



## 신약개발을 위한 그리드컴퓨팅 기반의 e-Science 환경 구축

### 목 차

1. 서 론
2. 사용자에게 제공되는 서비스
3. 포탈 시스템 구조
4. 활용 현황 또는 방안
5. 결론 및 향후 발전 방향

이진복 · 최재영 · 노경태  
(송실대학교 · 연세대학교)

### <요 약>

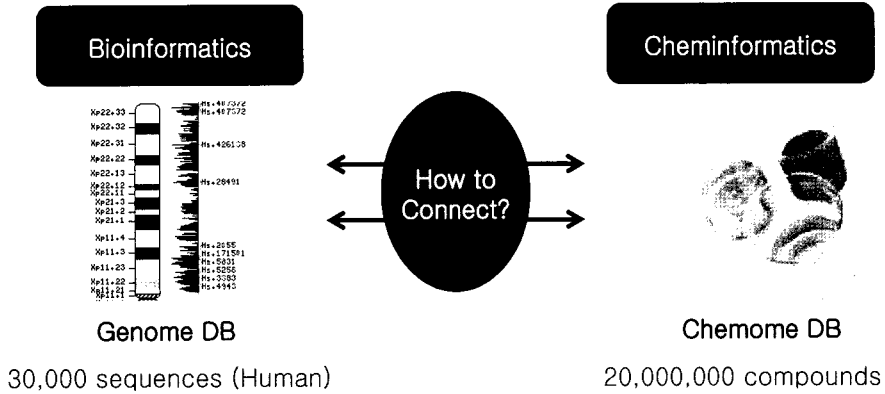
신약 개발의 효율상의 한계를 적극적으로 극복하기 위해서는, 인간 생명 현상에 관계되는 전체 유전자 정보와 이에 수반되는 단백질, 화합물 정보들을 그 의미론 정보와 동시에 전산화하고, 전산 모사와 계산 작업을 통해 정보를 역으로 재가공하는 작업이 필요하다. 또한 재사용성이 높은 대량의 전산 자원이 필요하며, 응용상의 실제 연구에 직결되는 유연하고 편리한 사용자 환경이 필수적이다. 따라서 본 연구에서는 생명과학과 화학 정보를 대량으로 관리하기 위한 정보 체계와 전산 자원 관리 및 작업 수행을 그리드컴퓨팅 환경에서 진행하였으며, 이를 통해 e-Science 포탈의 구축과 함께 현장 서비스를 통해 현실적으로 생명 연구의 혁신을 도모할 수 있는 결과를 제공할 수 있다.

### 1. 서 론

신약개발에 있어서 정보 기술의 활용 양상은 이미 완숙화 단계에 접어들어 있고, 신약 개발 기술의 첨단을 달리는 한 분야로 정착되어 있다.

전문 과학 기술 분야로서는 이례적으로 RCSB [1] 웹사이트는 현존하는 모든 웹페이지 중에 접속자 5,000위권에 속하고 있으며[2], 대표적인 소프트웨어 기술 기업인 마이크로소프트에서도 최근 본격적으로 신약개발에 관한 전망 및 지원을 대폭 강화하고 있다[3].

실제 신약 후보 물질은 대부분 화합물인데 유전자 정보와 화합물 정보의 관계를 그림 1에서 보이는 것과 같이 어떻게 규명하느냐가 중요하다. 이것과 관련된 이론적 및 실험적 방법이 필요하며 많은 연구가 진행되었지만, 기존의 신약개발 연구는 유전체나 단백질 등의 단위 분야별 연구가 주로 이루어졌다. 이러한 연구를 통해 인간 유전체 사업이 끝나고 현재 각 유전자의 기능에 대한 연구가 진행되어 질병과 유전자와의 연관성이 알려지고 있으나, 이를 통합적으로 연계시키는 시스템이 부족하였다. 따라서 세계적으로 유명한 Bioinformatics 및 Genomics 회사들이 질병관련 유전자 기능을 얻고도 유전자 기능을 조절할 수 있는 화합물 발굴에 어려움을 겪고 있다.



