프로테오믹스

연세대학교 의과대학 이비인후과학교실
김 창 훈

Proteomics

Chang-Hoon Kim, MD
Department of Otorhinolaryngology, Yonsei University College of Medicine, Seoul, Korea

서 론

근래의 생물학 및 의학 관련 연구는 연구의 결과로 얻어지는 수많은 정보들을 어떻게 체계화하고 저장하며 이들을 어떻게 활용하는가에 초점이 모아지고 있다. 이러한 현상은 인간 게놈 프로젝트가 발표된 이후 인간의 유전자 염기서열이 완전히 해독된 현 시점에서 더욱 본격화되고 있음을 알 수 있다. 인간의 유전자의 기능에 대한 연구가 활발히 진행되고 있고 그 결과로 방대한 양의 생물학 관련 정보가 쏟아져 나올 것을 예상한다면 이 결과들을 활용 할 수 있는 효과적인 방안들을 마련해 놓아야 하는 당위성은 자명한 것이다.

인간 게놈(유전체)의 완전 해독에 이어 앞으론 단백질 연구가 본격화할 전망이다. 인간 게놈 지도 완성 결과 유전자가 당초 예상보다 훨씬 적은 3만여 개에 불과하고, 하나의 유전자가 한 개의 단백질을 생성하고 이것이 한 질병과 연결될 것이라는 전제가 무너지면서 단백질의 정체를 밝히는 작업이 차세대 연구과제로 떠오르고 있는 것이다. 완전히 해독된 인간 게놈 지도는 한 개의 유전자가 많게는 수백 가지의 단백질을 만든다는 새로운 사실을 일깨워줬다. 따라서 중요한 것은 유전자 자체가 아니라 각각의 유전자가 복잡한 상호작용을 거쳐 만들어내는 단백질 조합이라는 게 과학자들의 일치된 견해다.

생물학 관련 대량 정보 생산방식은 genomics 연구의 DNA array, proteomics, CCL(combinatorial chemical library), robotics를 이용한 HTS(high throughput screening), 생물정보학(bioinformatics) 기술 등에 의해 주도되고 있으며 이미 많은 양의 생물학 및 의약 관련 정보가 축적되고 있다. 그러나 10만 여개로 추정되고 있는 인간의 유전자중에서 만여 개의 유전자만이 그 기능이 알려져 있으며, 이러한 기능이 알려져 있지 않은 유전자의 기능을 연구하는 functional genomics의 한 분야로서의 proteomics는 유전자 또는 단백질의 기능을 연구하는 효과적인 기술로서 활용되고 있다. Proteomics는 생물학 관련, 질병관련 유전자 또는 단백질의 발현변화와 변형 상태 또는 생화학적인 활성상태 등에 대한 정보를 대량으로 빠르게 제공해줄 수 있다. 따라서 이러한 정보 획득 능력을 바탕으로 유용한 정보들을 생산, 축적하며 이 정보들을 효과적으로 활용하여 생물학, 의약관련 연구 및 산업적으로 이용할 수 있도록 하기 위해서는 이러한 대량 정보 생산 방식을 도입하고 빠르게 정착시켜 생물 및 의학 관련 연구능력을 발전시켜야 할 것으로 본다.

프로테옴(Proteome)과 프로테오믹스(Proteomics)

Proteome(단백체)이란 PROTEin expressed by a genOME의 합성어로[1] 생명체가 특정한 환경 하에서 특정 유전자의 지시에 따라 생합성된 총 단백질을 의미한다. Proteomics(단백체학)는 프로테옴(proteome)의 어미에 학, 론을 의미하는 접미사 *-ics*가 붙어 프로테옴을 연구하는 방법과 기술을 포괄적으로 의미하는 말로, 굳이 번역한다면 ' 프로테옴(단백체) 분석학 '이라고 해야 할 것이다. 즉, 단백질의 성질을 발현, post-translational modification (번역후 변형), 다른 단백질과의 결합에 초점을 두어 연구함으로써 세포내 변형과정과 네트웍 형성을 질병의 진행과정과 연계시켜 총괄적으로 이해할 수 있는 연구 분야를 뜻한다.

프로테오믹스라 함은 결국은 단백질의 기능을 연구하는 분야로 기존의 단백질 연구와 목적상으로 볼 때 크게 다르지 않다. 다만 세포 또는 조직 등을 구성하는 단백질의 전체(프로테옴)를 대상으로 해서 대량으로 그리고 단백질의 상대적인 양을 훼손하지 않고 분석을 한다는 점에서, 기존

의 몇몇 개별적인 단백질을 대상으로 수행하던 방식과 차이를 가진다고 할 수 있다. 그러나 프로테오믹스에 국한된 것은 아니겠으나 이러한 연구(system biology)의 방법론은 가설에 의존해서 접근하던 종래의 방식과는 조금 괘를 달리하는 결과를 가져올 것으로 전망하고 있다. 즉 예측되지 않았던 새로운 정보들로부터 다른 해석을 가능하게 하고 통합된 정보들로부터 새로운 가설들을 빠르게 세워나감으로써 연구 분야에 대한 지식의 폭과 정보 생산 속도를 가속화할 수 있다는 것이다.

