

단백질 가시화 서비스 이용 현황 분석

변재희*, 최유주**, 서정근**

*한독미디어대학원대학교 입체영상미디어학과

**한독미디어대학원대학교 뉴미디어콘텐츠학과
e-mail:bjach9188@gmail.com

The Present Situation Analysis on Services of Protein Visualization

Jaehye Byeon*, Yoo-Joo Choi**, Jung-Keun Suh**

*Dept. of Stereoscopic Media, KGIT

**Dept. of Newmedia Contents, KGIT

요 약

단백질 의약품 개발 시 단백질 구조는 단백질 기능을 규명하기 위한 필수적인 정보이다. 따라서 단백질 구조를 효과적으로 전달할 수 있는 단백질 가시화 서비스가 증가하고 있으며, 대표적으로 Cn3D, Jmol, Chimera 등이 있다. 각 서비스는 단백질 가시화 서비스와 함께 단백질 분석을 위한 부가 기능을 제공하고 있다. 본 논문에서는 단백질 의약품의 효율적 정보전달 서비스를 위한 사전 연구로 업계종사 기간이 5년 이상인 피험자를 대상으로 단백질 가시화 이용현황을 설문하였다. 설문 분석 결과 자주 이용하는 단백질 가시화 서비스로 피험자 모두 PDB를 사용한다고 응답하였으며, 단백질 1, 2차구조 혼합 가시화 서비스를 일주일에 4회 이상 사용하는 응답자가 58.3%였다. 특히 단백질 의약품 개발 시 반드시 필요한 단백질 정보로는 91.6%가 단백질 1차구조, 단백질 2차구조, 단백질 2차구조 비율이라고 응답하였다.

1. 서론

단백질 데이터가 점차 증가하면서 단백질 정보를 효율적으로 전달하기 위한 연구 또한 증가하고 있다. 대표적 바이오데이터베이스인 NCBI에서는 단백질 가시화 서비스인 Cn3D를 자체 개발하여 제공하고 있으며, PDB에서는 오픈소스인 Jmol을 이용하여 다양한 단백질 가시화 서비스를 제공하고 있다. 대표적 단백질 가시화 서비스로는 Chimera와 Jmol이 있으며 두 서비스 모두 오픈소스로 단백질 정보 분석 등의 추가 기능을 제공한다는 특징이 있다. 하지만 단백질 의약품에 특화된 단백질 가시화 서비스에 대한 연구는 미진한 실정으로 단백질 의약품에 특화된 단백질 가시화 서비스 연구가 필요하다[1-2].

본 논문에서는 단백질 의약품에 대한 효율적 정보제공 서비스에 대한 사전 연구로 단백질 가시화 서비스의 이용현황을 설문 분석하고자 한다.

2. 단백질 가시화 서비스

단백질 가시화 서비스는 바이오데이터베이스에서 제공하는 서비스와 독립적인 단백질 가시화 서비스로 구분할 수 있다. 바이오데이터베이스에서 제공하는 단백질 가시화 서비스로는 NCBI의 Cn3D와 오픈소스인 Jmol을 이용하여 기능을 확장한 PDB 가시화 서비스가 있다. 독립적인 단백질

가시화 서비스로는 Chimera와 Jmol이 있다. 다음의 <표 1>은 단백질 가시화 서비스를 비교한 것으로 PDB의 경우 Jmol을 이용하여 가시화하므로 동일 항목으로 두었다[1-2].

<표 1> 단백질 가시화 서비스 비교

구분	Chimera	PDB(Jmol)	Cn3D
장점	<ul style="list-style-type: none"> • 오픈소스 • 다양한 기능 제공 • 단백질 서열정보 가시화 • subunit 가시화 	<ul style="list-style-type: none"> • 오픈소스 • 자바만 설치되어 있으면 작동 	<ul style="list-style-type: none"> • 오픈소스 • NCBI에서 개발
단점	<ul style="list-style-type: none"> • 복잡한 UI • 렌더링 속도가 느림 	<ul style="list-style-type: none"> • 고수준의 가시화 어려움 • 불편한 UI 	<ul style="list-style-type: none"> • 렌더링 속도가 느림 • 불편한 UI
가시화 요소	<ul style="list-style-type: none"> • 그래픽, 텍스트 	<ul style="list-style-type: none"> • 그래픽, 텍스트 	<ul style="list-style-type: none"> • 그래픽, 텍스트
Library	<ul style="list-style-type: none"> • OpenGL 		<ul style="list-style-type: none"> • OpenGL
Language	<ul style="list-style-type: none"> • Python, C++ 	<ul style="list-style-type: none"> • Java 	<ul style="list-style-type: none"> • C++

* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2014년도 산업계 맞춤형 인력양성지원사업의 연구결과로 수행되었음

3. 단백질 가시화 서비스 이용 현황

3.1 조사대상 및 방법

단백질 가시화 서비스의 이용 현황을 분석하기 위해 업계 종사기간이 5년 이상인 전문가 집단을 대상으로 설문을 진행하였다. 설문에 응답한 피험자는 총 12명으로 설문 응답에 대해서는 빈도분석을 실시하였다.

