เอกสารประกอบการสอน

วิชา 104 852 Principles in Molecular Biology Techniques

ระดับบัณฑิตศึกษา

โดย

ผศ. ทนพญ. ดร. วีไลรัตน์ ลื้อนันต์ศักดิ์ศิริ

สำนักวิชาวิทยาศาสตร์

มหาวิทยาลัยเทคโนโลยีสุรนารี

พ.ศ. 2550

Table of Contents

Part I

Principles of Techniques in DNA Study

Topic 1. DNA Extraction and Purification Technique

1. DNA Extraction, Purification and Concentration of DNA from Aqueous Solution

A. phenol extraction and ethanol precipitation of DNA
appropriate for purification of DNA from small volumes(<0.4 ml) at concentrations < 1mg/ml

DNA to be purified
25:24:1 (V/V/V) phenol/chloroform/isoamyl alcohol
ether extraction
3M sodium acetate, pH 5.2
100% ethanol, ice cold
70% ethanol, room temperature or ice cold (usually)
TE buffer, pH 8.0
Speedvac evaporator

-precipitation of DNA using isopropanol (equal volume of DNA and isopropanol)

B. purification of DNA using glass beads
: for rapid and efficient purification of DNA from contamination protein, RNA, or organic solvents (available as commercial kits e.g. Glas-Pac/National Scientific Supply, GeneClean, Bio101, Qiaex Gel Extraction Kit/Qiagen)

6M sodium iodide (NaI) solution
50 to 200 μl DNA
wash solution
TE buffer, pH 8.0
Glass beads suspension

2. Preparation of Genomic DNA from Mammalian Tissue

Liquid nitrogen
Digestion buffer
PBS, ice-cold
25:24:1 (V/V/V) phenol/chloroform/isoamyl alcohol
7.5 M ammonium acatate
100% and 70% ethanol
TE Buffer, pH 8.0
Sorvall H1000B rotor or equivalent

A. Beginning with whole tissue : freeze in N2 (l)/ grind with prechilled mortar and pestle or crush with hammer to fine powder

B. Beginning with tissue culture cells: trypsinization adherent cell

3. **Preparation of genomic DNA from Plant Tissue**
   CsCl centrifugation

4. **Preparation of genomic DNA from Bacteria**

   TE Buffer
   10% (w/v) sodium dodecyl sulfate (SDS)
   20 mg/ml proteinase K (stored in small single-use aliquots at –20°C)
   5 M NaCl
   CTAB/NaCl solution
   25:24:1 (V/V/V) phenol/chloroform/isoamyl alcohol
   24:1 (V/V) chloroform/isoamyl alcohol
   isopropanol
   70% and 100% ethanol

DNA Isolation from Rhizobium by Phenol Chloroform Method

It is very difficult to isolate Rhizobium DNA due to the gum production by the organism. Hence this protocol has been designed for efficient isolation of DNA from the organism for further gene amplification.

Procedure

1. Grow *Rhizobium* cells in 5 ml of YEMB at 200 rpm till the O.D (600 nm) reaches 0.6-0.8.
2. Pellet the cells by centrifugation at 10,000 rpm for 15 mins.
3. Wash the pellet with TE buffer (10T/1E).
4. Dissolve the pellet in 400 µl TE.
5. Add 40 µl 10% SDS and 5µl Proteinase K (20mg/ml).
6. Incubate at 56 ° C for 45 mins.
7. Add 400 µl of tris saturated phenol (pH 8.0).
8. Centrifuge at 10,000 rpm for 10 mins.
9. Take the supernatant and add 200 µl of tris saturated phenol and 200 µl of chloroform: isoamyl alcohol (24:1).
10. Again take the supernatant and repeat the above step.
11. Take the supernatant and add 400 µl of chloroform: isoamyl alcohol (24:1).
12. Repeat the above step.

13. Take the supernatant and add 0.1 volume of 3M sodium acetate and 2 volume of chilled absolute ethanol.
14. Incubate at -20 ° C for 2 hrs.
15. Centrifuge at 15,000 rpm for 30 mins in a cooling centrifuge at 10-15 ° C.
16. Decant the supernatant and wash the pellet with 70% ethanol by centrifugation at 10,000 rpm for 15 mins.
17. Decant the supernatant and dry the pellet at room temperature.
18. Suspend the pellet in 50 µl of TE.

Reagents

Composition of YEMB (Yeast Extract Mannitol Broth)	
Mannitol	10.0 g
K2HPO4	0.5 g
Yeast extract	0.5 g
MgSO4.7H2O	0.2 g
NaCl	0.1 g
Distilled water	1 litre
pH	6.8

DNA From Whole Blood for PCR

(from Higuchi, R. (1989) Amplifications 2: 1-3)

1. Obtain 65-100 µl of blood by retro-orbital bleed with a heparinized microcapillary tube. Expel blood immediately into a 1.5 ml microfuge tube containing 20 µl of 10 mM EDTA. Mix immediately to prevent clot formation. Store on ice until processing.
2. Add 200 µl Lysis Buffer to each tube and vortex to suspend evenly.
3. Microfuge 25 seconds at 16000xg to pellet nuclei.
4. Remove and discard supernatant and repeat steps 2-4 two more times, or until no hemoglobin remains.
5. Resuspend nuclear pellet in 100 µl PBND with 60 µg/ml proteinase K and incubate at 55 C for 60 minutes (or overnight, if convenient).
6. Heat samples to 97 C for 10 minutes to inactivate proteinase K.
7. Add 1-5 µl of DNA solution for a 25 µl PCR reaction.

Reagents:

1) Lysis Buffer
   0.32 M Sucrose
   10mM Tris-HCl (pH 7.5)
   5 mM MgCl2
   1% v/v Triton X-100
2) PBND (PCR Buffer with Nonionic Detergents)*
   50 mM KCl
   10 mM Tris-HCl (pH 8.3)
   2.5 mM MgCl2
   0.1 mg/ml gelatin
   0.45% (v/v) Nonidet P40
   0.45% (v/v) Tween 20

   Autoclave to sterilize and dissolve gelatin
   Store frozen

*Add proteinase K (60 μg/ml) immediately prior to use)

Typical 25 μl PCR reaction:
1-5 μl DNA
2.5 μl 10x Perkin Elmer buffer, 1.5 mM MgCl2 (final)
2 μl 2.5 mM dNTP mixture (2.5 mM each dNTP, 200 μM final)
0.5 μl 20 μM forward primer (0.4 μM final)
0.5 μl 20 μM reverse primer (0.4 μM final)
0.1 μl Taq DNA polymerase (can decrease to 0.05 μl)
dH2O to 25 μl

DNA Purification from Agarose Gel

1. Separate DNA fragments in an agarose gel cast with 0.5 mg/mL Ethidium bromide. Locate bands with a hand-held long-wave UV lamp.

2. Slice the gel with a razor blade above and below the bands of interest. Insert the membrane into the gel at each slice, avoiding air bubbles between the gel and the membrane. The membranes above the bands of interest prevent contamination with higher molecular weight DNA.

3. Electrophorese at high voltage to run the DNA onto the membrane.

4. Remove the membrane with the bands of interest and verify DNA with long-wave UV. Discard membranes placed above the bands of interest.

5. Rinse the membranes in Low Salt Buffer to remove agarose.

6. Place membrane in 1.5 mL tube. Add 150 ml High Salt Buffer.

7. Incubate at 70 C for 20 min. Remove supernatant and place in a fresh tube.

8. Add 150 ml High Salt Buffer to the membrane. Incubate at 70 C for 10 min.

9. Combine the supernatant with that removed in Step (10). Add 300 ml Buffer Saturated Phenol.

10. Mix well by vortexing. Centrifuge for 15 min at 4 C. Remove the aqueous phase being careful NOT to remove ANY of the interface, which contains fibers of DEAE membrane.