프로테오믹스와 지노믹스는 공히 유전자의 기능을 밝힌다는 점에서 목표는 같으나 방법이 매우 다르다. 기술적으로는 실험의 출발점도 전자가 단백질 차원에서, 후자는 유전자차원에서 각각 유전자의 기능을 연구를 하는 기술이므로 연구형태가 매우 다르다. 그러나 이 둘은 실제 유전자구조를 이용한다는 점에서 한 뿌리이며 상호 보완적인 연구분야이다. 가령 지노믹스가 세포의 핵산(DNA, RNA, 또는 이들의 유사체)을 대상으로 각 유전자발현의 정도차이를 분석하고, 유사유전자들의 수집 및 분류, 특정 생리조건에 반응하는 유전자들의 대량분석(DNA칩 및 올리고 칩), 특정 유전자의 발굴(mining)을 수행하는 연구 분야로서, 이미 알고 있는 유전자구조를 근간으로 연구하는 기술인데 비해, 프로테오믹스는 일단 유전자의 산물인 단백질을 대상으로 조직이나 세포, 체액과 같은 개체의 실시간 세포생리현장에서 얻어진 시료를 대상으로, 이들을 대량으로 분석하고, 상호기능관계의 지도를 작성하며, 구조분석을 통하여 궁극적으로 특정 단백질과 이를 만드는 유전자의 기능을 동시에 밝혀내는 기술이다.[3] 실제로, 특정 유전자의 발현과 단백질의 발현차이에는 밝혀진 유전자를 기준으로 30% 이상이 상관성이 거의 없는 것으로 알려져, 그만큼 지노믹스의 수단으로 단백질과 유전자의 발현(mRNA) 상관성을 분석하고 예측하기는 불가능하다. 이 두 분야에 속한 공통 분야가 있는데, 구조유전체학(sturctural genomics) 또는 구조프로테오믹스(structural proteomics)라고도 불리는 단백질 구조분석 연구이다. 이 분야는 프로테오믹스에 의해 밝혀진 특정 단백질이나 이미 알려진 게놈 DB 상의 특정유전자 정보를 바탕으로 단백질의 1차 구조를 예측하고, 구조와 기능간의 관계를 유추하여 궁극적으로는 생리적 활성상태인 가상적인 3-D 구조를 추측해 낸다.[4] 이러한 결과를 사용하여 drug ligand나 receptor의 공간적 관계를 추정할 수 있으며, 이를 drug design에 응용할 수 있다.

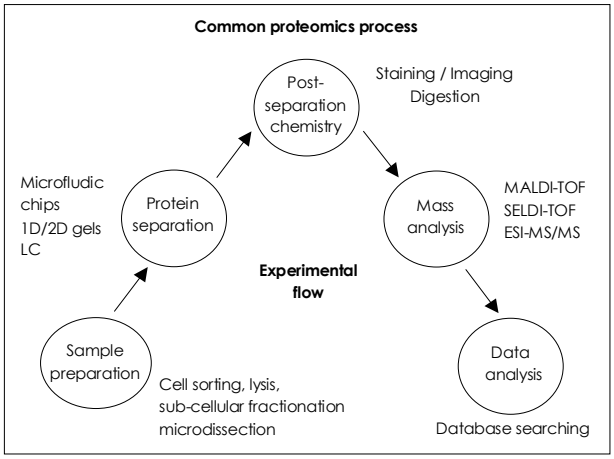

Fig. 1. Common Proteomic process. Sample preparation is the first step of the proteomic process and the second step is protein separation, such as protein chip, 1-D and 2-D gels, and the liquid chromatography. The third step is post-separation chemistry, such as staining, imaging, and in-gel digestion. The fourth step is the mass analysis step. MALDI-TOF, SELDI-MS, ESI-MS/MS can be used . The last step is data analysis.

프로테오믹스는 단백질의 분리와 단백질의 동정에 관한 기술이 그 기술적 핵심을 이루고 있다(Fig. 1). 이차원 전기영동법(2 dimensional gel electrophoresis)에 의한 단백질 분리와 MALDI-TOF(matrix assisted laser desorption/ionization time-of-flight), MS/MS(or tandem mass), ESI(electrospray ionization) MS에 의한 단백질 동정 및 특성연구는 현재의 프로테오믹스를 가능하게 한 중요한 기술들이며 이들 방법에 대한 보완 또는 대안으로서 LC(liquid chromatography) MS/MS, SELDI(surface enhanced laser desorption/ionization) technology(Ciphergen) 등이 기술적 발전을 거듭하고 있다.

프로테오믹스는 구체적으로 유전체구조와 유전자의 세포내 행동간 갭(gap)을 메우는 역할을 하는 단백질 생성물 및 그들 간의 네트워크를 연구하는 분야이다. 생명공학 분야에서 이 기술의발전은 대량발굴탐색(high-throughput screening)이 가능한 2차 전기영동(two dimensional electrophoresis : 2DE) 분석과 MALDI-TOF(matrix-assisted laser desorption/ionization time-of-flight)에 의한 자동화된 단백질 구조 분석 기술 및 이들을 지원하는 프로테옴 정보학(proteome informatics)의 발전과 긴밀하게 연계되어 있다. 따라서 프로테오믹스는 세포의 생리적 상태 변화에 따른 분자적인 현상과 세부적인 기전을 규명할 뿐만 아니라 어떻게 단백질의 표현형(phenotype)이 질병과 약물 처리에 따라 변화하는 지를 분석할 수 있고, 이로 인해 약물표적의 식별과 검증을 할 수 있게 한다.[2]

아무튼 한 유전자가 최종적으로 세포 내에서 어떠한 기능을 하는가는 얼마나 정교하고, 적절하게 단백질 합성후 변형(post-translational modification) 되는 가에 달려 있

어서, 프로테오믹스를 이용해 최종적으로 완벽한 모양과 기능이 갖추어진 단백질을 분석하지 않고는 그 유전자의 세포내 기능을 알 방법이 없기에 프로테오믹스의 독특한 가치가 여기에 있는 것이다.

방 법

현재 이루어지고 있는 프로테오믹스 연구는 크게 두 가지 기술로 집약될 수 있다. 먼저 한 세포에서 얻어진 총 단백질군을 서로 분리하여 개개의 단백질로 분리하는 기술과 분리된 단백질들을 분석할 수 있는 기술인데, 첫 번째 기술은 현 단계에서 isoelectric focusing(IEF)을 이용한 2D Gel이 개발되어 있으며 상용화된 기기들이 속속 출현하여 종래의 2D Gel의 문제점인 재현성을 크게 향상시켰고, 2D Gel에서 개개의 단백질들을 염색하고, 잘라내고 단백질 분해효소로 자르고 하는 단계들이 자동화 기기와 컴퓨터를 이용해 많은 시료를 손쉽고 간단히 처리 할 수 있지만, 처음 단계인 2D Gel 수행의 자동화가 아직도 이루어지지 않고 있다. 각각의 분리된 단백질의 분석을 위해 질량 분석기가 쓰이고 있으며 시대의 요구에 부응하여 고성능의 질량 분석기가 속속 개발되어, 미량의 단백질을 간편히 분석하고 아미노산 서열의 분석도 가능하게 하고 있다. 단백질을 질량분석기로 분석해 얻은 결과는 기존의 Database를 이용해 Web상에서 분석이 가능하며 대용량의 연구를 위해서는 독자적인 Database의 확보가 필요하다. 앞으로의 발전 방향은 총 단백질의 효과적 분리를 위한 여러 가지 방법의 개발과 발현량이 적은 단백질의 검출방법의 개발이 될 것이며, 분리된 단백질의 분석은 질량분석기가 가장 유용한 도구로서 쓰일 것이다.

