

Chiral-MGrid: 그리드 기반 광학이성질체 식별 서비스

강미란* 정갑주* 이종현* 최영진** 정선호**

건국대학교

{meelan*, jeongk*, lejohy*, ototo**, shjung**}@konkuk.ac.kr

Chiral-MGrid: Grid based Chiral Discrimination Service

Miran Kang* Karpjoo Jeong* Jonghyun Lee* Youngjin Choi** Seunho Jung**

Department of Technology of Fusion, Konkuk University*

Department of Microbial Engineering, Konkuk University**

요 약

현재 IT분야의 끊임없는 발전과 더불어 BT/NT 분야와의 협업이 활발히 진행되고 있는 가운데, 광학 이성질체에 대한 연구는 그 중요성이 점점 증가되고 있다. 광학 이성질체는 생체 내에서 작용할 경우 성질, 기능, 효과에서 정반대 결과를 일으키기에 제약 산업 등 응용분야에서는 사용을 위해 효과적으로 식별하여 분리할 필요성이 존재한다. 이런 실험은 단계가 복잡하고 결과를 분석하고 시각화하기 어렵다. 이를 해결하기 위해 본 논문에서는 광학 이성질체의 분리연구를 위한 그리드 기반의 Chiral-MGrid 시스템을 제공한다. Chiral-MGrid 시스템은 분리연구 관련 실험을 그리드 기반에서 one set operation으로 수행할 수 있는 서비스를 제공한다.

1. 서 론

그리드란 지리학적으로 분산되어 있는 컴퓨팅 자원을 네트워크로 상호 연동하여 조직과 지역에 관계없이 사용할 수 있는 환경 또는 서비스를 말한다. 그리드 플랫폼은 다양한 소프트웨어, 하드웨어로 구성되어 있기에 이런 그리드 환경에서 소프트웨어를 개발하고 적용하는데 어려움이 존재한다. 그리드 포털은 연구자들에게 웹 기반의 인터페이스를 통해 그리드 서비스를 제공함으로써 보다 편리한 작업 환경을 제공한다. Chiral-MGrid 시스템은 일반 컴퓨터 자원으로는 연산이 어려운 광학 이성질체 분리 모사실험을 그리드 기반의 웹 포털을 통해 수행하고, 그에 대한 결과를 분석, 시각화하여 연구자가 쉽게 확인할 수 있도록 설계, 구현되었다.

본 논문에서는 연구자가 분자구조 파일만을 정의하고, 이를 연구자 웹 인터페이스를 통해 제출하면, 시뮬레이션/분석/시각화 단계를 자동화하여 그에 대한 요약결과를 보여주는 Chiral-MGrid 시스템 기술한다. 2절에서는 Chiral-MGrid 시스템의 설계 및 구현의 배경인 광학 이성질체 및 MGrid 시스템에 대해 설명하고, 3절에서는 Chiral-MGrid 시스템 설계 및 구조에 대해, 4절에서는 Chiral-MGrid 시스템의 구현에 대해 기술한다. 마지막으로, 5절에서는 결론 및 향후 계획에 대해 기술한다.

2. 관련연구

2.1 광학 이성질체 분리

광학 이성질체(Chiral)란 원자조성이나 결합방식이 똑 같지만 원자들 간의 입체 배치만 다른 관계에 있는 일련

의 분자 화합물을 의미한다.[1] 광학 이성질체는 물리화학적 성상이 동일하여 화학반응성에 차이가 없지만 생체 내에서 작용할 경우 광학 이성질체 간에 차이가 나타나며, 심할 경우 광학이성체의 한쪽은 치료약으로 쓰일 수 있고 다른 한쪽은 독성을 나타내는 경우가 발생하기도 한다.[2][3]

