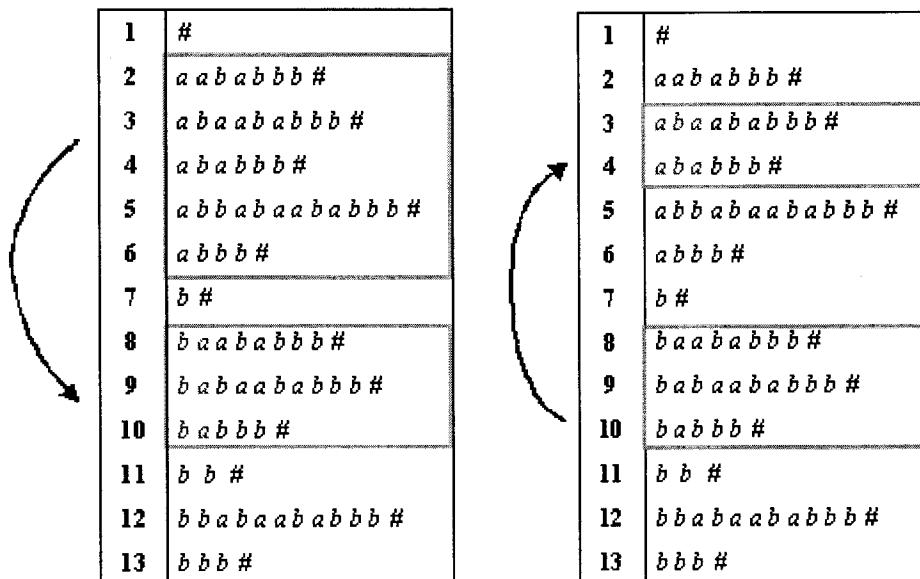
반면, 접미사 트리를 저장하는 데에는 많은 메모리가 필요하다는 단점이 있다. 따라서, 접미사 트리와 같은 역할을 하면서 메모리를 효율적으로 이용하는 자료 구조로 접미사 배열이 제안되었다. 아래 그림은 T=abbabaababbb#에 대한 접미사 배열의 예이다.

1	#
2	a a b a b b b b #
3	a b a a b a b b b b #
4	a b a b b b b #
5	a b b a b a a b a b b b b #
6	a b b b b #
7	b #
8	b a a b a b b b b #
9	b a b a a b a b b b b #
10	b a b b b b #
11	b b #
12	b b a b a a b a b b b b #
13	b b b #

접미사 배열은 접미사 트리에 비하여 적은 공간을 차지하나, 텍스트에서 패턴을 찾는 문제에 대해서는 접미사 트리를 이용한 방법이  $O(p \log |\Sigma|)$  시간이 걸리는 반면, 접미사 배열을 이용하면  $O(p|\Sigma|)$  시간이 걸린다는 단점이 있었다. 여기에서  $p$ 는 패턴의 길이이며,  $|\Sigma|$ 는 알파벳의 크기이다.

본 연구에서는 접미사 배열을 이용하면서도 접미사 트리와 같이  $O(p \log |\Sigma|)$  시간에 패턴을 찾을 수 있는 방법을 제시하였다.

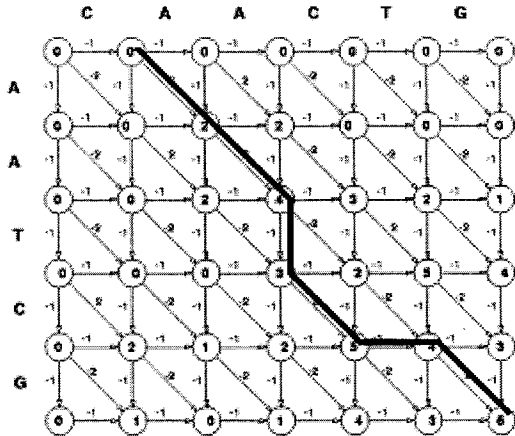
이를 위하여 본 연구에서는 패턴을 오른쪽에서 왼쪽으로 찾아나가면서 탐색을 진행한다. 즉, 패턴  $P = aba$ 인 경우,  $a \rightarrow ba \rightarrow aba$ 의 순으로 진행해나간다.



이 과정을 효율적으로 진행하기 위해서, 알파벳의 각각의 문자마다 접미사 배열에서 가장 마지막으로 나타난 위치를 기록한다. 또한 접미사 배열에 저장된 접미사의 바로 앞에 오는 문자를 이용한다.