단백체 연구의 기본 3단계는

1) 총 단백질의 분리(protein mapping) : 2D - electrophoresis

2) Peptide mass fingerprinting : MALDI - TOF MS (Matrix Assisted Laser Desorption/Ionisation Time - Of - Flight Mass Spectrometry) 또는 ESI - MS(Electrospray Ionization Mass Spectrometry)

3) Identification of protein : data base(bioinformatics)

로 이루어져 있다.

단백질의 분리 : 2D - electrophoresis

모든 단백질들은 구성하고 있는 amino acid에 따라 질량과 pI(isoelectric point)값이 다르다. 따라서 이 두 요인을 이용하여 단백질을 분리할 수 있다. pI를 이용하여 단백질을 펼치는 방법을 IEF(isoelectric focusing)라고하며 질량에 의해서 단백질을 펼치는 방법을 SDS - PAGE(SDS - polyacrylamide gel electrophoresis)라고 한다.

Sample preparation

Sample Preparation시 가장 중요한 점은 얼마나 잘 녹는가, 적은 양의 단백질로 2 - D spot을 잘 볼 수 있는가이다. 주로 Urea, Thiourea, CHAPS(3 - [(3 - cholamidopropyl) dimethylammonio] - 1 - propanesulfonate)를 위주로 한 sample buffer를 사용한다.[5] 조직을 이용하는 경우 초고속 원심분리기(ultra - high speed centrifuge)를 이용하기도 한다.

First dimension-IEF

단백질을 pI값에 의해 펼치는 방법으로 1975년 O 'Farrell에 의해서 개발되었다.[6] Immobilized pH Gradient (IPG)라는 개념이 도입되어 immobiline이라는 zwitter ion을 이용

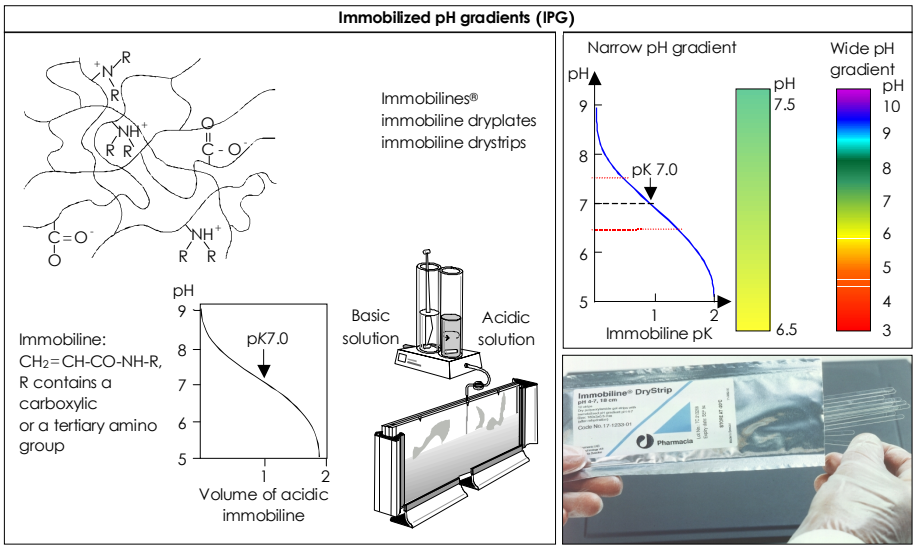

Fig. 2. Immobilized pH gradient (IPG) concept in 1-D separation. A linear pH gradient has to be cast. These IPG gels are cast on film support and one-dimensional separations are run on a horizontal gel.

pH gradient를 규칙적으로 생성한다(Fig. 2). 요즈음 IPG gel strip이 상품화되어 재현성이 증가되었다(Figs. 3 and 4).

Second dimension-SDS-PAGE

단백질을 분자량(molecular weight)에 의해 분리하는 방법으로 horizontal system과 vertical system(Fig. 5)을 이용한다.

Staining

Gel 안에서 구현된 단백질의 spot을 염색하는 방법으로 silver staining(detection limit 0.5~1.2 ng)과 commassie stain(detection limit 8~26 ng), SYPRO Orange 또는 SYPRO Ruby stain(detection limit 2~10 ng)을 사용한다.

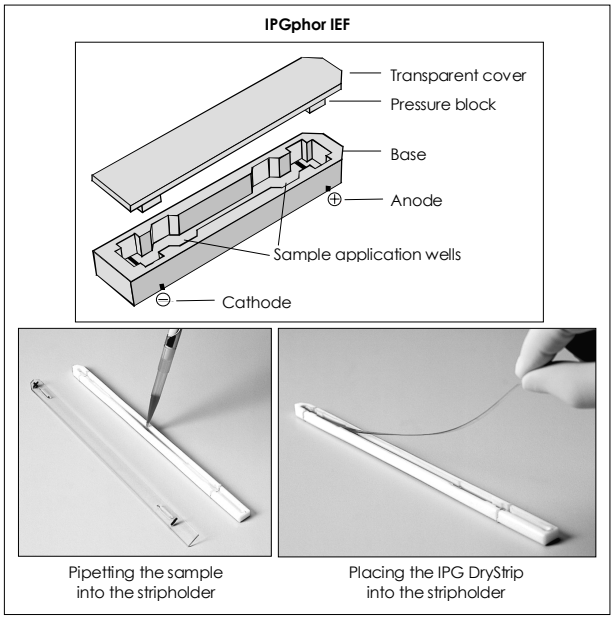

Fig. 3. Isoelectric focusing on IPGphor (1). The sample is being pipetted into the base in the rehydration solution.