광학이성질체의 쌍을 각각 다르게 인식할 수 있는 것을 Chiral recognition이라고 하며 이를 통해 각 광학이성질체를 분리할 수 있게 되면 Chiral Discrimination되었다고 한다. 이때, 연구자는 광학이성질체 분리를 위하여 다양한 종류의 분리물질들을 첨가하는 일련의 유기화학 실험을 진행하여 연구대상이 되는 광학이성질체를 분리할 수 있는 조건을 찾아야 한다. 이 중에서 Cyclodextrin과 같은 환형 탄수화물을 이용하여 광학이성질체와 포집 복합체(Inclusion Complex)를 형성시키고 이때의 결합력 차이를 이용하여 광학 이성질체를 분리하는 방식이 널리 쓰이고 있다. 이러한 실험들은 고가의 시약을 구입하여 소모해야 하며 원하는 수준의 분리능을 찾을 때까지 다양하게 실험조건을 바꾸어가며 많은 시간을 투자해야 하기 때문에 사전에 컴퓨터 시뮬레이션이 필수적이다. 그러나, 이러한 장점에도 불구하고, 시뮬레이션/분석/시각화의 단계가 너무 복잡하여 실제 연구자에게 많이 사용되지 못하고 있다. Chiral-MGrid 시스템은 광학이성질체 식별 실험에 대한 간단하고, 쉬운 모의실험 환경을 제공한다. [4][5][6][7]

2.2 MGrid(Molecular Simulation Grid)

MGrid 시스템[7]은 분자 그리드 시스템 아키텍처를 기반으로 바이오/나노 연구에 최적화된 분자 시뮬레이션 시스템을 제공하고, 이 시스템을 통해 연구자의 연구 생산성을 높이는데 그 목적을 두고 있다. 분자 시뮬레이션

그리드 아키텍처란 최적화된 분자 시뮬레이션 지원을 위한 그리드 시스템의 기본 구성요소를 정의하고, 그에 대한 요소간 기능정의와 인터페이스 및 상호작용을 정의한 것으로 MGrid 시스템의 핵심이다. MGrid 시스템은 1) 연구자 친화적이고, 직관적 연구자 인터페이스와 싱글 시스템 뷰를 제공하는 포털 기반 PSE, 2) 위치 투명한 방법으로 분자 시뮬레이션 서비스를 제공하는 분산 작업 서버, 3) 시뮬레이션 작업에 대한 결과를 위한 분산 시뮬레이션 저장소로 구성되어 있다. Chiral-MGrid 시스템은 이를 기반으로 설계, 구현되었다.

3. 시스템 설계

3.0 Experiment 와 Job 정의

하나의 실험(Experiment)은 광학이성질체 식별 연구의 최소단위이다. 하나의 작업(Job)은 컴퓨터 시뮬레이션을 위한 최소단위이다. 광학이성질체 식별 실험은 두 개의 컴퓨터 시뮬레이션을 통해 이루어진다. 즉, 연구자의 하나의 실험은 한 쌍의 물질인 R, S-enantiomer, 두 개의 Job으로 이루어진다. 각각의 Job은 이를 표현하는 논리적인 메타데이터(e.g. 실험명, 실험 부가설명, 실험시간, Host/Guest 물질명)와 물리적인 파일(시뮬레이션 스크립트 파일, Host/Guest 분자 구조 파일, 분자 파라미터 파일)로 이루어진다. [그림 1]