### 3. An Efficient Local Alignment Algorithm for Masked Sequences

Local alignment algorithm은 주어진 두 서열에서 서로 유사한 부분을 찾는 알고리즘이다. 예를 들어, 아래 그림처럼 두 서열 AATCG와 CAACTG가 주어지면, AATC-G와 AA-CTG가 가장 유사한 부분임을 찾아준다. Local alignment algorithm은 DNA 시퀀싱 프로그램을 비롯한 다양한 곳에 응용되어진다. 가장 잘 알려진 local alignment algorithm은 Smith와 Waterman이 제시한 것으로 길이가 각각  $n$ ,  $m$ 인 두 서열  $a$ ,  $b$ 가 주어지면  $O(nm)$  시간에 local alignment를 찾는다.



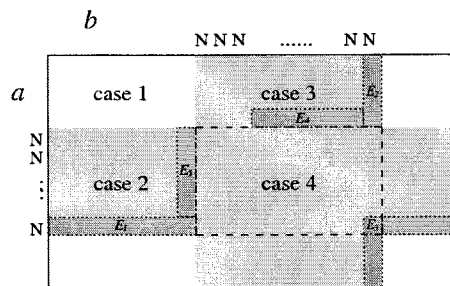
한편, DNA 시퀀싱에 사용되는 fragment들은 여러 가지 이유로 일부 base들의 정보가 없는 경우가 많다. 알려진 repeat들과 low-quality base들은 screen되어 사라져 버리기도 하고, base calling 단계에서 base의 값이 결정되지 않은 base들도 있다. 이러한 이유들로 mask된 부분은 N으로 표시된다. Mask된 부분들은 base 정보가 없기 때문에 alignment에서 의미가 없게 된다.

본 연구에서는, mask된 부분을 건너뛰면서, local alignment를 찾는 효율적인 알고리즘을 제시한다.

제시한 알고리즘은 mask된 부분을 건너뛰지만 Smith-Waterman 알고리즘과 동일한 local alignment를 찾는다. 먼저 mask된 부분을 건너뛰기 위해 mask된 부분에서 사용할 수 있는 새로운 recurrence를 제시하고, 새로운 recurrence를 이용한 효율적인 알고리즘을 제시한다.

$a_{i-1..i}$ 가 mask된 부분인 경우, 본 연구에서 제시하는 새로운 recurrence는  $H_{i-t, j-t}$ 부터  $H_{i-t, j}$ 까지만 있으면  $H_{ij}$ 를 계산할 수 있다. 즉, 이를 이용함으로써 mask된 부분을 계산하지 않고 건너뛸 수 있다.

Mask된 부분에 따라 아래와 같이 네 가지의 case로 나누어 생각한다. 먼저, case 1 영역은 기존 Smith-Waterman 알고리즘과 동일하게 계산한다. 다음, case 2와 case 3의 경우 마지막 라인인  $E_1$ 과  $E_2$ 만 계산한다. 마지막으로 case 4의 경우에는  $E_3$ 만 계산하면 되고 이를 위해  $E_4$ 와  $E_5$ 를 미리 계산한다.



주어진 두 시퀀스의 길이가 각각  $n$ ,  $m$ , 각 서열의 인접하지 않은 mask된 영역의 개수가 각각  $v$ ,  $w$ , 각 서열에 있는 mask된 base의 총 개수가 각각  $T$ ,  $S$  인 경우, 본 연구에서 제시하는 알고리즘의 시간-복잡도는  $O((n-T)(m-S) + vm + wn)$ 이다.

#### 4. Deleting keys of B-trees in parallel

B-tree는 많은 키를 접근하고 갱신하는 데에 많이 사용되는 기본적인 자료구조이다. 본 연구에서는 B-tree가 수행하는 작업 중의 하나인 삭제 작업을 개선하여, EREW PRAM상에서 동작하는 병렬 알고리즘을 개발하였다. 개발한 알고리즘은 키의 개수를  $n$ 이라고 하고, 프로세서의 개수가  $k$ , B-tree의 각 node에 있는 최소한의 키의 개수를  $B$ 라고 할 때,  $O(B(\log k + \log_B n))$ 의 시간복잡도를 갖는다.

B-tree는 외부 저장소에 저장된, 매우 많은 키를 효율적으로 관리하기 위하여 개발된 balanced search tree 중의 하나이다. Root 노드를 제외한 모든 B-tree의 노드는 최소  $B$ 개, 최대  $2B-1$ 개의 키를 갖는다. 그리고  $n$ 개의 키를 가지고 있는 B-tree의 높이는  $O(\log_B n)$ 이다. B-tree는 찾기, 삽입, 삭제 등의 여러 작업을 수행할 수 있다.

B-tree에 대한 연구는 다양하게 이루어져 왔다. 그 중의 하나가 병렬 알고리즘이다. EREW PRAM상에서  $k$ 개의 프로세서를 이용하여 B-tree의 찾기, 삽입, 삭제 등을 수행하는 방법들이 연구되었으며, 그 시간복잡도는 다음과 같다.