(그림 1) HG2C(Human Genome to Chemicals)

인간 유전자 정보는 알려져 있으므로 이를 이용하여 단백질 구조를 결정할 수 있고, 단백질 구조는 이들과 상호작용하는 화합물을 찾는 데 사용될 수 있으며, 이 화합물은 다시 신약이 가져야 할 특성을 기준으로 분석될 수 있다. 이 상관관계를 미리 데이터베이스로 구축해 놓으면 새로운 질병의 원인이 되는 유전자가 알려졌을 경우 바로 신약 후보 물질을 선택하여 실험을 할 수 있으므로 신약탐색 시간과 비용을 획기적으로 줄일 수 있다.

세계 시장에서 상대적으로 국내 신약 개발 연구 수준이 그 역사가 일천한 상황을 고려한다면, 후발 연구자의 입장에서 연구를 수행하는 국내 연구진에게 이러한 새로운 기술 양상에 적극적으로 주목할 필요가 있다. 더욱이 국내 정보 기술의 독특한 학문적 및 산업적 경쟁력을 고려한다면 신약개발 연구에 있어서 전산 처리의 소프트웨어와 하드웨어 지원을 적극적으로 활용하고 강화하는 관점은 비교우위와 경쟁우위를 동시에 만족하는 좋은 방향이다.

신약 개발 연구에서 정보 기술을 활용하는 것은 가장 대표적인 학제간 교류에 의한 연구 분야이다. 따라서 다양한 이질적인 분야의 기술과 정보 및 자료들이 통합되어 해석될 때에 지식으로서의 의미가 강화되는 영역이다. 그러나 기존에

연구의 바탕이 되는 데이터베이스와 정보처리 체계는 학제간 교류의 부족과 정보 처리량의 부족으로 특정 형태의 자료에 분화된 고립형 정보로 관리되어 온 한계가 있었으며, 그에 따라 학제간의 유관 정보 탐색 및 그에 따른 지식 창출은 전적으로 연구자의 직관과 단순 검색 기능에 의존하는 것이 현실이었다.

그러나 최근 분산처리기술이 발전되고 보다 확대된 대용량 정보 처리가 가능해지면서, 인간 유전체 전체를 처리하는 대규모 전산 작업이 실용화되고 있다. 또 다양한 분야에서 시도되고 있는 수치적인 자료 해석 및 전수조사 처리가 신약 개발 데이터베이스에 대해서도 연구되고 있는 실정이다. 이러한 맥락에서 e-Science 인프라를 확충하고 활용하는 연구 방향에서 대규모 전산 모사를 이질적인 생명과 화학 정보를 통합하고 이러한 정보를 공개 및 가공 하는 시도는 문제해법의 용이성과 결과 유의미함에 있어서 신약 연구 현실 문제에 잘 부합된다.

본 연구에서는 이와 같은 필요성에 부합하는 소프트웨어를 개발하고, 실제 시스템에 활용할 수 있도록 구축되었다. 또한 이 시스템을 통해 실제 인간 유전체 전체에 대해 계산 가능한 정보를 전량 계산하여, 이를 데이터베이스화하고 관련 단백질 화합물 관련 정보를 모두 전산화하는

방대한 연구 작업을 수행해 완료하였다. 그리고 HG2C[4] 포털을 제공하여 정보가 필요한 사용자는 웹으로 접근할 수 있다.

## 2. 사용자에게 제공되는 서비스

HG2C 연구는 인간 생체의 분자생물학적 및 화학적 묘사를 위해 필요한 모든 객체들과 그 연관 관계를 저장하는 표준화된 데이터베이스를 구축하고, 그 자료를 공공에게 공개하여 검색, 조회하는 것을 목표로 진행되었다. 따라서 인간 유전체와 그 발현 단백질, 단백질에 활성이 있는 화학물질의 정보들 모두를 전산 모사를 통해 관계를 예측해왔고, 그 결과를 공공에 공개하고 있다.

구체적으로 HG2C 연구의 진행과정에서 고성능 분산처리 작업을 통해, 밝혀진 인간 유전체 46,162개 전량에 대해 모두 탐색작업을 완료하였다. 우리는 그동안 HG2C 연구를 진행하면서 병렬처리를 통해 대규모 계산 작업을 수행하였고, 2기의 파일 서버를 통해 인간 유전체 전체를 탐색한 전산 모사의 계산 결과를 데이터베이스화하여 이를 웹을 통해 저장하고 공유하였다. 이로써 현시점에서 인간 유전체에 대한 이론 연구가 목표로 하는 가능한 모든 조합에 대한 계산 결과를 전량 확보하였다.