11. Add 300 ml 24:1 Chloroform:Isoamyl alcohol. Mix well by shaking.

12. Centrifuge 2 min at 4 C. Remove the aqueous phase. Add 5 ml 1% linear polyacrylamide carrier and mix well.

13. Add 700 ml 95% ethanol. Incubate at -20 C for 2 hr.

14. Centrifuge for 15 min at 4 C. Discard supernatant. Add 500 ml 70% ethanol. Mix well to dislodge the pellet.

15. Centrifuge for 2 min. Discard supernatant. Air dry.

16. Resuspend DNA in 50 mL TE.

DNA Preparation from Adherent Cells

Reagents

Chloroform
Mallinckrodt, cat. 4440
EDTA, 0.5 M
Ethanol, 100%
Ethanol, 70%
Isoamyl Alcohol
Sigma, cat. I-3643
Phenol
Gibco BRL, Cat. 15513-039
Phosphate Buffered Saline (PBS), 10X and 1X
Gibco BRL, Cat. 10010-023
Proteinase K
Gibco BRL, Cat. 24568-2
Sodium acetate, 3 M, pH 5.2
Sodium dodecyl sulfate (SDS), 10%
Tris-HCl, 1 M, pH 8.0
TAE buffer (Tris acetate/disodium EDTA), 1X
Bio Whittaker, Cat. 16-011V
Trypsin
Gibco BRL, Cat. 25200-056
Distilled Water
Gibco BRL, Cat. 15230-170

Preparation

DNA buffer
1M Tris-HCl, 1 M, pH 8.0 100 ml
0.5 M EDTA 100 ml
dH2O water 300 ml
Chloroform/Isoamyl alcohol 24:1
Chloroform 24 ml
Isoamyl alcohol 1 ml

Procedure
1. Use trypsin or cell scraper to remove cells from tissue culture flask (T-75). Centrifuge
cultured cells for 10 min at 10°C (1200 rpm). Remove supernatant and re-suspend cell pellet in 1X PBS and wash twice with 10 ml 1X PBS, centrifuging between washes.
2. Resuspend pellet in 10 ml DNA buffer. Centrifuge cells for 10 min at 10 °C (1200 rpm). Remove supernatant.
3. Add 3 ml DNA-buffer, re-suspend the pellet, add 125 ml Proteinase K (10 mg/ml) and
400 ml 10% SDS; shake gently and incubate overnight at 45°C.
4. Add 3.6 ml of phenol, shake by hand for 10 minutes (RT); centrifuge for 10 min at 10°C (3000 rpm).
5. Transfer the supernatant into a new tube (15 ml); measure the volume. Add 1.8 ml phenol and 1.8 ml chloroform/isoamylalcohol (24:1) or a total amount equal to the volume of the supernatant. Shake by hand for 10 min (RT); centrifuge for 10 min at 10°C (3000 rpm).
6. Transfer the supernatant into a new tube (15 ml); measure the volume. Add 3.6 ml chloroform/isoamylalcohol (24:1) or an amount equal to the volume of the supernatant. Shake by hand for10 min (RT); centrifuge for 10 min at 10°C (3000 rpm).
7. Transfer the supernatant into new tube, measure the volume. Add 1/10 volume 3 M sodium acetate (pH 5.2) and 3 x the volume 100% isopropanol (2-propanol); shake gently until the DNA is precipitated.
8. Use a sterile glass pipette to transfer the precipitated DNA into a tube with 30 ml of 70% ethanol tube. Place on inverting rack and invert for 2 hr to thoroughly rinse. Transfer DNA into a sterile eppendorf tube.
9. Centrifuge for 20 min at 14,000 rpm. Dry pellet in a SpeedVac for 5 min. Dissolve the DNA in 300-500 µl sterile water and place in an eppendorf thermomixer shaker overnight at 37°C.
10. Measure the DNA concentration and run 1-5 µl (approximately 200 ng) for gel electrophoresis on agarose gel (1%) in 1X TAE buffer.

DNA Preparation from Fresh/Frozen Tissue

Reagents

Chloroform
EDTA, 0.5 M
Ethanol, absolute
Isoamyl alcohol
Sigma, Cat. I-3643
Phenol
Phosphate Buffered Saline (PBS), 1X
Proteinase K
EM Science, Gibbstown, WV Cat. 24568-2 (100 mg)
RNase A
Boehringer Mannheim, Cat. 109 169
Sodium dodecyl sulfate (SDS) solution, 10%
Digene Diagnostics, Beltsville, MD, Cat. 3400-1016

Preparation
DNA buffer (Tris-EDTA)
1 M Tris pH 8.0 20 ml
0.5 M EDTA 20 ml
Sterile water 100 ml
Proteinase K (10mg/ml)
Dissolve 100 mg Proteinase K in 10 ml TE for 30 min at room temperature (RT)
Aliquot and store at –20°C
RNase A (20 mg/ml)
Dissolve 200 mg RNase A in 10 ml sterile water, boil for 15 min, and cool to RT.
Aliquot and store at –20°C

Procedure
1. Put 60-80 mg of tissue in a petri dish with culture media and divide the tissue into two pieces.
2. Put the tissue into two sterile 15 ml tubes and centrifuge for 2 min at 4°C at 1500 rpm.
3. Remove the supernatant, and wash twice with 1 ml 1X PBS or DNA-buffer. (It is possible to store the pellet at -80°C; in that case, add 1 ml 1X PBS and resuspend the pellet. Use a cryo-tube and centrifuge at 1500 rpm for 2 min at 4°C. Remove the supernatant, and freeze the pellet.)
4. Remove supernatant and resuspend the pellet in 2.06 ml DNA-buffer.
5. Add 100 µl proteinase K (10 mg/ml) and 240 µl 10% SDS, shake gently, and incubate overnight at 45°C in a waterbath.
6. If there are still some tissue pieces visible, add proteinase K again, shake gently, and incubate for another 5 hr at 45°C.
7. Add 2.4 ml of phenol, shake by hand for 5-10 min, and centrifuge at 3000 rpm for 5 min at 10°C.

8. Pipette the supernatant into a new tube, add 1.2 ml phenol, and 1.2 ml chloroform/isoamyl alcohol (24:1); shake by hand for 5-10 min, and centrifuge at 3000 rpm for 5 min at 10°C.
9. Pipette the supernatant into a new tube, add 2.4 ml chloroform/isoamyl alcohol (24:1), shake by hand for 5-10 min, and centrifuge at 3000 rpm for 5 min at 10°C.
10. Pipette the supernatant into a new tube, add 25 µl 3 M sodium acetate (pH 5.2) and 5 ml ethanol, shake gently until the DNA precipitates.
11. Take a glass pipette, heat it over a gas burner, and bend the end to a hook. Fish the DNA thread out of the solution using the hook and transfer DNA to a new tube.
12. Wash the DNA in 70% ethanol and dry it in the speed vac.
13. Dissolve the DNA in 0.5-1 ml sterile water overnight (or longer if necessary) at 4°C on a rotating shaker.
3
14. Measure the DNA concentration in a spectrophotometer and run 200 ng on a 1% agarose gel.

Tissue (mg)	5	10	15	20	40	60	80	100
Volume in µl								
Total	400	800	1200	1800	3200	4800	6400	8000
DNA buffer	360	680	1020	1360	2720	4080	5440	6800
Proteinase	20	40	60	80	160	240	320	400
10% SDS	40	80	120	160	320	480	640	800

ES CELL DNA EXTRACTION: TUBE METHOD

Protocol for extracting DNA from ES Cells, starting from the 96-well plate but processing in an eppendorf tube to recover more of the DNA.

NOTE- THIS TAKES A LOT OF TIME if you do the whole plate this way!