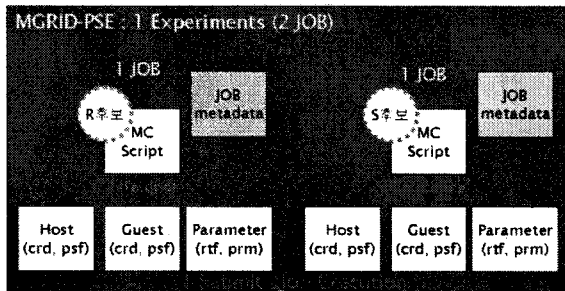


그림 1. Experiment 와 Chiral Job 정의

3.1 시스템 구조

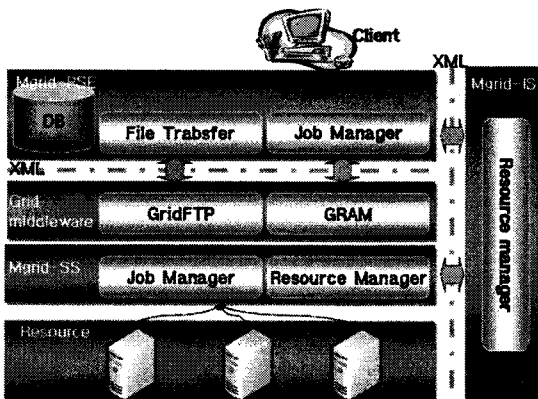


그림 2. 시스템 구조

시스템은 크게 MGrid-PSE, MGrid-IS(Information Server), MGrid-SS(Simulation Server) 3개 부분으로 구성된다. [그림 2]

MGrid-PSE는 광학이성질체 식별 실험을 위한 연구자 인터페이스로 연구자로부터 실험에 필요한 메타 정보(e.g. Job 메타데이터와 파일)를 받아들이고, 이를 통해 하나의 실험(Experiment)을 준비하여 수행하고, 그 결과를 연구자에게 웹을 통해 보여주는 역할을 한다. MGrid-PSE는 하나의 실험을 내부적으로 두 개의 작업(Job)으로 분리하고, 이를 MGrid-SS로 제출한다. 메시지는 XML을 통해 기술되고, 그리드 미들웨어인 GRAM을 통해 전달된다. 작업(Job) 수행을 위한 파일은 그리드 미들웨어의 데이터 전송 컴포넌트인 GridFTP를 이용하여 전송한다.

MGrid-SS는 MGrid-PSE로부터의 작업 요청을 받아 이를 시뮬레이션/분석/시각화 단계를 수행하고, 이에 대한 결과(e.g. 성공/실패, 결과파일)를 MGrid-PSE로 보내주는 역할을 한다. 이것은 연구자가 제출한 시뮬레이션 작업의 관리 및 제어(등록, 수행, 멈춤)를 수행하는 Job Manager와 해당 자원의 정보(e.g. CPU, 노드정보, 메모리 정보 등)를 수집/관리하는 Resource Manager로 구성된다.

MGrid-IS는 각각의 분산된 MGrid-SS로부터 자원 정보를 수집/관리하고, 이에 대한 정보를 XML 형태로 MGrid-PSE에게 전달해주는 역할을 한다. 이것은 MGrid-SS로부터 해당 자원의 정보를 수집하는 수집자와 수집된 정보를 MGrid-PSE의 요청에 따라 전해주는 제공자로 구성된다. 자원 정보로는 작업의 수행에 필요한 호스트명과 IP 주소, 시뮬레이션을 위한 CPU, HDD, 메모리 등 하드웨어 정보와 소프트웨어의 종류, 버전이 포함된다.

각 컴포넌트 간 통신은 XML스키마로 정의되어 있으며 XML문서로 이루어진다.

3.2 실행 시나리오

Chiral-MGrid 시스템은 광학이성질체 분리를 위한 시뮬레이션, 분석, 시각화 단계를 하나의 operation으로 통합된 서비스를 제공하도록 설계된다. 작업의 수행은 [그림 3]와 같이 크게 4단계를 거쳐 수행된다. 연구자가 웹 포털을 통해 작업 제출을 수행하면, 시스템 내부적으로 아래와 같이 동작한다.