이후 저장된 계산 결과를 작년 2006년 11월부터 현재까지 일반 공공에 대하여 완전히 공개하여 실제 현장 운영에 돌입하였다. 현재 만1년째를 맞이하고 있는 HG2C 실제 서비스는 무제한의 전송과 조회 작업을 허용하고 있다. 따라서 HG2C 연구는 2006년 11월 공개 시점부터, 46,162개 유전자에서 전산 모사 계산을 통해 도출한 8,487개의 단백질 3차원 구조를 모두 공개하여 자유롭게 조회할 수 있도록 하고 있으며, 모든 단백질에 활성을 갖는 3,136 군집의 화학물질 정보 역시 모두 조회할 수 있도록 제공하여, 실제 연구 현장에 도움이 되고자 하였다.

HG2C 연구를 통해 계산되어 데이터베이스화

된 막대한 자료를 사용자에게 원활히 제공하기 위해, HG2C 서비스는 46,162개 유전자, 8,487개 단백질, 3,136 군집 화학물질간의 관계를 웹을 통해 탐색할 수 있도록 하였으며, 단백질과 화학물질의 3차원 구조 역시 웹에 통합된 인터페이스를 통해 그 3차원 정보를 컴퓨터 그래픽스로 확인할 수 있도록 했다. 그리하여 2006년 11월에 공개된 시점에서, 이론적인 전산 모사로 예측할 수 있는 인간 유전체에 대해 단백질 및 화합물에 이르는 모든 유관정보 전량을 전 세계 어느 사용자라도 자유롭게 활용할 수 있도록 하고 있다.

이러한 정보를 통해 사용자는 모든 인간의 유전체 중에서 임의의 유전자가 동질성(homology) 기준으로 예측 가능한 단백질 구조를 가질 수 있는지 웹을 통해 확인할 수 있으며, 그러한 단백질의 구조와 그 단백질에 활성을 갖는 화학물질의 구조까지 바로 웹을 통해 손쉽게 조회할 수 있다. 따라서 이론상 예측 가능한 인간 유전체상의 모든 질병에 대해 그에 따른 약성 및 독성을 지니는 것으로 예측되는 모든 화학물질들을 자유롭게 확인할 수 있다.

한편, 사용자는 원격지에서 실시간으로 전산 모사 계산을 재수행하여 갱신된 정보를 얻을 수 있도록, 이에 대한 소프트웨어와 하드웨어 자원 역시 제공하고 있다. HG2C 연구 과정에서 대량 계산을 통해 얻은 계산 결과들은 고속 계산을 위해 계산 방식과 탐색 범위를 일정하게 통일한 자료들이다. 그러므로 특정한 유전체에 대해 특이성이 있는 계산 방식을 도입한다거나, 예측 화합물의 특성을 사용자 임의로 조작하고 싶은 경우, 혹은 특별히 연구상의 중요성을 인식하여, 보다 높은 정확도의 전산 모사 계산을 수행하고 싶은 경우 등에 대해서는 조건을 변경하여 새로운 계산을 진행할 필요가 있다.

그러므로 사용자가 임의의 유전자를 이용해 단백질 3차원 구조를 예측하고, 단백질 3차원 구조를 바탕으로 활성이 있는 화학물질을 찾아내

는 계산을 진행하기 위한 소프트웨어를 개발하였다. 이러한 소프트웨어는 현재 보유하고 있는 HG2C 서버 하드웨어에 할당되어 사용자의 요청에 따라 가동되며, 그 결과는 가상 파일 서버에 보관되어 계산을 지시한 사용자가 조회할 수 있도록 하였다. 이렇게 실시간으로 사용자가 계산을 재수행하는 작업은 2006년 11월부터 가입한 회원에 대해 허용하고 있으며, 모든 작업 조작성 마찬가지로 웹을 통해 가능하도록 구현하였다.

현재 사용자의 실시간 계산을 지원하기 위하여, (주)이쿠스팜의 in silico 신약개발 소프트웨어 제품인 IDPro™와 IDPharmo™를 설치하여 원격 사용 환경으로 연동시켰다. 이를 통해 원격지 사용자는 웹 접속을 통해 내부 서버에서 가동되는 해당 소프트웨어와 소프트웨어에 연동된 기타 라이브러리 자원을 활용하여 갱신된 계산 결과를 얻을 수 있다.