Lysis buffer: (100 ml Recipe)
10 mM Tris-HCl pH7.5 - 0.5 ml of 2M
10 mM EDTA - 2 ml of 0.5M
10 mM NaCl - 0.2 ml of 5M
0.5% (w/v) Sarkosyl - 0.5gm (N-Lauroylsarcosine, Sigma # L-9150)
1mg/ml Proteinase K - add fresh each time (stored in freezer).

1. Add 250µl Lysis Buffer (with proteinase K added) to each well.

2. Incubate the plate in a sealed humid container at 60°ree;C for 1 hour. (Meanwhile, label eppy tubes!)

3. Transfer the contents of each well into a separate 1.5ml eppendorf tube (µ-cent. tube).

4. Continue lysing the cells for 2 more hours. (60°ree;C). (Float tubes in waterbath- don't need tupperware).

5. Add an equal volume (250 µl) of phenol:chloroform (phenol: chloroform: isoamylalcohol, 25:24:1).
(The pH of the phenol:chloroform needs to be ~7.8 - 8.0, or else the DNA will partition into the organic phase).

6. Mix the contents of the tube until an emulsion forms. (Shake by hand for ~1 min, (or vortex ~2sec),because vortexing shears long (genomic) DNA).

7. Centrifuge 5-10 min in a microfuge at TRm. (The phases should be well-separated).

8. Transfer (pipet) the aqueous phase (the top phase) to a fresh tube. (Toss the interface/organic phase-- chemical waste in hood).

9. Repeat steps 5-8 until no protein is visible at the interface of the aqueous and organic phases.

10. Add an equal volume (250 µl) of chloroform and repeat steps 6-8.

11. Optional: Repeat chloroform extraction. (The chloroform extraction removes the traces of phenol).

12. Add 1/25th volume of 5M NaCl (10µl of 5M NaCl for 250 µl), for a final concentration of 0.2M NaCl. Mix well. (use NaCl rather than other salts because of detergent in lysis buffer).

13. Add exactly 2 volumes of ice-cold ethanol and again mix the solution well.

14. Store the ethanolic solution on ice for 30 min to precipitate the DNA. (can be stored indefinitely at 0°ree;C or at -20°ree;C).

15. Centrifuge at 12,000xg for 10 min, 0°ree;C. (Try 4°ree;C, microfuge at max (16,000), 10 min). (Remember to point tube hinges out, to locate pellet even if it's invisible).

16. Suck off the supernatant (dispo pipet tip, with vacuum flask). Do not disturb the pellet. Vacuum droplets from walls of tube as well.

17. Half fill the tube with 70% EtOH and re-spin for 2 minutes at 4°ree;C.

18. Suck off supernatant as in step 16.

19. Store the open tube on the bench at TRm until the last traces of fluid have evaporated. (i.e.: Air dry).

20. Dissolve the DNA in 30µl (or 50 µl, and use half of sample) of 10mM Tris pH 8.5 (or TE). (pipet around the walls of the tube to dissolve the DNA off the walls of the tube). Let dissolve for at least an hour at 37°ree;C.

21. Restriction digest:
- DNA in Tris pH8.5, from above protocol
- restriction enzyme
- restriction enzyme buffer
- X (BSA if required by enzyme)
- spermidine (to a final conc. of 1mM; stored in freezer)

- (also add:) DTT to a final conc. of 1mM
Final volume = 40µl or less.
Digest overnight at 37°ree;C.

Next: Run gels; Southern blot and hybridize.

Mouse and Human Genomic DNA Extraction

Treat human blood and supernatants as potentially infectious. Discard supernatants by an appropriate method, eg into bromochlor solution.

DNA from blood

1. Collect 10ml blood into a heparin or EDTA tube.
2. Transfer to a 50ml falcon tube and freeze overnight at -20°C.
3. Allow to thaw in iced water for several hours.
4. Add 40ml of triton/sucrose lysis buffer to blood and invert to mix. Spin samples at 3000rpm for 15 minutes. Discard supernatant.
5. Add 20ml of triton/sucrose lysis buffer and resuspend white cell pellet with a blue tip. Spin samples at 3000rpm for 10 minutes. Discard supernatant.
6. Add 3ml salting out/lysis buffer to the white cell pellet. Resuspend using a blue tip. Add 200µl 20% SDS and 300µl 5mg/ml proteinase K. Incubate samples overnight at 50°C in a gently shaking incubator. If sample has not been completely digested overnight, add 50µl 5mg/ml proteinase K and continue incubation for several more hours.
7. Add 1.5ml 6M NaCl and shake vigorously for 15 seconds. Spin sample at 3000rpm for 15 minutes.
8. Transfer supernatant to a new tube, and add 10ml absolute ethanol. Allow the DNA to come out of solution. Hook out DNA and transfer to an ependorf tube. Spin briefly and remove ethanol. Add 400µl TE and allow DNA to resuspend overnight at 4°C.
9. Extract once with phenol/chloroform and once with chloroform. Ethanol precipitate and rinse pellet with 70% ethanol. Allow pellet to resuspend overnight at 4°C.

Triton/sucrose lysis buffer1 litre:

1M $MgCl_2$	5ml
1M Tris-Cl pH 7.5	10ml
Triton X-100	10ml
sucrose	109.54g

Salting out/lysis buffer100mls:

1M Tris-Cl pH 7.5	1ml
5M NaCl	8mls
500mM EDTA pH 8	400µl

Mouse:

1. Dissect one lobe of liver tissue, and freeze in liquid nitrogen.
2. Grind tissue into a fine powder under liquid nitrogen.
3. Transfer tissue to a cold 15 ml blue top tube, add 5 ml of salting out/lysis buffer (10 mM Tris-Cl pH 7.5, 400 mM NaCl, 2 mM EDTA pH 8) and resuspend tissue with a blue tip.
4. Add 200 µl 20% SDS and 300 µl 5 mg/ml proteinase K. Incubate samples for
5. several hours at 50°C, until tissue had been completely digested.

Extraction and precipitation:

1. Add 5 ml phenol equilibrated with 500 mM Tris pH 8, place on slowly rotating wheel for 15 minutes. Centrifuge 3500 rpm for 10 minutes. Remove aqueous phase to another tube.
2. Add 5 ml phenol/chloroform equilibrated with 500 mM Tris pH 8, place on slowly rotating wheel for 15 minutes. Centrifuge 3500 rpm for 10 minutes. Remove aqueous phase to another tube.
3. Add 5 ml chloroform equilibrated with TE, place on slowly rotating wheel for 15 minutes. Centrifuge 3500 rpm for 10 minutes. Remove aqueous phase to another tube.
4. Add 0.5 vol isopropanol and allow DNA to come out of solution. Gently swirl the tube for a few minutes. (**At this stage, tube can be placed at -20°C overnight.)
5. Hook out the DNA with a pasteur pipette hook, and allow excess solution to drain from the DNA by holding the blob to the side of the tube.

6. Place hook with DNA on it into a 15 ml blue top tube with 1 ml TE. Break off pasteur pipette and leave hook with DNA in tube. Place on very gently rocking platform for several hours until DNA has completely gone into solution.

7. Quantitate and resuspend DNA at a final concentration of 500 μg./ml.

Measurement of DNA concentration : O.D.260 nm.
Purity O.D.260/280 nm.

Electrophoresis of DNA

1. Agarose Gel Electrophoresis

Electrophoresis buffer (TAE or TBE)
Ethidium bromide solution
Agarose, electrophoresis-grade
10X loading buffer
DNA molecular weight markers
Horizontal gel electrophoresis apparatus
Gel casting platform
Gel combs (slot formers)
DC power supply

Appropriate Agarose Concentration for Separating DNA fragments of Various Sizes

Agarose (%)	Effective range of resolution of linear DNA fragments (kb)
1.5	0.2-3
1.2	0.4-7
1.0	0.5-10
0.7	0.8-12
0.5	1-30

2. Nondenaturing Polyacylamide Gel Electrophoresis

Polyacrylamide gels are formed by the polymerization of acrylamide monomer and N, N-methylene-bis-acrylamide (Bis). The reaction is initiated by adding ammonium persulphate and accelerated by TEMED.