Image acquisition

염색된 gel을 스캔하여 image file로 전환하는 과정으로 GS-800 Calibrated Imaging Densitometer, Fluor-S MAX MultiImage System, Molecular Imager FX 등을 사용한다.

Image analysis

스캔된 image file을 분석하여 비교함으로써 발현의 증가 또는 감소, spot의 이동 등을 자동으로 분석해 주는 software로써 PDQuest 2-D Software와 Melanie III 2-D software 등이 널리 쓰인다(Figs. 6 and 7).

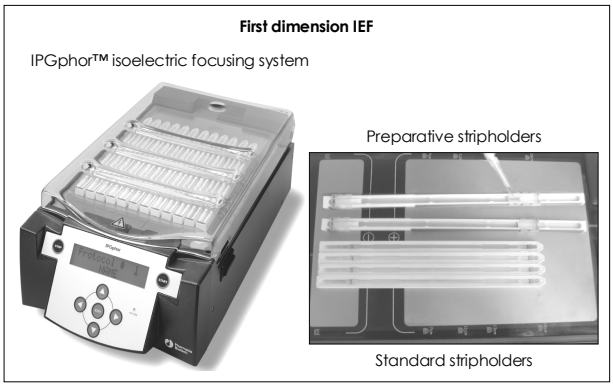

Fig. 4. Isoelectric focusing on IPGphor (2). We need an instrument to run the Immobiline DryStrips on the IPGphor Isoelectric Focusing System. IPGphor is a dedicated system optimized to perform the first dimension IEF step using Immobiline DryStrip. An integrated high voltage power supply and a cooling system minimize focusing time and accurately control voltage and temperature to give high reproducibility between runs.

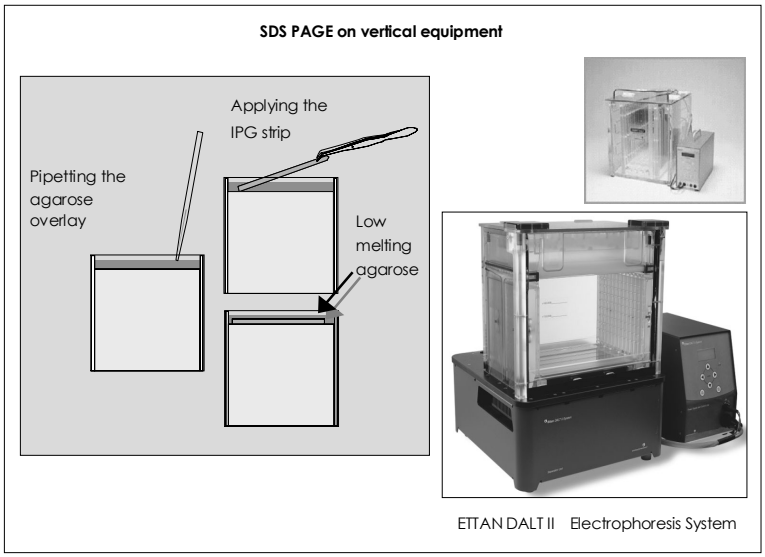

Fig. 5. 2-D separation with SDS-PAGE. Focused-dry-strip is placed on the 2-D gel and embeded with agarose overlay. This picture is Ettan Dalt II electrophoresis system of Amersham biotech, which is most popular 2-D system.

Peptide mass fingerprinting : MALDI - TOF MS(Matrix Assisted Laser Desorption/Ionisation Time - Of - Flight Mass Spectrometry)

Trypsin digested peptides

Image analysis를 통해 선정된 단백질 spot을 채취하여 분해하는 과정으로 modified trypsin을 이용하여 gel 안에 존재하는 단백에 따라 특정 크기의 peptide로 분해한다.

MALDI-TOF MS

Mass spectrometer의 원리는 전하를 띤 물질은 전기장 하에서 움직이게 되는데 이 운동의 크기는 물질의 질량과 전하의 비에 따라 결정된다(m/z). 이 때 물질이 움직인 거리나 비행시간(time of flight)을 알면 질량을 구할 수 있다. MALDI - TOF MS의 원리는 단백질 용액과 matrix 분자(CHCA[cyano - 4 - hydroxy cinnamic acid])를 혼합하고 laser pulse(N_2 laser, 337 nm)가 matrix에 박혀 있는 peptide를 치게 되면molecular ion들이 전기장 하에서 힘을 받아 이동하게 되고 이동한 시간과 거리를 통해서 각 peptide의 질량을 detection하게 된다(Fig. 8).

Identification of proteins from data base

Detection한 질량의 peptide를 가지는 단백을 web 상의 database와 비교하여 이 단백이 어떤 단백인지를 알아내

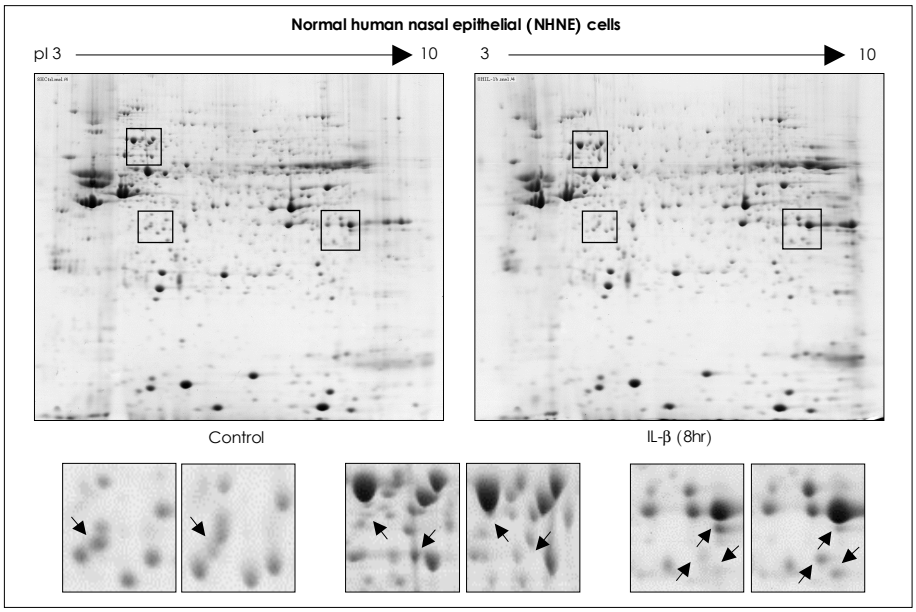

Fig. 6. Example of 2D result that compares normal human nasal epithelial cells treated with IL-1 β and its control. Here, we can see a number of spots that have changed.