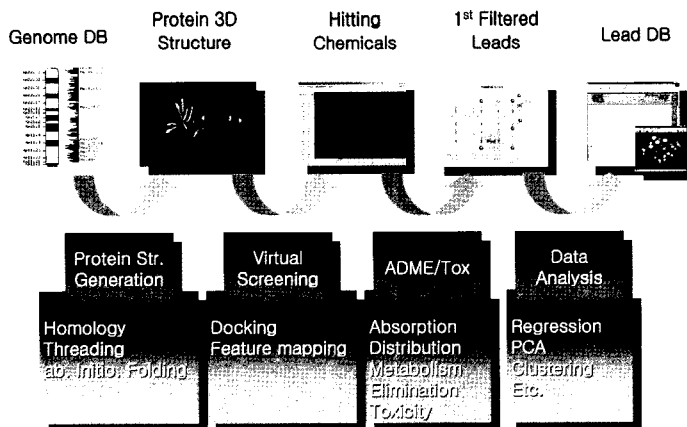
또한 2006년 11월부터 HG2C 연구의 주요 기능은 SOAP 기반의 공개 웹서비스로도 제공되고 있다. 기타 e-Science 연구와 확장 연결에 대한 요청을 수용하여, HG2C 연구는 그 실시간 전산 모사 계산 기능 중 단백질-화합물 상호작용을 예측하는 단계를 웹서비스로 제공하고 있다.

웹서비스는 세 종목으로 구분되어 단백질 3차원 구조를 분석하여 활성부위를 예측하는 기능,

지정된 활성부위에서 화합물과 상호작용을 할 수 있는 화학적 작용기를 예측하는 기능, 지정된 화학적 작용기들 간의 상호작용을 고려하여 단백질-화합물 상호작용의 결과를 예측하는 기능이 제공되고 있다. 실제 현장 운영을 개시하여 공개적으로 서비스를 진행하는 과정에서 이처럼 내부 소프트웨어 작동에 직접 연결된 RPC를 제공하는 것은 서비스 안정성과 보안에 매우 중대한 사안이다. 따라서 현재 HG2C 연구에서 제공하는 웹서비스는 그 보안상의 문제를 예방하기 위하여 최대한 예민하게 관리하며 가동 중에 있다.

HG2C 연구는 그 초기에 무엇보다 이질적인 형태의 정보와 그 정보들의 연관관계를 저장하는 표준화된 데이터베이스가 설계하는 것을 우선으로 작업하였다. 그러므로 HG2C 연구가 그 결과물로 생성하고 갱신하는 대규모의 방대한 자료는 2006년에 공개 운영을 앞두고 표준 SQL 질의를 수용하고 호환성있는 데이터베이스 형태로 구현되었다.

공개된 데이터베이스 구조는 자유로운 검색과 조회를 유전자를 제 1 기준으로 하는 것으로, 유전체 의미론 정보, 유관 문헌 정보, 단백질 구조 정보, 화합물 구조 정보 등의 다양한 정보를 각각 그 참조 인식자로 저장하고 표현하였다. 이러한 데이터베이스는 기본적으로 공개정책에 따라

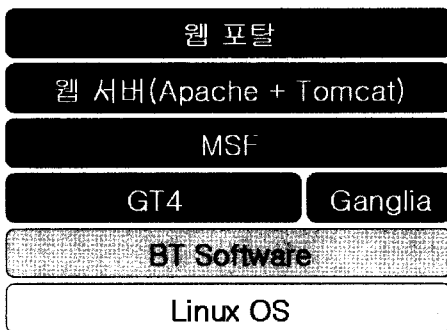


(그림 2) HG2C 시스템의 진행절차

운영되어 표준 SQL 질의와 mySQL 데이터베이스 접속이 가능한 모든 클라이언트에서 접속 가능하다. 이로써 2006년 공개 운영 이후부터 현재까지 신청 연구자에 대해 SQL 질의 수준에서 데이터베이스에 대해 직접적인 접근과 조작을 제공하고 있다. 인간의 유전자 정보 데이터베이스로부터 신약개발에 이용할 수 있는 정보를 얻기 위한 HG2C 시스템의 처리 절차를 (그림 2)에서 보여주고 있다.