- 찾기 :  $O(B(\log k + \log_B n))$
- 삽입 :  $O(B(\log k + \log_B n))$
- 삭제 :  $O(B^2(\log_B k + \log_B n))$

본 연구에서는 새로운 삭제 알고리즘을 개발하였다. 본 알고리즘은 같은 조건에서,  $O(B(\log k + \log_B n))$ 의 시간복잡도를 갖는다.

#### 5. 간염 바이러스 감염이 NFκB pathway에 끼치는 영향의 수학적 모델링

생명 현상을 시스템적으로 이해하기 위해서는 현상에 대한 수학적 모델링이 필수적이다. 여러 가지 수학적 모델 가운데 상미분 방정식(Ordinary Differential Equation) 모델은 여러 가지 생화학 반응을 모델링 하는데 널리 사용되고 있다.

본 연구에서는 NF-κB 신호전달 경로에 B형 간염 바이러스(HBV)의 감염이 미치는 영향을 ODE 모델로 모델링하고, MATLAB을 이용한 시뮬레이션을 통해 동역학적 특성을 분석하였다. 또한, 신호전달 경로 내 단백질 간의 연결 구조를 나타내는 그래프 표현과 이에 대한 ODE 수식 표현을 분리하여, 확장성과 재사용성, 수정의 용이성을 갖는 모델 명세 언어를 제안하였다.

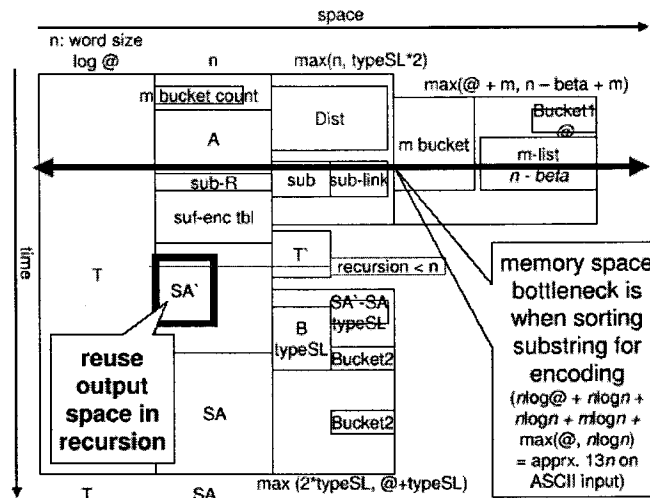
#### 6. Efficient Implementation of Suffix Array Construction Algorithms

Suffix array는 DNA 서열내에 존재하는 패턴을 검색(search) 작업을 수행하는 데에 자주 사용되는 자료구조이다. Suffix array를 생성하는 알고리즘으로는 Karkkainen-Sanders, Ko-aluru, Manber-Myers 알고리즘 등이 있다. Karkkainen-Sanders 알고리즘과 Ko-aluru 알고리즘은 스트링의 길이를  $n$ 이라고 할 때, 이론적으로  $O(n)$ 의 시간복잡도를 갖는다. 그러나 Manber-Myers의 알고리즘은 평균 수행시간(average time)은  $O(n)$ 이지만, 최악의 경우에는  $O(n \log n)$ 의 시간복잡도를 갖는다. 위의 세 알고리즘이 프로그램 수행 도중에 사용하는 메모리의 크기는  $O(n)$ 이다.

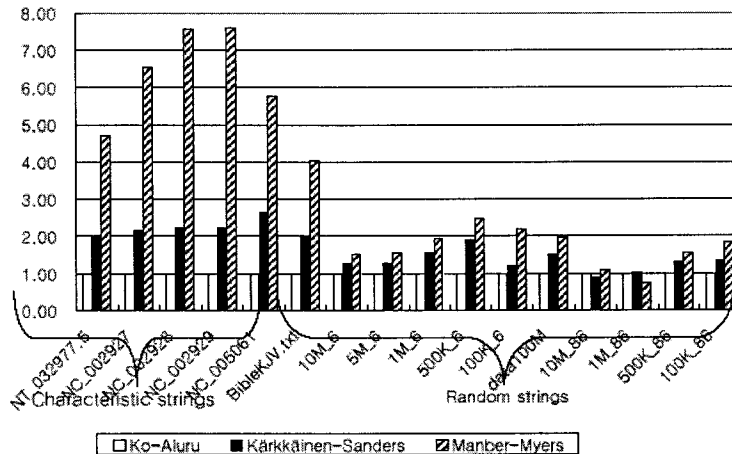
본 연구에서는 이들 알고리즘들을 최대한 빨리 동작하도록 구현하는 방법들을 개발하였다. 본

연구에서 사용한 방법들은 다음과 같다.