Concentration of Acrylamide Giving Maximum Resolution of DNA Fragment

Acrylamide (%)	Size fragments separated (bp) marker	Migration of bromphenol blue
3.5	100 to 1000	100
5.0	100 to 500	65
8.0	60 to 400	4
12.0	50 to 200	20
20.0	5 to 100	12

3. Pulsed-Field Gel Electrophoresis (PFGE)

3.1 Verticle Single Inhomogenous Electric Field PFGE (Figure A)and Verticle Double Inhomogenous Electric Field PFGE (Figure B)

3.2 Orthigonal-Field-Alteration Gel Electrophoresis (OFAGE) (Figure C)

3.3 Transverse Alternating Field Electrophoresis (TAFE) (Figure D)

3.4 Field Inversion Gel Electrophoresis (FIGE) (Figure E)

3.5 Coutour-Clamped Homogeneous Electric Field (CHEF) (Figure F)

3.6 Crossed Field Gel Electrophoresis (CFGE) (Figure G)

3.7 Pulsed Homogeneous Orthogonal Field Gel Electrophoresis (PHOGE) (Figure H)

Factors:

a. Temperature
b. Agarose gel concentration
c. Electric field strength
d. Pulse time
e. Reorientation angle

Advantages:

a. Electrophoretic karyotyping
b. Macrorestriction mapping
c. Use in cloning

Ethidium bromide Staining : usually use 0.5 ng/ml

limitation of detection is 1-2 ng
UV wavelength 260, 300, or 360
emitted light 590 nm

Isolation and Purification of DNA from Agarose Gels

e.g. Spin columns : Quantum Prep Freeze-squeeze DNA gel extraction spin columns (Biorad)

Practical requirements for DNA work

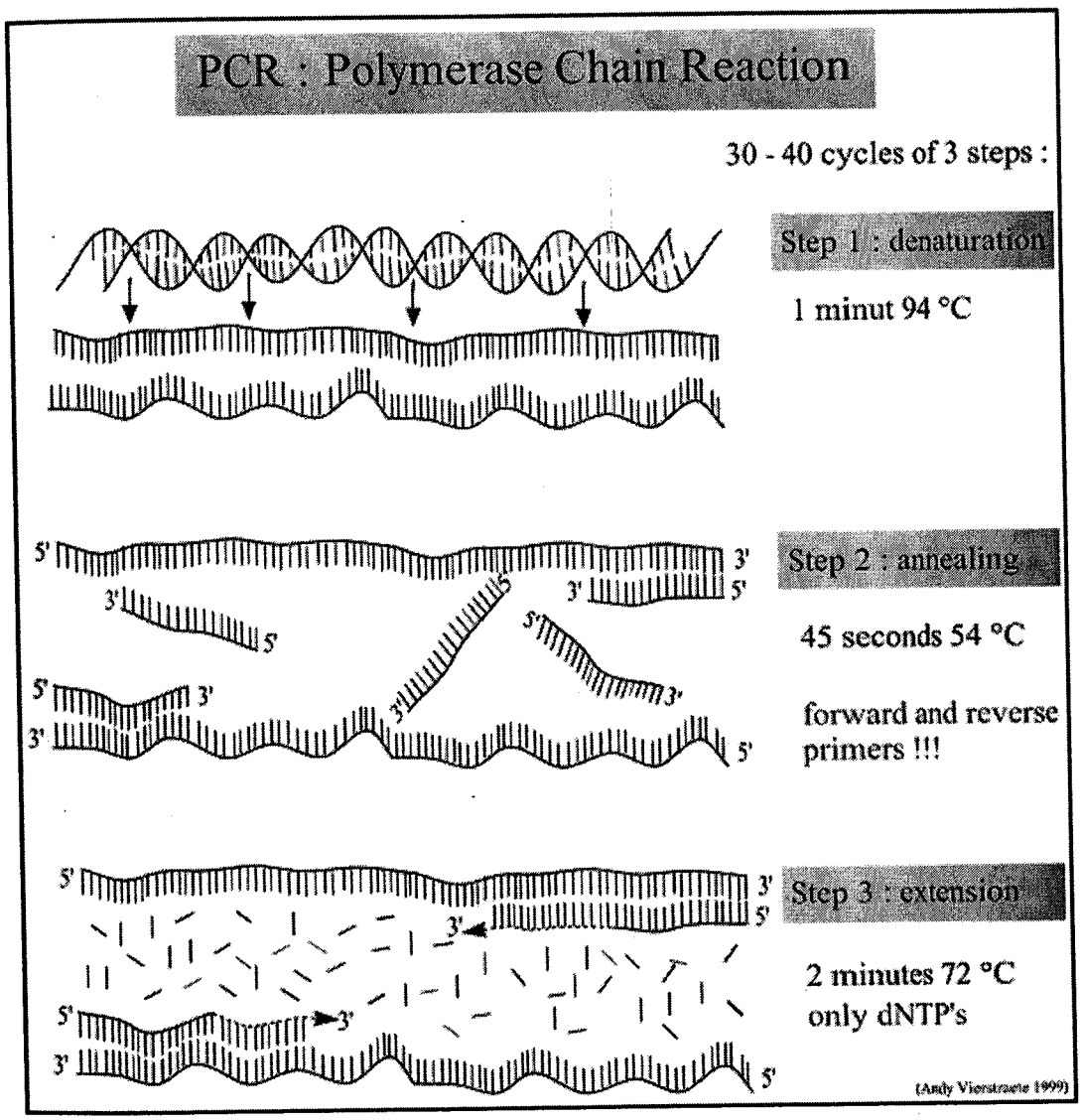
1. learn how to pipette down to 1 μl accurately and reproducibly;
2. develop a steady hand for loading samples on to gels and other manipulations: give up caffeine if it makes you shake
3. do not allow stocks of enzymes to warm up to room temperature: keep in the freezer and take out only for as long as it take to remove an aliquot;
4. autoclave plasticware (pipette tips, microfuge tubes, etc.) and keep sterile before use; do not handle pipette tips;
5. always use double-distilled water; make sure the glassware cleaning programme includes a fina; rinse in double –distilled water;
6. get in the habit of wearing disposable **gloves at all times**.