### 3. 포탈 시스템 구조

HG2C 과정을 그리드 환경으로 통합하고 이를 사용자에게 제공하기 위한 포탈이 필요하며, (그림 3)과 같은 웹포탈 형태로 시스템을 구축하였다.

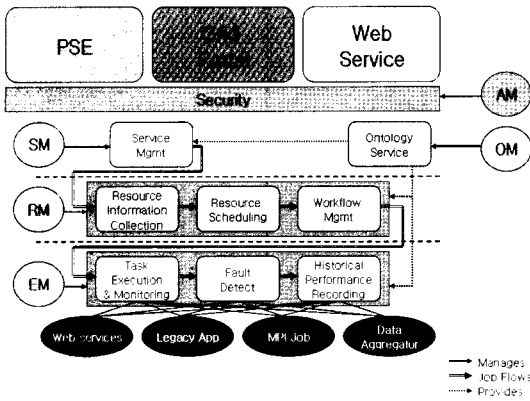


(그림 3) HG2C 시스템의 진행절차

HG2C를 위한 시스템은 BT 소프트웨어가 설치되어 있는 일반적인 Linux OS를 사용하며, 그리드컴퓨팅을 위한 Globus Toolkit[5]이 설치되어 있다. 그리고 MSF (Meta Service Framework)[6]는 그리드 환경에서 작업을 생성하고 실행하기 위해 이를 제출하는 과정을 통합하는 미들웨어이다. 이러한 다수의 작업으로 부터의 결과물을 사용하기 위해 사용자는 웹을 통하여 쉽게 접근할 수 있으며, 웹서비스로 이용하기 위해서 표준 WSDL[7]을 제공하고 있다.

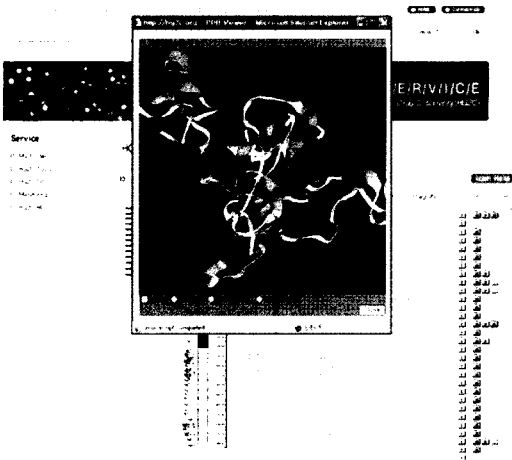
HG2C 시스템에서 복잡하고 작업량이 많은 BT 소프트웨어를 효율적으로 실행하고 관리하기 위한 미들웨어가 필요하게 되었다. 그리드 자원을 관리하는 계층부터 사용자에게 워크플로우 인터페이스를 제공하는 계층을 통합하여 그리드 포탈을 구축할 수 있는 미들웨어인 MSF를 개발하게 되었으며, (그림 4)와 같은 구조로 구성되어 있다. MSF는 서비스 지향적인 워크플로우를 기반으로 워크플로우의 재사용성을 증대시키며, 기능에 따라 다섯 개의 에이전트로 구성되어 요 구되는 그리드 환경에 따라 간단한 코드 수정 등으로 최적화할 수 있다.

MSF를 구성하는 다섯 개의 에이전트는 모두 자바로 구현되어 있으며, 각 에이전트의 역할과 기능을 살펴보면 다음과 같다. AM(Access Manager)은 전체 시스템의 인증을 담당하며, OM(Ontology Manager)은 워크플로우의 재사용성과 편의성을 제공하기 위한 계층적인 메타 서비스를 기술하는 XML 파일들을 저장하고 관리한다[8]. SM(Service Manager)은 사용자가 OM에 접속해서 서비스를 검색하고 호출하면 해당 서비스의 메타서비스 정보를 이용해서 워크플로우를 생성하는 기능을 가지고 있다. RM(Resource Manager)은 생성된 워크플로우를 해석하여 태스크로 분할하고 미리 수집된 자원 정보를 이용해서 스케줄링하며, EM(Execution Manager)에 할당하게 된다. EM에서는 할당받은 작업을 실행, 모니터링 등의 제어 기능을 가지고 있다. 이와 같이 MSF를 다섯 개의 에이전트로 구성하면 다양하게 변화되는 그리드 환경에 최적화된 미들웨어를 구축할 수 있게 되며, 이로써 적응력이 뛰어나고 높은 유연성을 제공하게 된다. 또한 새로운 워크플로우를 작성하거나 기존 워크플로우의 정보를 수정할 때에는 함께 개발된 GUI 환경의 워크플로우 편집기를 사용하면 직관적이고 편리하게 작업할 수 있다[9].