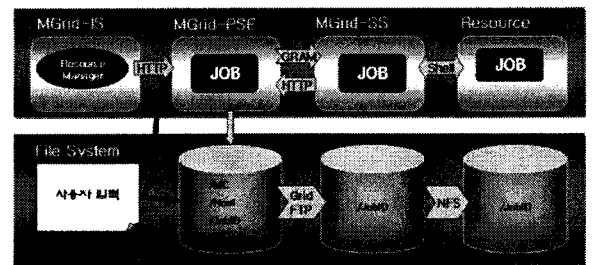


그림 3. 제어 메시지 및 데이터 흐름

① 연구자로부터 광학이성질체 식별 모사실험 요청을 받으면 MGrid-PSE는 MGrid-IS로부터 시뮬레이션을 수행하기 위해 필요한 자원 정보를 읽어 온다. MGrid-IS는 각각의 MGrid-SS로부터 수집된 자원을 확인하고, 가능한 자원을 스케줄링한다. 자원 정보는 XML 문서로 구성되며, HTTP 프로토콜을 통해 MGrid-PSE로 전송된다.

② 연구자의 입력된 정보를 통해 MGrid-PSE는 하나의 실험, 즉 두 개의 Job을 생성하고, 이를 MGrid-SS에 GRAM을 통해 제출한다. 작업 요청은 1)작업 등록, 2)작업시작 두 단계를 통해 이루어진다. 작업 수행에 필요한 데이터 파일은 연구자의 업로드에 의해 분자 구조파일과 템플릿 스크립트(/MC), 호스트(/Host), 파라미터(/Parameter) 파일의 조합으로 구성되어 유일한 JobID로 구별되는 디렉토리에 저장되어 있다가, GridFTP를 통해 전송된다.

③ GRAM을 통해 작업 요청이 MGrid-SS로 들어오면 해당 Job의 정보를 이용해 지역 자원 관리자(e.g. openPBS) 혹은 Shell을 통해 특정 컴퓨팅 자원에 시뮬레이션/분석/시각화 워크플로우를 수행시킨다. 작업 수행에 필요한 데이터 파일은 NFS를 통해 MGrid-SS와 자원 간에 공유된다.

④ MGrid-SS는 현재 진행되고 있는 작업의 상황을 주기적으로 MGrid-PSE에 전송한다. 이것은 분자 시뮬레이션 작업의 특성상 시간이 매우 오래 걸리는 작업이기 때문에, 연구자로 하여금 작업의 현재 진행정도를 알려주기 위해서이다. 시뮬레이션 작업이 완료되면 MGrid-SS는 분석/시각화 작업을 수행하고, 그 결과를 XML문서로 구성하여 HTTP 프로토콜을 통해 MGrid-PSE에 전송한다.

3.3 통신 프로토콜

Chiral-MGrid 시스템의 통신 프로토콜은 XML 스키마로 정의된다. XML[9]은 이질적인 자원으로 이루어진 그리드 환경에서 확장성과 이식성을 제공한다. Chiral-MGrid 시스템에서 정의한 XML 스키마는 다음과 같다.

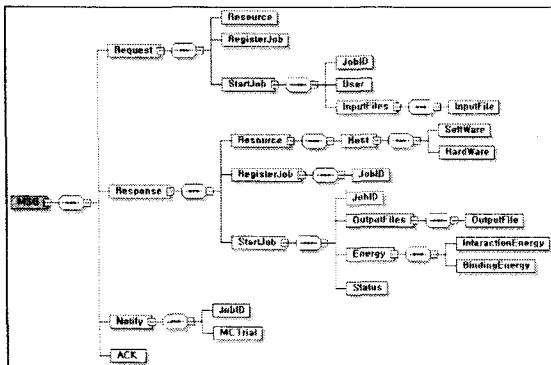


그림 4. XML 데이터 스키마

통신 프로토콜 구조는 크게 4개의 메시지(Request, Response, Notify, ACK)로 나뉘어진다.