Topic 2. Polymerase Chain Reaction (PCR) Technique

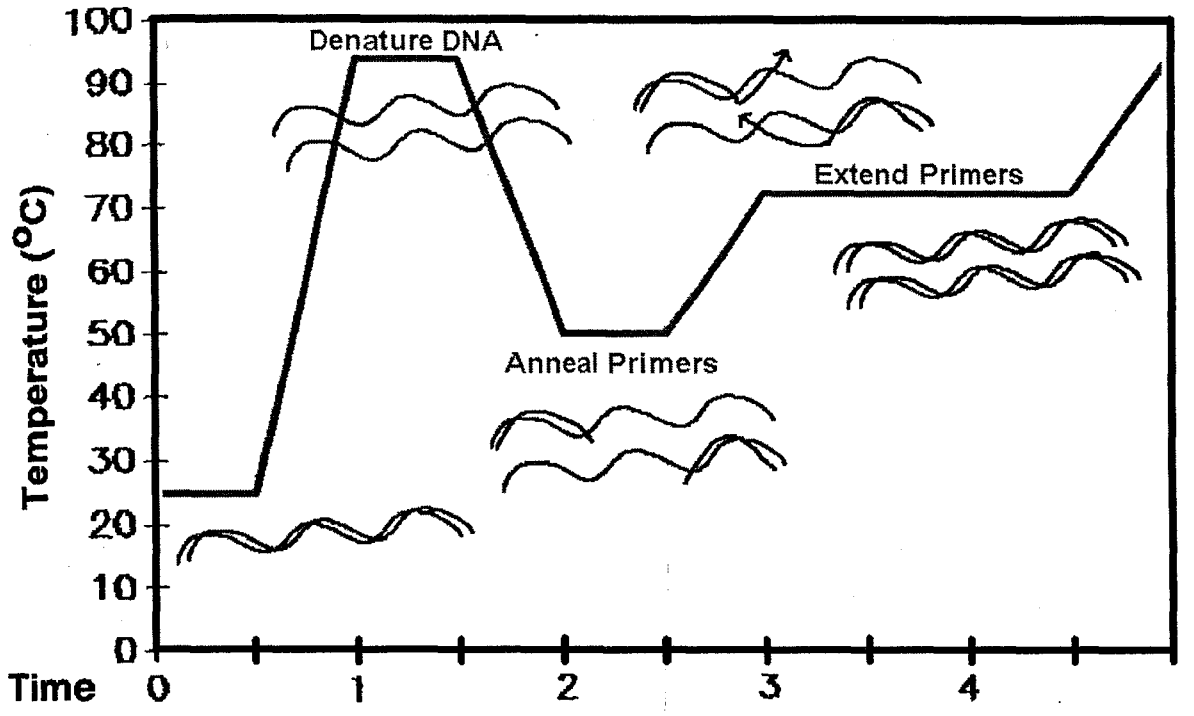
The polymerase chain reaction (PCR) is a rapid procedure for in vitro enzymatic amphification of a specific segment of DNA.

Basic components : 1. DNA to be amplified (genomic, cDNA or plasmid)
2. two single stranded oligonucleotide primers
3. protein component (a DNA polymerase)
4. dNTPs, Buffer and Salts

Principle: VDO



Source: http://users.ugent.be/~avierstr/principles/pcr.html



Source: http://www.mun.ca/biology/scarr/PCR_simplified.html

PCR begins with a mixture containing a **dsDNA template**, a pair of short **ssDNA oligonucleotide primers**, a pool of the four **dNTPs**, and a **heat-resistant DNA polymerase**, ***Taq* Enzyme**. The reaction is carried out in a computer-regulated heating block, a **thermal cycler**, which permits rapid, controlled heating & cooling. The primers are chosen so that they are base-complementary to opposite ends of either strand of a short stretch of DNA containing the gene region of interest: **PCR** thus requires some prior knowledge of the gene.

The reaction is first heated to **95°C** to melt (**denature**) the **dsDNA** into separate starnds. The reaction is then cooled to **~50°C**, at which temperature the **primers** will find base-complementary regions in the **ssDNA**, to which they will stick (**anneal**). The reaction is finally heated to **72°C**, at which temperature the ***Taq*** enzyme replicates the primed **ssDNA (extension)**. At the end of one cycle, the region between the two primers has been copied once, producing two copies of the original gene region.

Because a heat-resistant polymerase is used, the reaction can be repeated continuously without addition of more enzyme. Each cycle *doubles* the copy number of the amplified gene: ten cycles produces 2 → 4 → 8 → 16 → 32 → 64 → 128 → 256 → 512 → 1,024 (2^{10}) copies. Thus, **30 cycles** yields a (2^{10x3}) = **10^9-fold amplification.** This produces a sufficient quantity of the gene region of interest for direct analysis, for example by **DNA sequencing**.

General protocol : 1. Denature 1 min at 94 °C
2. Anneal 1 min at 60 °C
3. Extend 1 min at 72 °C
30 Cycles

Mater mixes for optimizing reaction components

10XPCR buffer (KCl, Tris-HCl pH 8.4 at 20°C, $MgCl_2$, Glycerol)
forward primer and reverse primers
DNA template, dNTPs, $MgCl_2$, *Taq* polymerase, DMSO, Glycerol, H_2O

Specificity and fidelity are the major concerning Parameters :

Sample preparation :
Shearing (especially for genomic DNA template), RNA contamination

Primers design:
- 50-60 % G+C content
- Tm 55-80 °C (Add up 2°C for a T or A, 4°C for a C or G)
- Not complementary at the 3′ ends: primer-dimer formation

- ACTGTCCCGTATGCAATTGGCATGC
TACGTATGTTCCGGCATGCAT

- Can add other useful sequence (e.g. multiple cloning sites, anchors, etc) at the 5 ends without effects on the PCR

Primer concentrations:
0.1-0.5 μM in reaction
too much: nonspecific products

Enzyme selection and concentration

- recommended range for *Taq* DNA polymerase (Perkin-Elmer) is 1-2.5 units/100 μl reaction
- Too high : nonspecific background
- Too low : an insufficient amount of desired product is made
- Different Taq DNA polymerase from different company may behave differently because of different formulations, assay condition, and or unit definitions
- Thermo-resistant Taq DNA polymerase : higher fedelity, less problem but more expensive e.g. pfu *Taq* polymerase, TaqGold polymerase, Platinum *Taq* polymerase, etc.

dNTP

- use 20-200 μM concentration of each
- the lower the concentration, the better the fidelity
- same concentration for all three nucleotides
- 20 μM will make 2.5 μg of DNA

$MgCl_2$ concentration

- very often can have major impact on PCR reaction if optimized
- If standard conditions poor, try change concentration by 0.5 mM up to 2.5

Denaturation time and temperature

- typical time = 30 second at 95 °C or 97 °C for 15 seconds
- too little denaturation- very poor results
- too much- decrease Taq polymerase
  at 92.5 °C T1/2 > 2 hr.
  at 95 °C, 40 minutes
  at 97.5 °C, 5 minutes

Primer annealing

- 5°C below Tm
- the higher the priming temp, the less mispriming-better specificity

Primer extension

- around 72 °C are typical
- 30-100 NT per second
- 1 minute extension allows up to 2 kb synthesis

Cycle number

- too many cycle, PCR against target will stop, other products come up

Negative (can use water blank control) and Positive controls

RT-PCR : Reverase Transcriptase Polymerase Chain Reaction
Amplification of cDNA from RNA

RT step: to generate cDNA (use this cDNA as DNA template in PCR step)
RNA template (total RNA or mRNA)
Reserve Transcriptase
DNTPs, PCR buffer, RNAsin (RNase inhibitor)

General protocol: 90 °C, 5 min
quick chill on ice
incubate at 42 °C, 60 min
95 °C, 10 min

PCR

1. Amplification of DNA (PCR) or RNA (RT-PCR)
2. Competitive PCR for quantitation of mRNA
3. PCR with labeled probe or primer

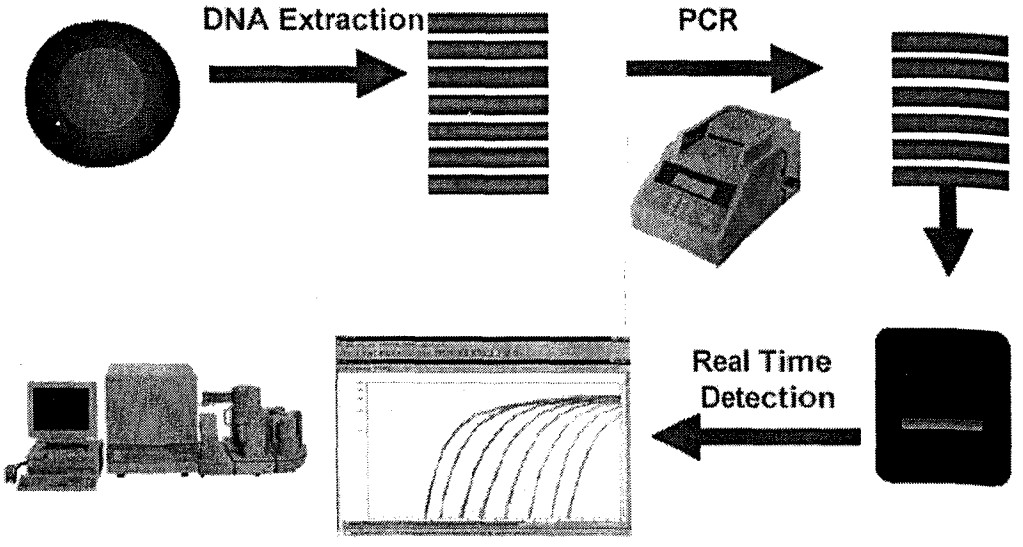
etc.