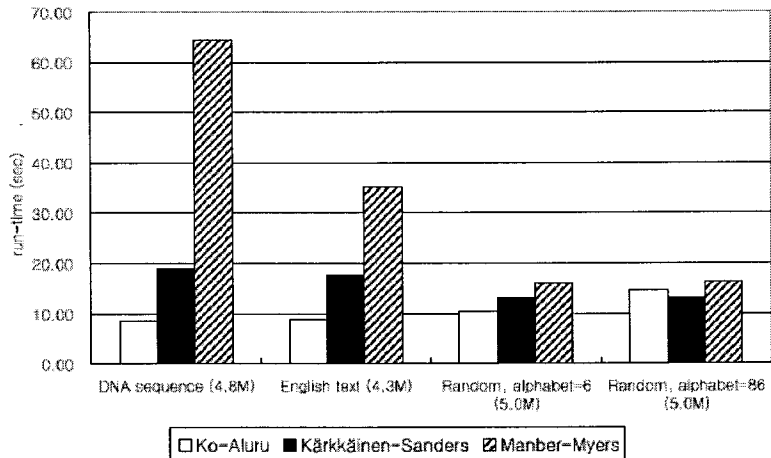
- 입력 알파벳의 크기를 프로그램의 수행에 맞도록 수정.
- 재귀적인 방법(recursion)이 수행될 때, 메모리를 재사용하도록 수정.
- Ko-Aluru 알고리즘의 수행시, 메모리 사용을 효율적으로 하도록 아래와 같은 메모리 맵을 구성하여 메모리를 관리.



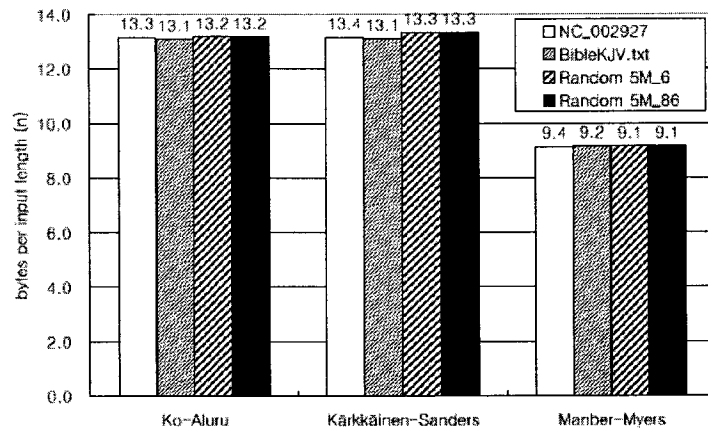
다양한 데이터에 대하여 이들 알고리즘을 수행하여 그 시간과 메모리 사용량을 측정하였다.



위의 그림은 다양한 스트링 데이터에 대하여 실험하여 수행 시간을 측정한 결과이다. Ko-Aluru 알고리즘의 결과를 1로 하여 수행시간을 비교할 수 있도록 하였다. Manber-Myers 알고리즘은 랜덤하게 생성한 서열을 이용하여 수행했을 때보다, DNA 서열을 이용하여 수행하였을 때 다른 두 알고리즘보다 훨씬 더 느린 결과를 보였다. 그리고 Ko-Aluru 알고리즘과 Kärkkäinen-Sanders 알고리즘을 비교해 볼때는, Ko-Aluru 알고리즘이 대부분의 경우에 좀 더 빠르게 suffix array를 생성함을 알 수 있었다.



위의 그래프는 서로 비슷한 크기를 갖는 4가지 데이터를 이용하여, 위의 세 알고리즘을 수행한 결과이다. Ko-Aluru 알고리즘과 Karkkainen-Sanders 알고리즘은 데이터의 종류에 크게 영향을 받지 않는 것을 확인할 수 있다. 그러나 Manber-Myers 알고리즘은 유난히 DNA 서열에서 더 느리게 수행되는 것을 확인할 수 있었다.



위의 그림은 세가지 알고리즘을 4가지 데이터에 대하여 수행하면서 메모리 사용량을 잰 것이다. Manber-Myers 알고리즘이 가장 메모리를 적게 사용하였고, 나머지 두 알고리즘은 서로 비슷한 정도의 메모리를 사용하였다.