자주 이용하는 단백질 가시화 서비스의 설문 항목은 Chimera, PDB, Cn3D, 기타였으며 단백질 가시화 서비스 이용 횟수는 일주일에 1~2회 이내, 3~4회, 4회 이상 기타였다. 단백질 의약품 개발을 위해 반드시 필요한 단백질 정보 설문 항목은 단백질 1차구조, 아미노산 위치, 단백질 2차구조, 단백질 2차구조 비율, 기타로 중복 선택이 가능하도록 하였다.

3.2 단백질 가시화 서비스 이용자 특성

단백질 가시화 서비스 이용 현황 분석 전 피험자가 전문가 집단에 속하는지 판단하기 위해 개인정보에 대한 설문을 진행하였다. 개인정보는 업계종사기간과 학력 정보를 수집하였다.

피험자 중 업계종사기간이 10년 이상인 경우는 총 9명으로 75%였으며, 5년 이상인 경우는 3명으로 25%였다. 학력은 박사졸업이 총 9명으로 75%였으며, 석사졸업은 3명으로 25%였다. 다음의 <표 2>는 피험자 분석 결과이다.

<표 2> 피험자 분석

구 분		빈도 (명)	백분율 (%)
학력	대학원 졸(석사)	3	25
	대학원 졸(박사)	9	75
업계종사 기간	5년 이상~10년 미만	3	25
	10년 이상	9	75

3.3 단백질 가시화 서비스 이용 현황

자주 이용하는 단백질 가시화 서비스로 피험자 모두 PDB라고 응답하였다.

<표 3> 자주 이용하는 단백질 가시화 서비스

구 분	빈 도(명)	백분율(%)
Chimera	0	0
Cn3D	0	0
PDB	12	100
기타	0	0
합 계	85	100

3.4 단백질 가시화 이용 횟수

피험자 중 단백질 가시화 이용 횟수는 일주일에 4회 이상이 7명인 58.3%로 가장 많았으며, 3~4회는 4명으로 33.3%, 1~2회 이상은 1명으로 0.1%였다.

<표 4> 단백질 가시화 서비스 이용 횟수

구 분	빈 도(명)	백분율(%)
Chimera	0	0
Cn3D	0	0
PDB	12	100
기타	0	0
합 계	85	100

3.5 단백질 의약품 개발 시 필요한 단백질 정보

피험자 중 단백질 의약품 개발 시 반드시 필요한 단백질 정보로 11명인 91.6%가 단백질 1차구조, 단백질 2차구조, 단백질 2차구조 비율이라고 응답하였다. 아미노산 위치에 대해서는 3명인 25%가 반드시 필요하다고 응답하였으며, 기타 의견은 없었다.

<표 5> 단백질 의약품 개발 시 필요한 단백질 정보

구 분	빈 도(명)	백분율(%)
단백질 1차구조	11	91.6
아미노산 위치	3	25
단백질 2차구조	11	91.6
단백질 2차구조 비율	11	91.6
기타	0	0

4. 결론

단백질 의약품의 효율적 정보제공 서비스를 위한 사전 연구로 본 논문에서는 단백질 가시화 서비스 이용 현황을 분석하였다. 이를 위해 업계 종사기간이 5년 이상인 전문가를 대상으로 설문을 진행하였다.

설문의 결과 자주 이용하는 단백질 가시화 서비스로 피험자 모두 PDB를 이용하였으며, 이용 횟수는 일주일에 4회 이상이 가장 많았다. 단백질 의약품 개발 시 반드시 필요한 단백질 정보로는 단백질 1차구조, 단백질 2차구조, 단백질 2차구조 비율에 대한 응답이 가장 많았다.

PDB는 가장 대표적인 단백질 데이터베이스이며, 피험자가 단백질 개발 시 필요한 정보라고 응답한 것은 PDB에서 단백질 1, 2차구조 혼합가시화 형태로 제공하고 있다. 따라서 본 설문 결과를 통해 단백질 의약품에 특화된 가시화 서비스의 필요성이 필요성을 확인할 수 있었다.

본 연구의 결과를 바탕으로 향후 다양한 단백질 1, 2차 혼합가시화 형태들의 단백질 정보표현적합도를 분석하고자 한다.

참고문헌

- [1] 변재희, “단백질 의약품의 효율적 정보 전달을 위한 웹사이트 설계”, 한독미디어대학원대학교 석사학위논문, 2014.
- [2] 송광은, 최유주, 서정근, “단백질-단백질 상호작용 인터페이스 정보 가시화”, 한국정보처리학회 학술대회 발표집, Vol. 21, No.1, pp. 677~679, 2014.