<Request>엘리먼트는 MGrid-PSE에서 MGrid-SS, MGrid-IS에 보내는 요청 메시지로서 MGrid-IS에 자원정보에 대한 요청 메시지와 MGrid-SS에 Job을 등록, 시작하는 요청메시지로 구성된다. 여기서 <InputFile> 엘리먼트는 Job에 필요한 파일명을 나타낸다. <Response>엘리먼트는 MGrid-IS, MGrid-SS에서 MGrid-PSE에 보내는 MGrid-PSE의 요청에 대한 응답메시지로서 MGrid-IS에서 보내는 유효한 자원정보에 대한 응답메시지와 MGrid-SS에서 보내는 Job 등록, 시작에 대한 응답메시지로 구성된다. <OutputFile> 엘리먼트는 시뮬레이션 결과를 시각화한 그래프 파일을 나타내고 Energy는 실험결과를 도출할 수 있는 Interaction 에너지와 Binding 에너지의 평균치를 나타낸다. 실험에서 보통 포집 복합체의 결합을 500차 진행하여 얻은 에너지의 평균값으로 결론을 도출하게 되는데 <McTrial>은 이 결합차수를 나타내는 엘리먼트이다. <ACK>엘리먼트는 GRAM을 통해 요청 메시지를 전송한 후 MGrid-SS로부터 받는 Acknowledgement를 말한다.

3. 4 데이터베이스 스키마

[그림 5]는 Chiral-MGrid 시스템의 데이터베이스 스키마를 나타낸다. 스키마는 3개의 테이블로 구성되며, 각 테이블간 JobID를 통해 관계된다. JobInfo 테이블은 시뮬레이션 작업에 대한 입력 메타데이터를 저장하고, JobResult 테이블은 시뮬레이션 작업에 대한 결과 메타데이터 및 결과 파일을 저장한다. 보낸 메시지와 받은 메시지는 시스템의 액션 로그로서 MSG 테이블에 저장된다.

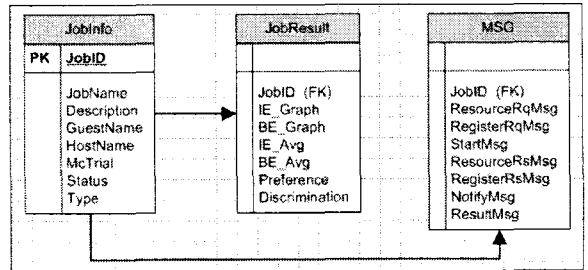


그림 5. DB 스키마

스키마를 구성하는 테이블 각 필드의 의미는 아래와 같다.

[표 1] 데이터 테이블 설명

테이블명	의미	
JobInfo	JobID	시뮬레이션 Job ID
	JobName	실험 이름
	Description	실험에 대한 부가설명
	GuestName	광학이성질체 물질명
	HostName	환경 탄수화물
	McTrial	Job 진행상황을 표시
	Status	Job 상태를 표시
Msg	Type	광학이성질체 쌍의 종류
	JobID	FK
	RegisterMSG	Job 등록요청 메시지
JobResult	StartMSG	Job 시작요청 메시지
	JobID	FK
	IE_Graph	Interaction Energy 그래프
	BE_Graph	Binding Energy 그래프
	IE_Avg	Interaction Energy 평균치
	BE_Avg	Binding Energy 평균치
	Preference	광학이성질체의 R, S 여부
Discrimination	분리가능 여부	

4. 시스템 구현

4.1 구현환경

Chiral-MGrid 시스템의 구현 환경은 아래와 같다.

[표 2] 구현 환경

	MGrid-PSE	MGrid-SS	MGrid-IS
운영체제	Linux RedHat 9.0		
DBMS	Oracle	-	-
구현언어	Gridsphere /JSP	Python	Gridsphere /JSP
웹 서버	Tomcat 5.0.28	-	Tomcat 5.0.28

Gridsphere[10]는 OGSA 그리드 서비스를 기반으로 개발된 그리드 포털 프레임워크로서 그리드 포털을 개발하기 위한 많은 기능을 제공할 뿐만 아니라 확장성과 재이용성이 뛰어나다. Python[11]은 스크립트 기반 언어로서 플랫폼간 독립성과 뛰어난 이식성, 빠른 개발에 대한 장점이 있다. 그리드 환경과 같은 이질적 자원 환경 상에서 python은 매우 편리하다.