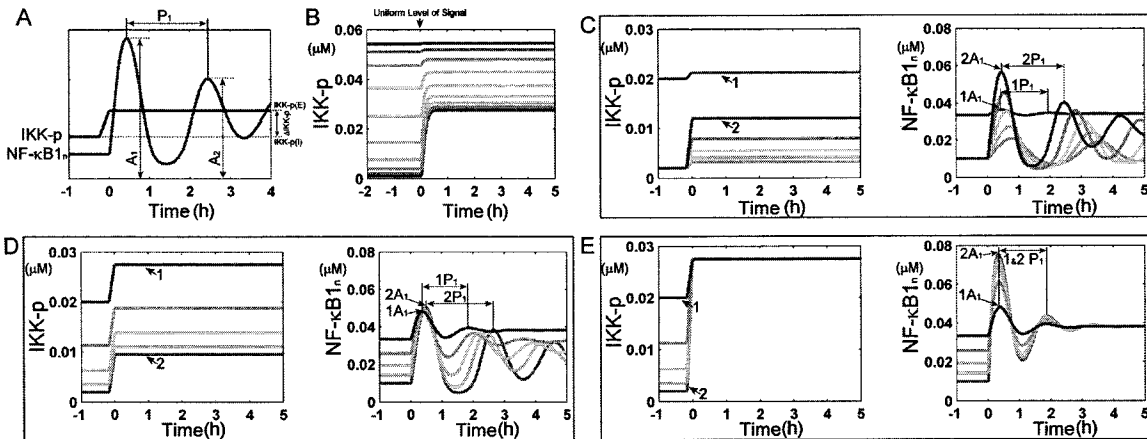
위의 결과들을 종합해 볼 때, Karkkainen-Sanders 알고리즘은 Ko-aluru 알고리즘과 유사한 메모리를 사용하고도 수행시간은 두배 정도를 사용하였으므로 Ko-Aluru 알고리즘이 더 좋은 알고리즘이라고 말할 수 있다. Manber-Myers는 메모리 사용에서 약간의 장점이 있었으나 수행시간이 상대적으로 너무 많이 걸렸다. 따라서 Ko-Aluru 알고리즘이 세 알고리즘 중에서 가장 좋은 알고리즘이라고 말할 수 있다.



## 7. Dynamic Patterns of the TNF- $\alpha$ Mediated NF- $\kappa$ B Activation Modulated by IKK for Hepatitis B Virus Infection: A Systems Biology Approach

NF- $\kappa$ B activation은 TNF- $\alpha$ 에 의한 apoptotic signal에 대해 세포가 anti-apoptotic 반응을 보이는데 핵심적인 역할을 하고 있다. 그러나 B형 간염 바이러스(HBV) 감염 세포에 대해서는 NF- $\kappa$ B activation이 높은 상태로 유지되고 있음에도, TNF- $\alpha$ 에 대해 오히려 apoptosis가 더 잘 일어나는 현상이 발견되고 있다.

본 연구에서는 *in silico* 시뮬레이션을 통하여 IKK-p의 초기 농도와 TNF- $\alpha$  신호에 의한 IKK-p 신호의 증가 정도( $\Delta$ IKK-p)가 NF- $\kappa$ B activation의 동적인 변화를 유발함으로써, 이러한 특징적인 apoptotic response이 나타나는데 중요한 역할을 하고 있음을 보였다.



IKK-p 초기 농도와 IKK-p(E),  $\Delta$ IKK-p에 따른 다양한 NF- $\kappa$ B activation

*in silico* 시뮬레이션을 통해, IKK-p의 초기 농도에 따라 동일한 upstream signal에 대해서도 다른  $\Delta$ IKK-p를 나타냄을 살펴볼 수 있다(그림 B). NF- $\kappa$ B activation은  $\Delta$ IKK-p에 따라 amplitude가 결정되는 특성을 나타내고 있다(그림 D). 반면에 upstream signal에 의해 결정되는 IKK-p의 평형 상태 level IKK-p(E)는 NF- $\kappa$ B activation의 주기에 영향을 주는 것으로 보인다(그림 E). 이러한 사실을 종합하여 볼 때, 초기 IKK-p의 level이 낮은 상태여도(그림 C, 2번) upstream signal로 인해 상대적으로 큰  $\Delta$ IKK-p를 갖게 되면 NF- $\kappa$ B activation이 상대적으로 높은 amplitude와 긴 주기를 갖는 반응을 보임을 살펴볼 수 있다(그림 C 1,2번 비교).

한편, *in vitro* 실험을 통해 실제로 HBV 감염 세포에서는 그렇지 않은 정상 세포보다 높은 IKK-p 초기 농도를 갖고 있음을 발견하였다. 또한 TNF- $\alpha$ 에 의한 NF- $\kappa$ B activation은 상대적으로 적은 주기와 낮은 amplitude를 보임을 확인하였다. 이러한 사실은 앞선 *in silico* 실험 결과에 부합하여, HBV 감염 세포에서 특징적인 apoptotic response가 IKK-p와 밀접한 관련이 있음을 시사하고 있다. 본 연구에서는 더 나아가 HBV 감염 세포에 대한 시뮬레이션을 통하여, *in vitro* NF- $\kappa$ B activation과 유사한 pattern을 갖는 *in silico* 결과를 보일 뿐만 아니라, *in vitro* 실험을 통해 HBV 감염 세포에서 anti-apoptotic gene의 활동이 감소하는 것을 발견하였다.