(그림 4) MSF 시스템의 구조

(그림 5)와 같이 사용자는 웹을 이용하여 HG2C 포털에 접근할 수 있으며, BLASTTM 및 IDProTM 등의 BT 소프트웨어를 실행하여 결과물을 얻을 수 있다.



(그림 5) HG2C 포털

#### 4. 활용 현황 또는 방안

현재 HG2C 연구는 이론상 접근 가능한 인간 유전체 전체에 대한 전산 모사 계산 결과를 무제한으로 공개하고 있다. 무엇보다 신약개발 과정의 시작 단계에서 인간 유전체상에서 생각할 수 있는 모든 질병 전체에 대하여 이론상 가능한 약

물 후보를 제시하여 준다. HG2C 연구는 이러한 작업을 복잡한 하드웨어와 소프트웨어의 고수준의 조합으로 수행하는 것이 아니라, 단순한 웹을 통해 누구나 접근할 수 있도록 구현되어 있다. 이것은 고수준의 전산 모사 계산을 이해하고 조작하는데 능숙하지 못한, 신약 개발 분야 실험 현장 연구자들의 제반에 크게 유리한 형태이다.

하지만 현재 이론 화학에 있어서 전산 모사 계산상의 예측 한계는 자명하며, 성공적인 예측이라 할지라도 생리적인 시간효과나 신진대사 상의 간접변화의 영향은 고려되지 않고 있다. HG2C 연구를 통해 구축한 공개된 데이터베이스 자료는 그러한 한계 내에서 얻은 결과이므로, 이렇게 제시된 약물 후보가 당장 투약 가능한 극단적인 정확도를 갖는 것은 역시 아니다.

제공되는 약물 후보의 목록은 구체적인 작업 대상과 연구 실체가 막연한 신약개발 시작 단계에서 연구자에게 구체적인 화학적 탐색의 관점을 제안할 수 있다. 또한 수백만 종에 이르는 대상 화합물 중에서 1차적으로 고려할 수 있는 화합물의 범위를 한정해주는 의미에서, 신약개발의 시간을 비약적으로 축소시킬 수 있을 것으로 기대되므로 구축된 결과 데이터베이스의 조회와 열람은 활발히 진행되고 있다.

나아가 모든 인간 유전체-화합물의 전산화된 상호작용 지도를 구축하여 열람할 수 있으므로, 이를 근거로 하여 특정 화학물질의 독성 연구 및 부작용 연구에 활용할 수 있으며, 화합물 기준으로 단백질/유전자를 역분류, 군집화하여 단백질의 유사성 및 계통 연구에 적용할 수 있다. 또한 유전자에 대해 진화론적인 연구를 수행하는데 대해 통계적인 자료로 구축된 데이터베이스 정보를 활용할 수도 있다. 이상과 같은 다양한 활용과 분석 연구 작업을 위해, 현재 서비스는 최대한 공개 정책을 위주로 운영되고 있으며, 이러한 공개 운영을 토대로 이질적인 대규모 정보를 저장하고 조회하는 표준 데이터베이스 구조안을 표준화 작

업에서의 성과도 기대할 수 있을 것이다.

한편 e-Science 분야 내부의 연구에 있어서 본 과제는 전 세계의 공공에 대하여 실제로 공개를 게시하고, 현장에서 서비스를 장기간 운영해 본 선도적인 경우로 주목할 만하다. 이를 통해 실제 현장 연구가 아니면 얻기 힘든 서버 운영과 미들웨어 안정성에 대한 귀중한 정보 자료를 수집할 수 있었으며, 특히 e-Science 응용의 실제에서 반드시 필요한 보안상의 다양한 실제 환경에 대한 경험을 갖게 된 것은 어느 응용 분야보다 앞서는 성과로 볼 수 있다.