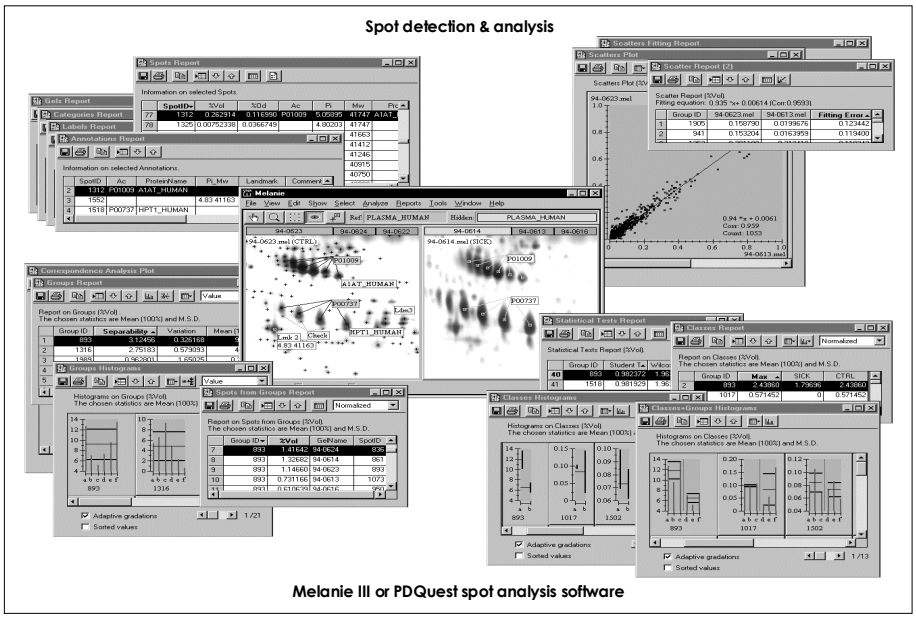

Fig. 7. Spot detection and analysis. Melanie III or PDQuest spot analysis software are used as the programs for spot detection.

는 과정이다(Figs. 9 and 10).

단백질 정체를 밝히는데 이용할 수 있는 유용한 internet site들은 다음과 같다.

1) Intact protein molecular mass

- peptideSearch :
  http://www.mann.embl - heidelberg. de/Services/PeptideSearch/PeptideSearchIntro.html
- TagIdent :
  http://www.expasy.ch/www/quess - prot.html

2) Proteinm sequence tags

- peptideSearch :
  http://www.mann.embl - heidelberg. de/Services/PeptideSearch/FR_sequenceOnlyForm.html
- TagIdent :

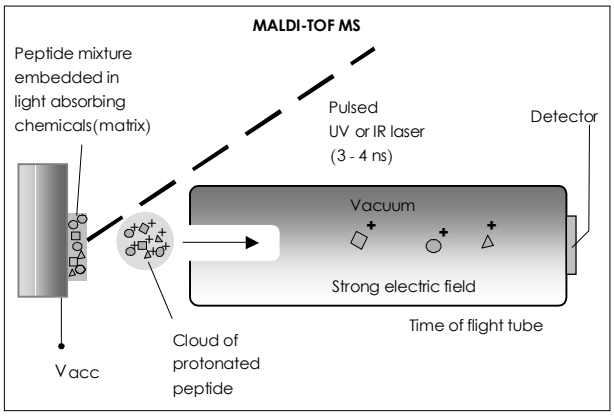

Fig. 8. Mechanism of MALDI TOF MS. The high energy of the laser is absorbed by the matrix and is transferred into excitation energy. The peptides are sublimated and protonated (up to 300 kDa). All ions with different masses are accelerated to the same kinetic energy. When leaving the acceleration region, the ions move down the field, velocities are reversely proportional to the square roots of their masses. 10 - 50 spectra from a single laser shot are summed for signal averaging to achieve accurate mass determination.

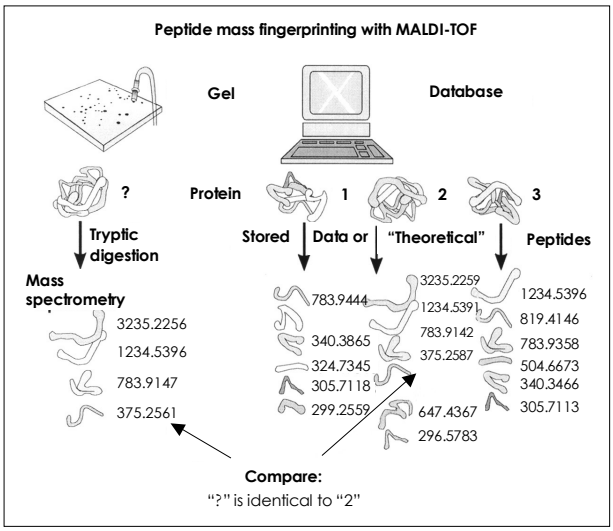

Fig. 9. Peptide mass fingerprinting with MALDI-TOF (1). The petide mass that has been obtained through mass spectrometry is compared with the peptide mass in the protein database. This process allows us to identify the protein.