## 제 4 장 목표달성도 및 관련분야에의 기여도

년도	세부연구목표	가중치	평가의 착안점	자체평가 및 기여도	달성도
1 차 년 도	○ 데이터 베이스 및 인터페이스의 설계	100%	데이터 베이스/인터페이스 설계	<ol style="list-style-type: none"> <li>1. 기존의 genetic network model인 Boolean model, Linear model, Differential model 에 대해 비교, 조사하였다.</li> <li>2. Systems Biology 관련 데이터 표현의 표준인 SBML 언어에 대해 조사하였다.</li> <li>3. HBV, HCV의 boolean network model을 구성하고, 이를 SBML 언어로 기술하였다.</li> <li>4. EST 정보를 이용한 pathway 데이터베이스 검색 프로그램을 개발하였다.</li> </ol>	100%
2 차 년 도	○ 실험 결과 분석 및 단백질 네트워크 모델 개발	50%	논문 1 편	<ol style="list-style-type: none"> <li>1. 협동연구기관의 실험결과를 분석하여 HBV, HCV 관련 신호전달경로의 <i>in vitro</i> 모델을 구성하고 시뮬레이션 결과를 국내 학술회의에 발표하였다. (논문 1편)</li> <li>2. 개발된 ODE 모델을 이용하여 단백질 억제 효과 및 바이러스에 의한 영향 시뮬레이션을 진행하고 이를 <i>in vitro</i> 실험 결과와 비교, 검증하였다. (목표 초과 달성)</li> </ol> <p>(관련분야 기여도)</p> <p>- HBV, HCV 관련 신호전달경로 모델 확보함으로써 관련 연구 촉진 기대</p>	100%
	○ 데이터 베이스 및 인터페이스의 prototype 완성	50%	데이터 베이스/인터페이스의 prototype	<ol style="list-style-type: none"> <li>1. 그래프 모델과 ODE 모델의 인터페이스 역할을 하는 ODE 모델 명세 언어를 개발하였다.</li> <li>2. 데이터베이스에 활용되는 자료구조를 효율적으로 처리하는 알고리즘을 개발하였다.</li> </ol> <p>(관련분야 기여도)</p> <p>- 효율적인 알고리즘 개발함으로써 대용량 생물정보처리 데이터베이스에 활용 기대</p>	100%

(1) 연구개발의 최종목표

최종목표	내용
- HBV, HCV 네트워크 모델 개발 및 검증	- HBV, HCV에 대한 그래프 및 differential 모델 개발
- HBV, HCV 데이터베이스 및 인터페이스 완성	- HBV, HCV 관련 정보를 저장하는 데이터베이스와 이를 접근하는 인터페이스개발

(2) 계획대비 달성도

연차	변경 전 (선정연도 연구계획서)			변경 후			비고
	연구목표	연구내용	연구범위	연구목표	연구내용	연구범위	
2	- 개발된 단백질 네트워크 모델의 검증 - 데이터베이스 및 인터페이스 완성	- 1차년도에서 협동연구기관의 실험으로 얻어진 데이터의 분석 및 모델의 개발 - 데이터베이스/인터페이스의 prototype 완성		- HBV, HCV 네트워크 모델 개발 및 검증 - HBV, HCV 데이터베이스 및 인터페이스 완성	- HBV, HCV에 대한 그래프 및 differential 모델 개발 - HBV, HCV 관련 정보를 저장하는 데이터베이스와 이를 접근하는 인터페이스 개발		2차년도 협약시 승인