4.2 기능

본 논문에서 구현된 Chiral-MGrid 시스템은 1)작업 제출 페이지, 2)시뮬레이션 상태 페이지, 3)작업 결과 요약 페이지, 세 부분으로 구성된다.

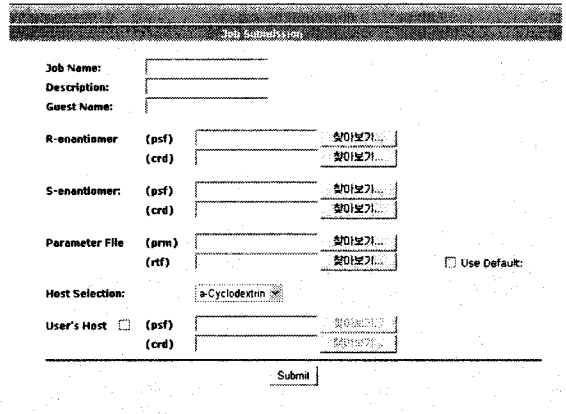


그림 5. 작업 제출 페이지

작업 제출 페이지는 연구자로부터 시뮬레이션작업 수행에 필요한 정보를 입력받는다.[그림 5] 연구자 입력정보는 광학이성질체 분리 모사실험명, 실험에 대한 부가설명, 광학이성질체명, 작업 수행에 필요한 분자 구조파일과 파라미터, 호스트 파일로 구성된다. 파라미터 파일은 연구자가 직접 입력할 수도 있고 실험에 보통 사용되는 디폴트로 설정된 파일을 사용할 수도 있다. 호스트 파일은 자주 사용되는 파일들이 제공되지만 연구자가 파일을 업로드 할 수도 있다. 연구자가 작업을 제출하면 해당 작업의 진행상황 및 결과는 시뮬레이션 상태 페이지에서 확인할 수 있다.

그림 6. 시뮬레이션 상태 페이지

시뮬레이션 상태 페이지는 작업 리스트를 보여주는 화면으로서 실행중인 시뮬레이션의 진행상황과 완료된 시뮬레이션의 결과를 보여준다.[그림 6] 여기서 MCTrial은 시뮬레이션에서 Host물질과 광학이성질체 물질을 결합시키면서 에너지를 측정하는 차수를 나타낸다. Status는 작업 진행상황을 나타내는 "ready", "running", "done" 세 개 단계로 구분된다. "ready"는 시뮬레이션이 요청되었지만 자원에 할당되지 않은 상태이고 "running"은 시뮬레이션이 실행중인 상태이며 "done"은 작업이 완료된 상태를 의미한다. 작업 제출 페이지에서 제출한 작업은 한 쌍의 물질인 R, S-enantiomer, 두 개의 Job으로 나누어 수행되는데 시뮬레이션이 완료된 한 쌍의 작업을 선택하면 분석결과를 작업 결과 요약 페이지에서 확인할

수 있다.