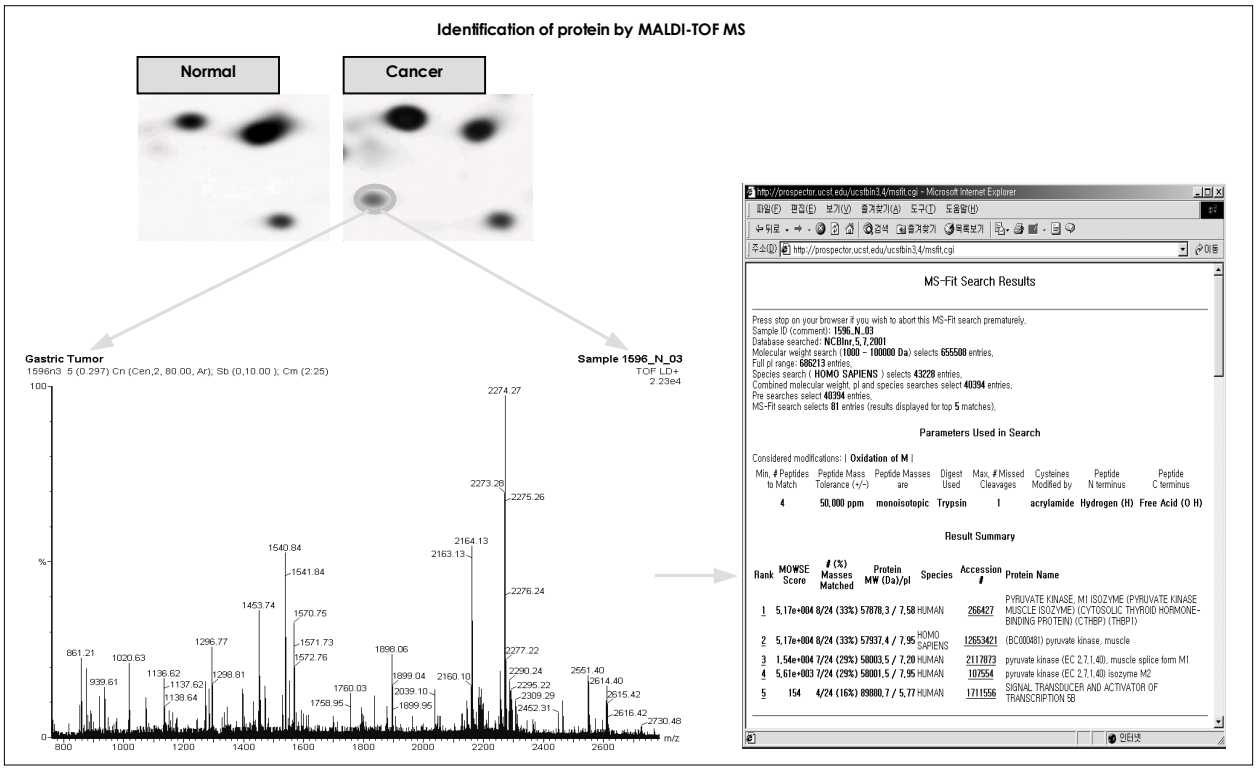

Fig. 10. Peptide mass fingerprinting with MALDI-TOF (2). The changed spot is picked up and analyzed wi-th mass spectrometry. This mass is then compared with the mass in the database for peptide mass fin-gerprinting.

http://www.expasy.ch/www/quess - prot.html
3) Peptide mass fingerprinting
- MassSearch :
http://cbrg.inf.ethz.ch/subsection3_ 1_3.html
- MS - Fit :
http://falcon.ludwig.ucl.ac.uk/MS - Fit.html
- PeptideSearch :
http://www.mann.embl - heidelberg.de/Services/PeptideSearch/FR_PeptideSearchForm.html
- ProFound :
http://chait - sgi.rockefeller.edu/cgi - bin/prot - id
4) Peptide mass fingerprinting and sequence tag data
- MS - Edman :
http://falcon.ludwig.ucl.ac.uk/msedman.html
- Mowse :
http://gserv1.dl.ac.uk/SEQNET/mowse.html
5) Peptide sequence tags(from MS/MS or MALDI - PSD)
- MS - Tag :
http://falcon.ludwig.ucl.ac.uk/mstag.html
- PepFrag :
http://chait - sgi.rockefeller.edu/cgi - bin/prot - id - frag
- PeptideSearch :
http://www.mann.embl - heidelberg.de/Services/PeptideSearch/FR_PeptidePatternForm.html
- Sequest :
http://thompson.mbt.washington.edu/sequest.html
6) Amino acid composition(with or without sequence tag)
- AACompIdent :
http://www.expasy.ch/ch2d/aacompi.html
- PropSearch :
http://www.embl - heidelberg.de/aaa.html
7) Amino acid composition, protein sequence tag, peptide mass fingerprinting
- MultiIdent :
http://www.expasy.ch/sprot/multiident.html
8) Programs that assist interpretation of analytical data
- Amino Acid Sequence :
http://chait - sgi.rockefeller.edu/cgi - bin/sequence
- Compute pI/MW :
http://www.expasy.ch/ch2d/pi_tool.html
- MS - Digest :
http://falcon.ludwig.ucl.ac.uk/msdigest.html
- PeptideMass :
http://www.expasy.ch/sprot/peptide - mass.html

Proteomics의 이용

기능성 측면에서는 세포내 단백질의 생리상태별(예, 암과 정상조직) 발현차이와 세포내 상태별 특정단백질의 정량적 차이 및 변형상태, 그리고 상호관계 등을 규정하는데 긴요하다. 활용성 측면에서는 각종 질환 표지자(disease marker)와 치료표적 단백질의 발굴과 동식물성장조절단백질의 규명, 생식기전 및 약물독성과 신의약 개발을 들 수 있다.[7-11)]

프로테오믹스 핵심기술의 최신 연구 동향

프로테옴 분석시료의 선택과 조제 분야

프로테옴 분석의 첫 단계는 샘플의 선택과 적절한 조제 방법인데 보다 많은 양의 샘플을 gel에 loading 할 수 있고 이들을 펼쳐 전개할 수 있는 방향으로 연구가 진행되고 있다. 최근에는 ' Proteomics toolbox '라는 새로운 개념이 도입되었다. 이것은 프로테옴 분석에 필요한 필수시약과 방법, 장비 등이 포함되어 이들을 하나의 set로 규정하는 것이다.