Exercise: 1. Application and experimental design
2. Practicing data analysis.

Topic 3. Real-Time RT-PCR Technique

The Evolution of PCR to Real-Time

PCR has completely revolutionized the detection of RNA and DNA. Traditional PCR has advanced from detection at the end-point of the reaction to detection while the reaction is occurring.



Source: Appliedbiosystem company product catalog 2006

Figure 1: Real-Time PCR Evolution

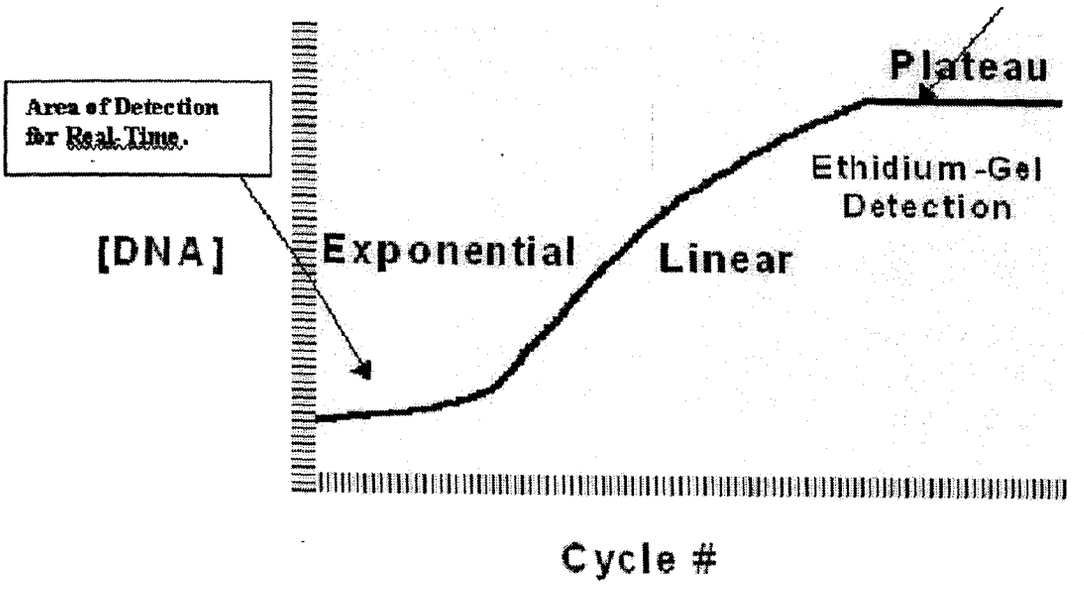
Real-Time Vs Traditional PCR
Real-Time chemistries allow for the detection of PCR amplification during the early phases of the reaction. Measuring the kinetics of the reaction in the early phases of PCR provides a distinct advantage over traditional PCR detection. Traditional methods use Agarose gels for detection of PCR amplification at the final phase or end-point of the PCR reaction.

Limitations of End-Point PCR
Agarose gel results are obtained from the end point of the reaction. Endpoint detection is very time consuming. Results may not be obtained for days. Results are based on size discrimination, which may not be very precise. As seen later in the section, the end point is variable from sample to sample. While gels may not be able to resolve these variabilities in yield, real-time PCR is sensitive enough to detect these changes. Agarose Gel resolution is very poor, about 10 fold. Real-Time PCR can detect as little as a two-fold change!
Some of the problems with **End-Point Detection**:
_ Poor Precision

- Low sensitivity
- Short dynamic range < 2 logs
- Low resolution
- Non - Automated
- Size-based discrimination only
- Results are not expressed as numbers
- Ethidium bromide for staining is not very quantitative
- Post PCR processing

Problems with detection in the Plateau phase of PCR

The following three figures show the plateau affect on 96 replicates and a fivefold dilution series. As stated earlier, the plateau region is the end-point of the reaction and is representative of the amount of product that you would see on Agarose Gels. The 96 replicates in the exponential phase are very tight in both the linear and logarithmic views.



Source: Appliedbiosystem system catalog product

Figure 2. In the logarithmic view, the plateau for each reaction seems to occur in the same place, but this is solely due to the log scaling of the plot. Figure 6 shows the same 96 replicates in linear view. The reactions show a clear separation in the plateau phase; therefore, if the measurements were taken in the plateau phase, quantitation would be affected.

Real-Time RT-PCR Technology

The advent of Polymerase Chain Reaction (PCR) by Kary B. Mullis in the mid-1980s revolutionized molecular biology as we know it. PCR is a fairly standard procedure now, and its use is extremely wide-ranging. At its most basic application, PCR can amplify a small amount of template DNA (or RNA) into large quantities in a few hours. This is performed by mixing the DNA with primers on either side of the DNA (forward and reverse), *Taq* polymerase (of the species *Thermus aquaticus*, a thermophile whose polymerase is able to withstand extremely high temperatures), free nucleotides (dNTPs for DNA, NTPs for RNA), and buffer. This movie shows PCR in action. The temperature is then alternated between hot and cold to denature and reanneal the DNA, with the polymerase adding new complementary strands each time. In addition to the basic use of PCR, specially designed primers can be made to ligate two different pieces of DNA together or add a restriction site, in addition to many other creative uses. Clearly, PCR is a procedure that is an integral addition to the molecular biologist's toolbox, and the method has been continually improved upon over the years.

Fairly recently, a new method of PCR quantification has been invented. This is called "real-time PCR" because it allows the scientist to actually view the increase in the amount of DNA as it is amplified. Several different types of real-time PCR are being marketed to the scientific community at this time, each with their advantages. This web site will explore one of these types, TaqMan® real-time PCR, as well as give an overview of the other two types of real-time PCR, molecular beacon and SYBR® Green.

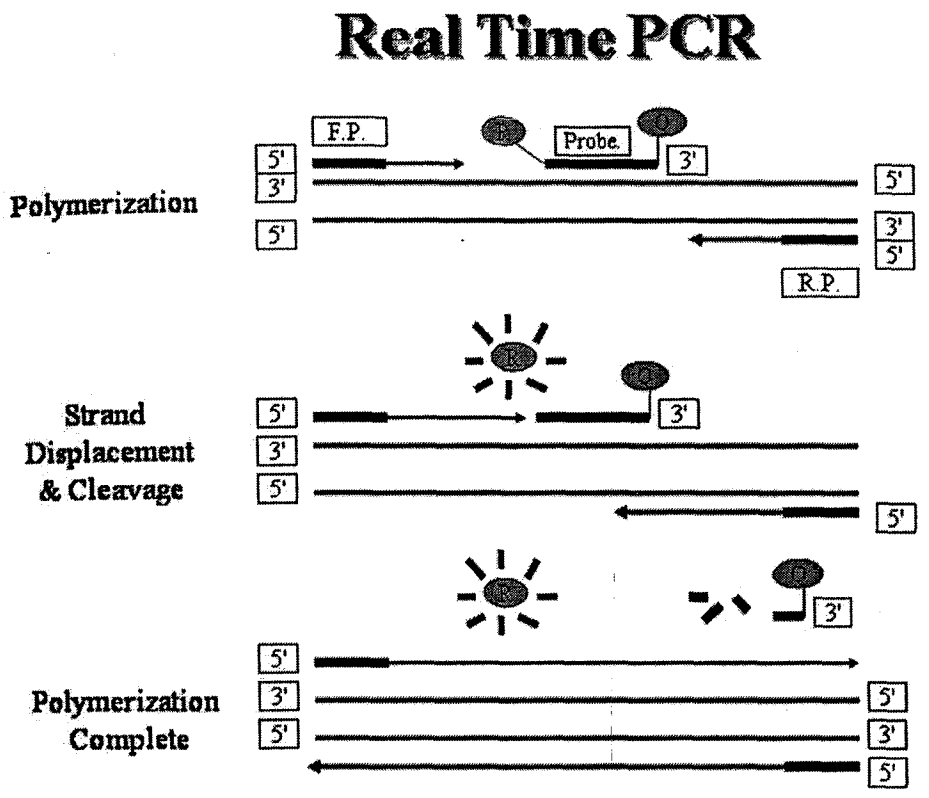
Principle and VDO Clip

1. Two fluorescent dyes, a reporter (R) and a quencher (Q), are attached to the probes used with the TaqMan PCR Reagent Kit. The 3' end of the probe is blocked, so it is not extended during the PCR reaction.