번호	세부연구목표	달성내용	달성도(%)
1	○ 데이터베이스 및 인터페이스 설계	<ol style="list-style-type: none"> <li>1. 기존의 genetic network model인 Boolean model, Linear model, Differential model 에 대해 비교, 조사하였다.</li> <li>2. Systems Biology 관련 데이터 표현의 표준인 SBML 언어에 대해 조사하였다.</li> <li>3. HBV, HCV의 boolean network model 을 구성하고, 이를 SBML 언어로 기술하였다.</li> <li>4. EST 정보를 이용한 pathway 데이터베이스 검색 프로그램을 개발하였다.</li> </ol>	100 %
2	○ 실험 결과 분석 및 단백질 네트워크 모델 개발	<ol style="list-style-type: none"> <li>1. 신호전달경로의 ODE 모델에 관한 국내외 연구 결과를 조사 분석하여 모델링 및 파라미터 추정, 모델의 동역학 분석, 파라미터 데이터베이스 활용 기법 등을 개발하였다.</li> <li>2. 제1, 2세부과제 연구기관에서 진행된 실험결과를 분석하여, HBV, HCV 관련 신호전달경로를 표현하는 ODE 모델을 4종 이상 개발하였다.</li> <li>3. 개발된 ODE 모델을 이용하여 신호전달경로의 동역학 특성을 분석하고, target molecules의 억제 효과 및 바이러스에 의한 신호전달경로의 변화 등을 시뮬레이션하고 <i>in vivo</i> 실험 결과를 비교하여 이를 검증하였다.</li> <li>4. 모델링 및 시뮬레이션 결과를 국내 학술대회에 발표하였다. (논문 1편)</li> </ol>	100%
3	○ 데이터베이스 및 인터페이스의 prototype 완성	<ol style="list-style-type: none"> <li>1. HBV, HCV 관련 신호전달경로를 표현하는 그래프 모델과 ODE 모델의 인터페이스의 역할을 하는 ODE 모델 명세 언어를 개발하였다.</li> <li>2. 대용량 생물정보 데이터베이스의 저장 및 정보처리에 활용되는 자료구조인 B-trees 및 Suffix Arrays에서 연산을 효율적으로 수행하는 알고리즘을 개발하였다.</li> </ol>	100%

## 제 5 장 연구개발결과의 활용계획

본 연구를 통해 구축된 간세포 신호전달경로 및 HBV/HCV 관련 *in silico* 모델은 향후 간염의 간질환 유도 작용에 대한 동역학적 메커니즘 규명 및 효과적 치료 방법 개발에 폭넓게 활용될 것으로 기대된다. 이를 위하여 wet/dry 연구 간 긴밀한 협력을 통한 모델 정립과 창의적인 *in silico* 모델 활용 및 분석 기법 개발이 필요하다.

추가 연구 과제	연구 내용
○ <i>in silico</i> 모델 구조의 정립	Literature 및 <i>in vitro/in vivo</i> 실험 결과 분석을 통하여, 구성 단백질 간의 상호 작용 관계 구조를 규명하고, 이를 통해 보다 정확한 <i>in silico</i> 모델을 정립한다.
○ <i>in silico</i> 상 주요 파라미터 선별 기법 개발	Sensitivity analysis 등을 공학적 기법을 활용하여 모델의 동역학적 변화에 가장 큰 영향을 주는 주요 파라미터를 선별하고, 이에 대한 <i>in vitro</i> 실험을 통해 모델을 보완한다.
○ <i>in silico</i> 모델의 동역학 특성 분석 기법 개발	비선형 시스템 분석에 사용되는 phase plane analysis, bifurcation analysis 등을 활용하여 모델이 가진 동역학적 특성을 발견한다.

위와 같은 연구를 통해 구성된 모델은 다음과 같은 연구에 응용 가능하다.

- 간염 관련 *in silico* 약물 target 선정
- 간염 치료 약물의 효과 및 부작용(side effect)에 대한 *in silico* 평가, 검증
- 간질환 환자의 임상 데이터 적용 및 *in silico* 예측
- 환자 임상 데이터에 기반한 최적의 치료 전략 시뮬레이션

## 제 6 장 연구개발과정에서 수집한 해외과학기술정보

### ◆ 해외 유용 웹사이트

▶ (HBV 관련) TNF $\alpha$ 신호전달경로 모델링 연구 그룹		
Wolkenhauer group, 독일	<a href="http://www.sbi.uni-rostock.de/default.htm">http://www.sbi.uni-rostock.de/default.htm</a>	
MPI, 독일	<a href="http://www.mpi-magdeburg.mpg.de/en/research/projects/1010/1022/1016">http://www.mpi-magdeburg.mpg.de/en/research/projects/1010/1022/1016</a>	
Universität Stuttgart, 독일	<a href="http://www.sysbio.de/projects/tnf/StudentProjects.shtml">http://www.sysbio.de/projects/tnf/StudentProjects.shtml</a>	
▶ 모델 repository		
DOQCS	<a href="http://doqcs.ncbs.res.in">http://doqcs.ncbs.res.in</a>	주요 모델 및 파라미터, reference 데이터베이스
SBML model repository	<a href="http://sbml.org/models/">http://sbml.org/models/</a>	각종 모델의 SBML 명세 파일 제공
SBML Simple ODE Solver oS	<a href="http://www.tbi.univie.ac.at/~raim/">http://www.tbi.univie.ac.at/~raim/</a>	ODE 시뮬레이터 및 MAPK 관련 모델 제공