이와 함께 각종의 단백질을 잘 용액화 할 수 있는 detergent cocktail 제조 조건과 대용량분석(highthroughput)에 편리한 IPG strip을 개발 중에 있고, alkaline pI(>pH 10)를 갖는 단백질의 분리문제도 매우 중요하게 다루어지고 있다. 암조직과 같이 혼합된 임상조직에 대하여는 순수조직군을 얻기 위해 laser - capture microdissection 방법이 도입되어 실용화 단계에 이르렀고, SELDI(surface enhanced laser desorption/ionization) 단백질 칩과의 혼용도 시도되고 있다. 분자량이 매우 큰 단백질들(>500 kDa)에 대한 분리는 Gradipore™를 사용하여 미리 사전에 큰 단백질을 분리한 후 전기영동을 시도하는 경향이 있다.

이차전기영동(2DE) 분리기술의 최근 경향

시료의 전처리 기법의 발달로 2DE 전에 미리 cysteinyl peptide capture 기법을 사용하는 것이 시도되고 있다. 일종의 enrichment 기법으로 사전에 특정 펩타이드 군을 선택해내는 방법으로, 사전 정제된 단백질들은 LC - MS/MS 등에 의해 분석된다. 또한, 2DE에 대한 대체방법도 개발되고 있는데, 예를 들면 capillary electrophoresis와 고해상의 Fourier transform ion cyclotron resonance(FTICR) MS

와 함께 사용되어 단백질을 규명하는 것이다. 특히 방사능 동위원소를 단백질의 N-terminal이나 biotin 등에 표지시켜[12] 세포의 주기별이나 생리상태별로 특정단백질군의 정량적인 차이를 규명해나는 방법이 시도되고 있다. 이와 같은 방법으로 일종의 micro-chip을 affinity ligand로 사용하여 대량으로 빠르게 단백질을 분리하거나 microscale solution isoelectrofocusing(sol-IEF)을 이용, 세포나 임상샘플을 전처리하여 2DE 분석을 하게 하는 방법 등도 개발 중에 있다.

그러나 아직은 2DE를 대체할 만한 확실한 방법이 상용화되지 않은 것은 사실이다. 2DE 중에서는 IPG-DALT 시스템이 가장 보편적으로 쓰이는 HTS로 인정받고 있다. 이 시스템은 다양한 pH range별 isoelectric focusing strip 및 pre-made gel을 사용하여 재현성에서도 매우 뛰어난 기능을 보여준다. 특히 1 pH range strip(Zoom-in gels : pH 3.0~4.0, 3.5~4.5, 4.0~5.0 단위)의 사용으로 획기적인 단백질 분석능을 가지고 있다.

MS 분석기술의 최신동향

MALDI는 보편적으로 펩타이드 분석에 활용되고 있는데, 이 장비에 일종의 microsystem(flow-through piezo dispenser)을 붙여서 희소 농도의 peptide를 분석하는 방법도 개발 중이다. FTICR MS는 ESI와 MALDI ionization을 사용하여 100 kDa 이상의 단백질을 분석하는 데 활용되고 있는데,[13] 특히 ESI-FTICR의 경우 affinity proteomics라는 새로운 시도를 가능하게 하고 있다. MALDI QIT(quadrapole ion trap) reflection TOF MS는 당쇄화된 당단백질들의 구조적 특성을 분석하는데 활용되고 있다. 이 기기는 기존의 MALDI 가 지니는 장점들과 매우 높은 해상도, 분해능을 갖고 있으며 질소 laser를 사용하는 것이 특징이다.

2DE gel 상의 단백질 탐지법

지금까지 사용하던 미소샘플탐지로 쓰이는 silver 염색법은 각 단백질 샘플의 MALDI 분석에는 적합하지 않기 때문에 여러 가지 변형방법을 시도하는 노력이 있어 왔다. 그러나 변형된 silver 염색법에도 한계가 있기 때문에 최근에는 형광염색법이 도입되어 CYPRO™ 같은 시약이 쓰여 왔는데, 이 또한 고가이고 부수적인 형광 detector가 필요로 하는 등 아직도 해결되어야할 것이 많다. 물론 샘플 양이 많아 종래의 coomassie 염색법을 사용하여 각 spot의 단백질을 gel에서 분리시켜 MALDI 분석을 하는 것은 통상적으로 하고 있다. 이러한 점을 감안하여 최근에 발표된 2D-Fluorescence difference gel electrophoresis(2D-DIGE)가 개발되어 분석 software와 함께 폭넓게 쓰일 전망이다. 이 방법은 2DE 전에 다른 단백질들을 각기 다른 형광염색으로 표지한 후 전기영동을 하고 이를 DeCyder라는 Image 분석 시스템으로 분석하여 샘플내 단백질별로 보이는 형광 발색의 차이로 정성적, 정량적인 분석을 고속으로 할 수 있게 한다. 이외에도 MPD(Multiple Photon Detection) 방법을 사용, 매우 희소한 단백질의 변형을 탐지할 수 있도록 개발 중이다.

프로테오믹스의 자동화

프로테오믹스의 자동화 과정은 인간 게놈에 적용되었던 자동염기서열분석기나 PCR처럼 쉽지 않은 것이 사실이다. 대부분 단백질은 DNA와는 달리 branch나 변형구조를 지니고 있기 때문에 초고속 구조 분석이나 분리가 어렵기 때문이다. 따라서 프로테오믹스의 자동화과정은 가장 큰 bottleneck 중의 하나이다. 여기서 자동화라 함은 2DE나 그 대체분석 방법으로부터 MALDI 분석을 거쳐 최종적으로 단백질을 annotation하는 일련의 과정을 의미한다. 부분적으로 이루어지고 있는 자동화 과정을 간략하게 소개하면 다음과 같다. 2DE 분석에서 MALDI까지 가는 robotics 과정에는 소위 Molecular scanner라는 개념이 SIB(Swiss Institute of Bioinformatics)에 의해 도입되어 부분적으로 성공한 바가 있다. 즉, 2DE를 한 후 gel 자체를 통째로 trypsin 처리하고, 이를 다시 PVDF membrane에 blotting 시켜서 각 spot 별로 MALDI 분석을 대량으로 하여 데이터를 분류하고 각각의 단백질을 annotation 하는 것이다. 그러나 여기에는 아직도 해결해야 할 문제점들이 많다. 예를 들면, trypsin 처리과정의 재현성, blotting의 효율성, 각 spot 별 MALDI 결과의 체계적인 management 등이다.