2. When both dyes are attached to the probe, reporter dye emission is quenched. During each extension cycle, the probe is displaced at the 5' end by the Taq DNA polymerase.

3. Taq DNA polymerase then cleaves the reporter dye from the probe via its 5'-3' exonuclease.

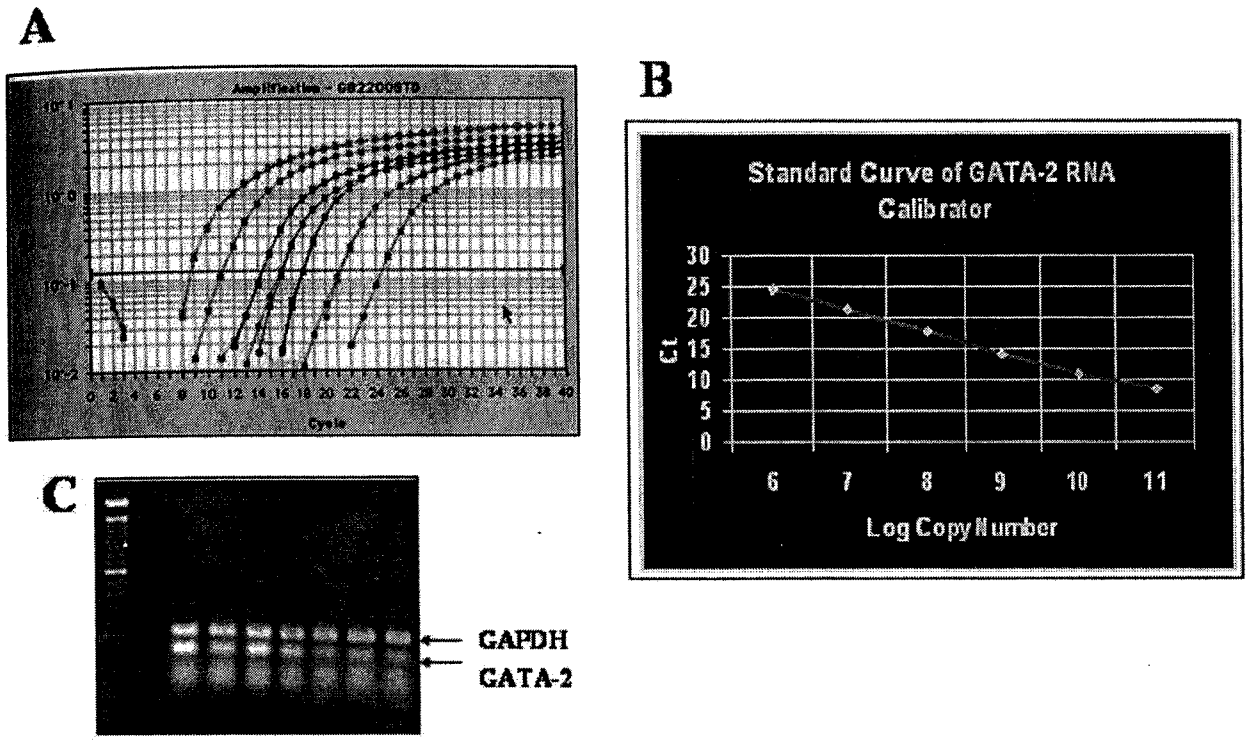
4. Once separated from the quencher, the reporter dye emits its characteristic fluorescence which can then be measured by the 7700. The amount of fluorescence measured is proportional to the amount of PCR product made



Source: Drawing by Wilairat Leeanansaksiri

The design of TaqMan probes and primers is critical to the success of the experiment. - - You may use the Primer Express program to assist in primer and probe design and ensure that all critical parameters are met.

Real Time PCR Results



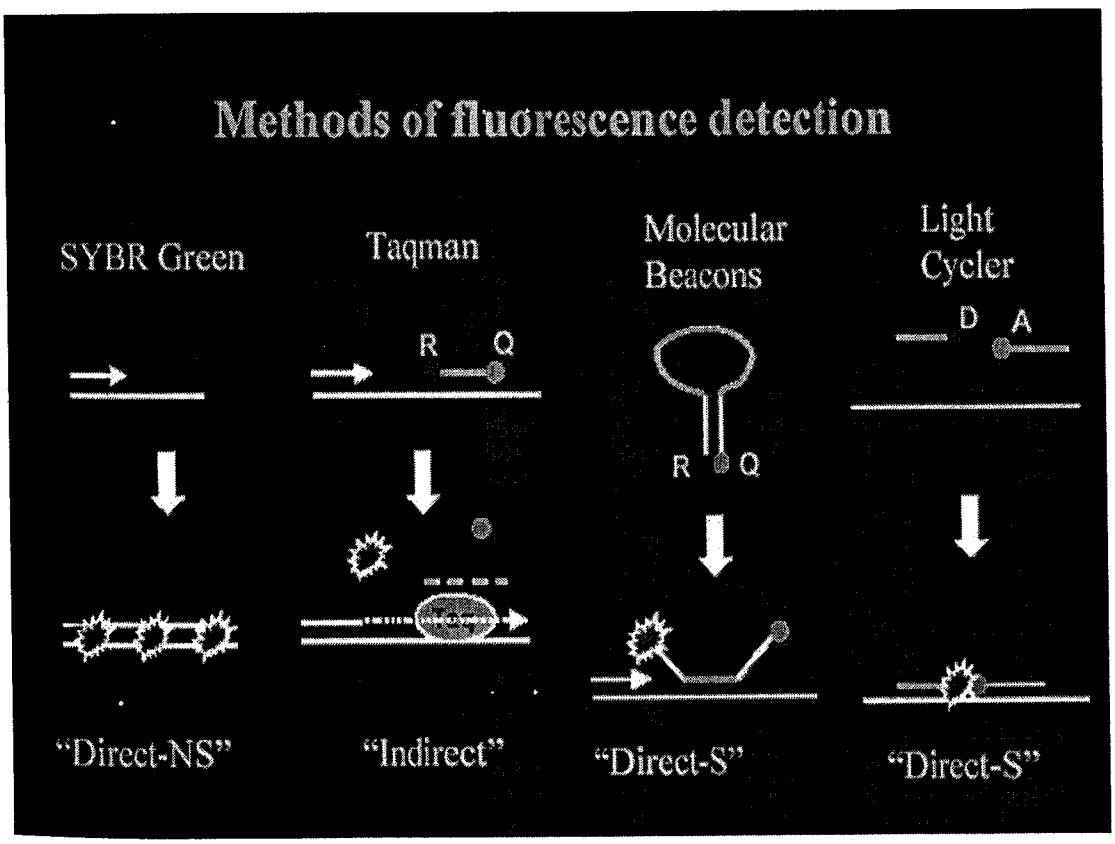
Source: Wilairat Leeanansaksiri - experiments on Real-time RT-PCR unpublished data

This Figure shows the amplification of standards from a dilution series. A. Normalized reporter signal is plotted against cycle number. The threshold cycle is the PCR cycle at which a significant increase in reporter fluorescence above baseline can be detected, and is marked by the darker horizontal line. B. The Standard curve is plotted in ct value and log copy number C. PCR analysis by gel electrophoresis demonstrates the specificity and sensitivity.