### ◆ 연구 과정에서 수집한 참고 문헌

저자	저널	년도	제목
A. Kremling	-	2002	A benchmark for problems in reverse engineering, parameter estimation and model discrimination for cellular systems
Bentele et al.	JCB	2004	Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis
Bhalla & Iyengar	Science	2002	MAP Kinase Phosphatase As a Locus of Flexibility in a Mitogen-Activated Protein Kinase Signaling Network
G.Nelson et al.	J.Cell Science	2002	Multi-parameter analysis of the kinetics of NF-kB signalling and transcription in single living cells
Gardner et al.	Science	2003	Inferring Genetic Networks and Identifying Compound Mode of Action via Expression Profiling
Ghoah & Karin	Cell	2002	Missing Pieces in the NF-kB Puzzle
Hayen & Ghosh	Genes & Dev.	2004	Signaling to NF-kB
Hoffmann et al.	Science	2002	The I $\kappa$ B-NF-kB Signaling Module: Temporal Control and Selective Gene Activation
Huang & Ferrell	PNAS	1996	Ultrasensitivity in the mitogen-activated protein kinase cascade
Issacs et al.	PNAS	2003	Prediction and measurement of an autoregulator genetic module
Iyengar et al.	BioEssays	2002	Modeling of signaling networks
J.Ferrell	SKTE	2000	What Do Scaffold Proteins Really Do?
Kalir & Alon	Cell	2004	Using a Quantitative Blueprint to Reprogram the Dynamics of the Flagella Gene Network
Kholodenko	Eur. J. Biochem	2000	Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades
Kholodenko et al.	JBC	1999	Quantification of Short Term Signaling by the Epidermal Growth Factor Receptor
Levchenko et al.	PNAS	2000	Scaffold proteins may biphasically affect the levels of MAPK signaling and reduce its threshold properties
M.Santoro et al.	EMBO	2003	NF-kB and virus infection: who controls whom
Nils Bluthgen	Diploma thesis	2002	Dynamical Models of Signal Transduction and the Influence of Feedback Loop
Nils Bluthgen & Hanspeter Hezel	JTB	2003	How robust are switches in intracellular signaling cascades?
Shapiro et al.	draft	2000	Automatic model generation for signal transduction with applications to MAP-KINASE pathway
Sung & Simon	draft	draft	<i>in vitro</i> Simulation of Inhibitor Drug Effects on NF-kB Pathway Dynamics
Ting & En dy	Science	2002	Decoding NF-kB Signaling
Wiley et al.	TRENDS in C.Bio	2003	Computational modelling of the EGF-receptor system: a paradigm for systems biology
Cho et al.	GenomeResearch	2003	Investigations Into the Analysis and Modeling of the TNF $\alpha$ -Mediated NF-kB Signaling Pathway
Cho et al.	-	2003	Mathematical Modeling of the Influence of RKIP on the ERK Signaling Pathway
Kholodenko et al.	PNAS	2002	Untangling the wires: A strategy to trace functional interactions in signaling and genetic networks

## 제 7 장 참고문헌

- [1] Asthagiri, A.R. and D.A. Lauffenburger. 2001. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnol Prog* **17**: 227-239.
- [2] Bhalla, U.S. 2004. Signaling in small subcellular volumes. II. Stochastic and diffusion effects on synaptic network properties. *Biophys J* **87**: 745-753.
- [3] Cho, K.H., S.Y. Shin, W. Kolch, and O. Wolkenhauer. 2003a. Experimental design in systems biology based on parameter sensitivity analysis with Monte Carlo simulation: A case study for the TNF-alpha mediated NF-kappaB signal transduction pathway. *Simulation* **79**: 726-739.
- [4] Cho, K.H., S.Y. Shin, H.W. Lee, and O. Wolkenhauer. 2003b. Investigations into the analysis and modeling of the TNF alpha-mediated NF-kappa B-signaling pathway. *Genome Res* **13**: 2413-2422.
- [5] Hartwell, L.H., J.J. Hopfield, S. Leibler, and A.W. Murray. 1999. From molecular to modular cell biology. *Nature* **402**: C47-52.
- [6] Hoffmann, A., A. Levchenko, M.L. Scott, and D. Baltimore. 2002. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **298**: 1241-1245.
- [7] Huang, C.Y. and J.E. Ferrell, Jr. 1996. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci U S A* **93**: 10078-10083.
- [8] Ihekweba, A.E., D. Broomhead, R. Grimley, N. Benson, and D.B. Kell. 2004. Sensitivity analysis of parameters controlling oscillatory signalling in the NF-kappaB pathway: the roles of IKK and IkappaBalpha. *Systems Biology* **1**: 93-103.