Imaging 분석, Data display 및 Database

2DE를 통해 얻은 이미지들을 상호 비교할 때 사용되는 프로그램에는 기존의 Melanie III(GenBio), Imgage Master(Amersham Pharmacia), PDQuest(BioRad) 등이 가장 보편적으로 사용된다. 최근에는 이스라엘의 Compugen 사에서 비교적 단백질 변형구조 식별이나 gel 서열 분석에 유용한 프로그램을 개발 중이나 20 여개 lab에서 시험 중이고 시판되기까지는 상당한 시간이 소요될 것으로 예상된다. DB 구축은 proteome informatics의 가장 중요한 부분으로, 프로테오믹스 실험을 통해 산출되는 많은 데이터의 효율적인 보관과 분석에 필수불가결하다. 현재 대다수의 2D-DB에서는 Make2ddb라는 프로그램을 사용하여 각 gel image를 관리 하는 것이 보편적이다. SIB에 소개된 여러 가지 2DE의 DB가 대부분 이 프로그램을 사용하고 있다. 그렇지만 query가 단순하게 구성되어 있고, 실제 실

험실에서 산출되는 많은 종류의 관련정보를 효과적으로 처리하기 위해서는 부족한 부분이 많은 것이 사실이다.[14] 최근에 BioRad와 MicroMass 공동으로 WorksBase System™을 개발하여 샘플제조로부터, 각 샘플의 임상적 history, 데이터 저장 및 분석 등이 가능하게 되었다. 시장에 출시되어 일부 사용 중이나 고가(>USD200,000)인 점이 부담이 되고 있다. 이외에도 PRIME, GlycoDB와 SIB의 HPI(Human Proteomics Initiative), HAMAP(High quality automated microbial annotation of proteomes) 등이 현재 구축 중에 있다.

Proteomics의 한계와 전망

각 기술들이 가지는 장점에도 불구하고 이들 각개의 기술들 만으로서는 단백질 연구에서 기대하는 충분한 기술적 조건들을 만족시켜주지는 못하고 있으며, 여전히 미량 존재(low abundant protein) 단백질들에 대한 분석이 용이하지 못한 기술적 한계점들을 지니고 있다. 또한 현재까지 단백질의 분리능력에 있어서는 가장 탁월한 결과를 보여주는 이차원 전기영동법을 이용한 프로테오믹스 기술은 복잡하고 긴 분석과정 때문에 완전 자동화가 어려운 상황이며 high throughput study를 위해서는 기업형의 설비가 갖추어 져야만 한다.

프로테오믹스 기술이 안고 있는 이러한 한계에도 불구하고, 이 기술을 도입해서 얻을 수 있는 정보의 양은 현재까지로는 지난 십수 년간 이어온 연구의 결과들의 많은 양들을 짧은 시간 내에 얻을 수 있음을 보아 왔고, 이러한 정보의 획득 능력은 단백질 연구의 새로운 돌파구가 찾아지지 않은 한 당분간은 지속될 것이라고 전망되고 있다.

맺 음 말

정보화시대에서의 생물 및 의학 관련 연구는 매우 빠른 속도로 발전하고 있으며, 그 방식 또한 체계적이고 유기적으로 진행되고 있다. 이러한 경향에서 뒤쳐지지 않고 경쟁력을 가지기 위해서는 생물학 연구 관련 새로운 기술들을 빠르게 도입하여야 하며, 이를 의학에 접목하여 이용함으로써 질병의 진단 및 치료에 필요한 도구로써 사용할 수 있어야 하겠다.

REFERENCES

1) Wilkins MR, Williams KL, Appel RK, Hochstrasser DF. *Proteome research: New frontiers in functional genomics. Berlin: Springer;2000.*
2) Blackstock WP, Weir MP. *Proteomics: Quantitative and physical mapping of cellular proteins. Trends Biotechnol 1999;17:121-7.*
3) Link AJ. *2D Proteome analysis protocol in methods in molecular biology. New Jersey: Humana Press;1999.*
4) *Proteomics: A trends guide. Elsevier Science;2000.*
5) Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, *et al. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 2000;21:1037-53.*
6) O'Farrell PH. *High resolution two-dimensional electrophoresis of proteins. J Biol Chem 1975;250:4007-21.*
7) Brewis IA. *Proteomics in reproductive research: The potential importance of proteomics to research in reproduction. Hum Reprod 1999;14:2927-9.*
8) Edvardsson U, Alexandersson M, Brockenhuus VL, Nystrom AC, Ljung B, Nilsson F, *et al. A proteome analysis of livers from obese (ob/ob) mice treated with the peroxisome proliferator WY14,643. Electrophoresis 1999;20:935-42.*
9) Jungblut PR, Zimny-Arndt U, Zeindl-Eberhart E, Stulik J, Koupilova K, Pleissner KP, *et al. Proteomics in human disease: Cancer, heart and infectious diseases. Electrophoresis 1999;20:2100-10.*
10) Mollenkopf HJ, Jungblut PR, Raupach B, Mattow J, Lamer S, Zimny-Arndt U, *et al. A dynamic two-dimensional polyacrylamide gel electrophoresis database: The mycobacterial proteome via internet. Electrophoresis 1999;20:2172-80.*
11) Myers TG, Dietz EC, Anderson NL, Khairallah EA, Cohen SD, Nelson SD. *A comparative study of mouse liver proteins arylated by reactive metabolites of acetaminophen and its nonhepatotoxic regioisomer, 3'- hydroxyacetanilide. Chem Res Toxicol 1995;8:403-13.*
12) Mann M. *Quantitative proteomics? Nat Biotechnol 1999;17:954-5.*
13) Despeyroux D, Walker N, Pearce M, Fisher M, McDonnell M, Bailey SC, *et al. Characterization of ricin heterogeneity by electrospray mass spectrometry, capillary electrophoresis, and resonant mirror. Anal Biochem 2000;279:23-36.*
14) Baxevanis AD. *The molecular biology database collection: An online compilation of relevant database resources. Nucleic Acids Res 2000;28:1-7.*