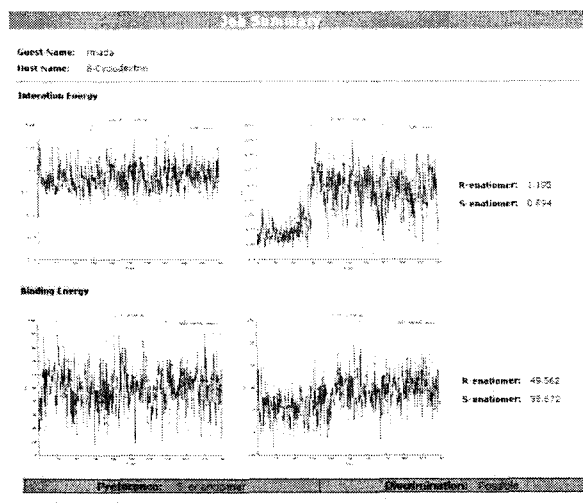


그림 7. 작업 결과 요약 페이지

작업 결과 요약 페이지는 시뮬레이션결과를 분석, 시각화하여 연구자에게 보여준다. [그림 7] 출력 데이터는 Guest명, Host명, 시뮬레이션 결과 데이터로부터 추출한 R, S-enantiomer 각각의 Interaction 에너지, Binding 에너지 그래프와 각각의 에너지 평균치, 결론으로 구성된다. 그래프는 시각적으로 에너지를 비교할 수 있도록 해주었고 에너지 평균치는 결론을 도출하는 근거가 된다. R, S-enantiomer의 Binding 에너지 평균치로 광학이성질체의 Discrimination 가능여부를 판단할 수 있는데 에너지 값의 차이가 0.5kcal 이하(0.5kcal 포함)이면 광학이성질체는 Discrimination이 불가능하다. 에너지 값의 차이가 0.5kcal 이상이면 광학이성질체는 Discrimination이 가능하며 에너지가 낮은 물질이 Preference를 갖는다.

5. 결론 및 향후계획

본 논문에서는 광학이성질체 식별 모의실험을 그리드 기반의 웹 서비스를 통해 시뮬레이션/분석/시각화를 하나의 오퍼레이션으로 수행하도록 하여, 연구자로 하여금 분자 구조 파일 만을 정의하여 그 식별 여부를 알아볼 수 있도록 하는 Chiral-MGrid 시스템을 설계, 구현하였다.

Chiral-MGrid 시스템은 웹 기반으로 설계되었기에 연구자들은 장소의 제한을 받지 않을 뿐만 아니라, 익숙한 환경을 제공하며, 그리드의 환경을 감춤으로서 연구자에게 사용하기 쉬운 환경을 제공한다. 또한 광학이성질체 모의실험의 복잡한 단계를 하나의 오퍼레이션으로 수행함으로써, 연구자로 하여금 구체적인 시스템의 흐름이나 응용 애플리케이션의 이해 없이 원 클릭으로 모든 과정을 수행하고 결과를 확인할 수 있다는 장점이 있다.

Chiral-MGrid 시스템은 MGrid시스템을 기반으로 설계,

구현되었다. 향후, 이 시스템은 현재, 연구자를 위한 분자 구조 Builder나 향상된 작업 브라우징, 다른 시뮬레이션 소프트웨어 지원 등을 위해 MGrid 시스템과 통합, MGrid 시스템상의 하나의 서비스로서 제공할 예정이다.

6. 참고문헌

- [1] Pirkle, W.H. and Pochapsky, T. C. *Chem. Rev.* 89, 347-362, 1989
- [2] Lipkowitz, K. B. Coner, R Peterson, M. A. *J. Am. Chem. Soc.* 1997, 119, 11269-11276.
- [3] Dodziuk, H. and Lukin, O. *Chem. Phys. Lett.* 327, 18-22, 2000
- [4] Wolbach, J. P. Lloyd, D. K. Wainer, I. W. *J. Chromatogr. A.* 914, 299-314, 2001
- [5] Booth, T. D. Azzaoui, K.; Wainer, I. W. *Anal. Chem.* 69, 3879-3883, 1997
- [6] Natrajan, A.; Crowley, M.; Wilkins, N.; Humphrey, M.A.; Fox, A. D.; Grimshaw, A. S.; Brooks, C. L. III. *High Perform. Distrib. compu.* 10, 1-7, 2001
- [7] Ferguson, D. M.; Raber, D. J. *J. AM. Chem. Soc.* 111, 4371-4378, 1989
- [8] MGrid 시스템 <http://www.mgrid.or.kr>
- [9] XML <http://xml.org>
- [10] GridSphere <http://www.gridspheer.org>
- [11] Python <http://python.org>