### 5. 결론 및 향후 발전 방향

본 연구는 국내에서 e-Science의 활용 연구와 그에 실제로 요구된 미들웨어 개발을 동시에 진행하여 실제 공개 운영에 들어간 최초의 서비스 중 하나이다. 그러므로 안정적인 서비스를 지속적으로 운영하여 e-Science 연구의 모범적인 형태를 제안할 수 있다. 이러한 서비스 운영을 통해 기존의 신약개발연구에서는 기대하는 것이 불가능했던 미지의 신약에 대한 후보물질 정보를 전산모사를 통해 정보검색단계에서부터 제시할 수 있으며, 또한 대규모의 이질적인 전산 자원과 대상 정보를 실제 서비스에서 활용하기 위해, 연구에 사용하여 검증된 미들웨어 및 사용자 편의 소프트웨어를 제공할 수 있다.

인간 유전체 정보에 포함되는 모든 신약개발에 필요한 정보를 가능한 전산모사로 미리 제안함으로써, 실험 정보가 부족한 도전적인 신약개발 연구에서 최초의 연구 방향성을 제시할 수 있게 되었다. 또한 이질적인 정보를 통합 관리하는 데이터베이스 및 정보 기술 체계를 실용화함으로써 관련 표준안을 선점할 수 있으며, 상용 워크플로우 엔진 및 상용 그리드 미들웨어를 개발하는 소프트웨어로 대체를 기대할 수 있다. 더불어 웹사이트 접속자 5만위권 수준의 외국 접속 DB를 국내의 통합 DB로 대체할 기반이 되며,

e-Science의 응용 연구와 소프트웨어 개발 연구를 동시에 진행하여 현장 연구자에게 서비스화한 성공적인 사례로서 그 운용 경험을 확보할 수 있다.

### 참고문헌

- [1] RCSB, Research Collaboratory for Structural Bioinformatics, <http://www.rcsb.org>
- [2] <http://www.alex.com>
- [3] <http://www.microsoft.com/industry/healthcare/lifesciences/solutions/lisdrugdiscovery.msp>
- [4] HG2C, Human Genome to Chemicals for Drug Discovery, <http://www.hg2c.org>
- [5] Globus Toolkit, <http://www.globus.org/toolkit>
- [6] 이상근, 최재영, 조금원, "HG2C 연구를 위한 이벤트 주도 구조 기반의 모듈화된 그리드 미들웨어", 한국정보과학회 HPC연구회 학술발표회 논문집, 제17권 제2호, pp.57-65, 2006.
- [7] WSDL, Web Services Description Language, <http://www.w3.org/TR/wsdl>
- [8] 이진복, 이상근, 최재영, 변옥환, "그리드 환경에서 메타서비스 기반의 워크플로우 시스템", 한국정보과학회, 정보과학회논문지, 시스템 및 이론, 제34권 제9호, pp.385-394, 2007.
- [9] 이진복, 이상근, 최재영, "그리드 환경에서 3개의 계층으로 분리된 MSF를 위한 사용자 중심 워크플로우 편집기의 구현", 한국정보과학회 HPC연구회 제18권 제1호, pp.65-72, 2007.

### 저자약력



**이진복**

2007년 숭실대학교 컴퓨터학부 졸업 (학사)  
2007년~현재 숭실대학교 컴퓨터학과 석사과정  
관심분야 : 분산처리, 그리드컴퓨팅, 가상화



**노경태**

1978년 연세대학교 화학과 졸업  
1980년 한국과학기술원 화학과 (석사)  
1983년 한국과학기술원 화학과 (박사)  
1983년~2003년 숭실대학교 화학과/생명정보학과 교수  
2003년~현재 연세대학교 생명공학과 교수  
1997년~2003년 숭실대학교 분자설계연구소 소장  
2002년~2003년 사단법인 분자설계연구소 소장  
관심분야 : 컴퓨터를 이용한 신약개발, 생물정



**최재영**

1984년 서울대학교 제어계측공학과 (학사)  
1986년 미국 남가주대학교 전기공학과 (석사)  
1991년 미국 코넬대학교 전공학과 (박사)  
1992년~1994년 미국 국립 오크리지연구소 연구원  
1994년~1995년 미국 테네시 주립대학교 연구교수  
2001년~2002년 미국 국립 슈퍼컴퓨팅 응용센터 (NCSA) 초빙연구원  
1995년~현재 숭실대학교 컴퓨터학부 교수  
관심분야 : 고성능컴퓨팅(HPC), 병렬/분산처리, 시스템 소프트웨어