- In addition to the sequence-specific TaqMan probes, it is also possible to use SYBR Green as a fluorescent dye. This dye fluoresces only when bound to double-stranded DNA, which offers an alternative method of quantitating PCR products without the necessity of constructing a new TaqMan probe for each.
  - During the cycling process, fluorescence emission is measured every 6 seconds within each well. After the run, the results are analyzed and can be examined in a number of ways, one of which is the amplification plot shown below:

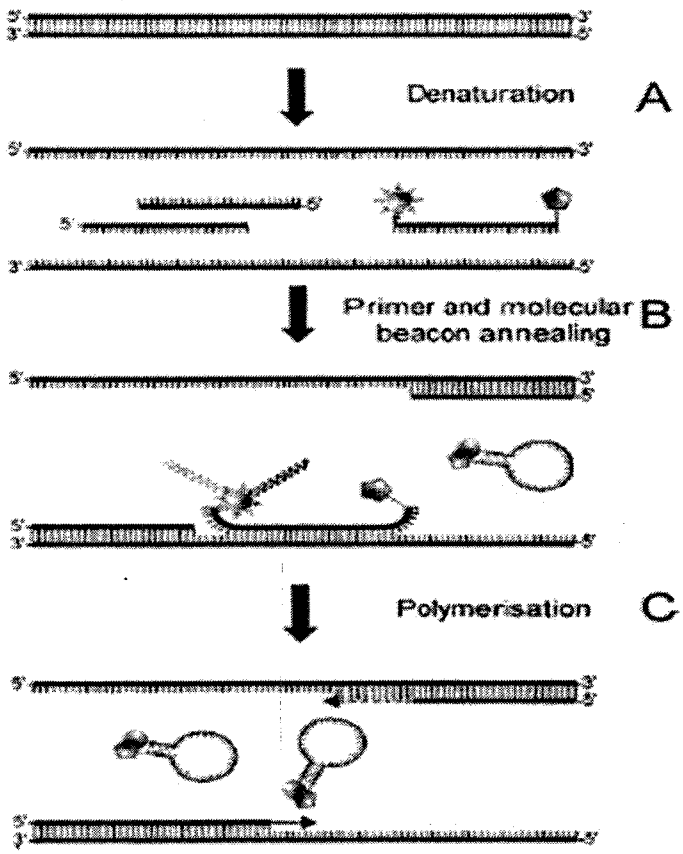
Application of PCR

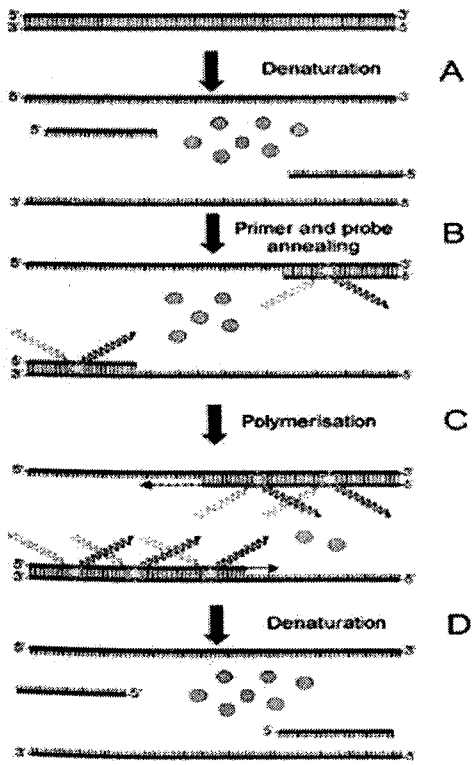
1. amplification of ancient DNA
2. screening mutation
3. genetic engineering (delete, insert, mismatch etc)
4. sequencing
5. cloning
6. screening of libraries
7. Genetic and evolution
8. Diagnostics
9. Forensics



Source: Powerpoint- Wilairat Leeanansaksiri

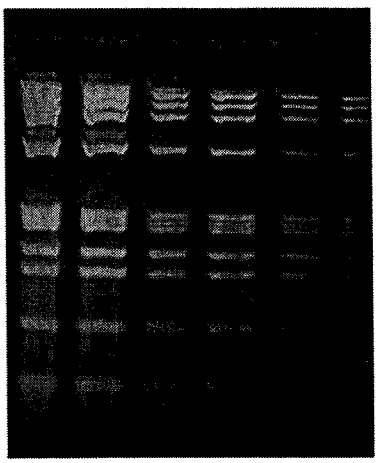
Molecular Beacons





Dye Incorporation Method

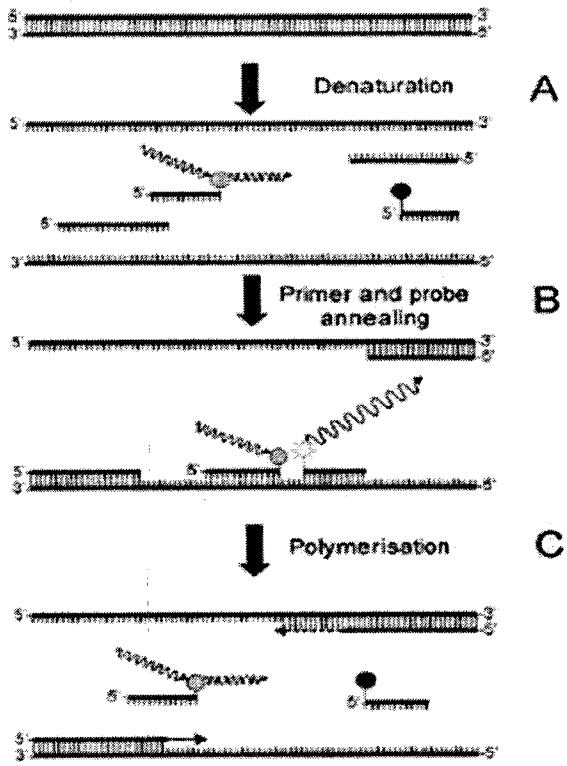
SYBR Green



Source: Product Catalog: In vitrogen 2003

Hybridization Probes Method

Light Cycler



Source: Product Catalog: Roche 2003

Organizing a Laboratory for Real-Time RT-PCR Work

Basic Equipment and Supply

Exercise: Experimental Design and Practicing data analysis of the real-time PCR.

Topic 4. DNA sequencing Technique

Maxam & Gilbert Method al clevage at specific bases
Not used for sequencing anymore

Sanger Method : enzymatic production of DNA that is terminated at specific bases

- Although it has undergone many refinements including ; improved enzymes, use fluorescent sequencing instead of radioactive, high throughput automation
- The principle has remained the same
- dye primers and dye terminators

Dye Primers	Dye Terminators
- Fluorescent dyes on the same primer	- Fluorescent dye on ddNTP, one color for each ddNTP
- 4 separate reactions with one color of dye-primer and one ddNTP	- one reaction only
- after reactions are complete, combine to run in one lane on gel Expensive if use many differentprimers	- ecomonic and convenient especially if you use many different primers
- Excellent signal to noise	- prone to more peak variability noise

Principle of sequencing and VDO clip

(This is only an explanation of the method used for sequencing on an automated sequencer ABI)

The purpose of sequencing is to determine the order of the nucleotides of a gene. For sequencing, we don't start from gDNA (like in PCR) but mostly from PCR fragments or cloned genes.

1. **The sequencing reaction :**

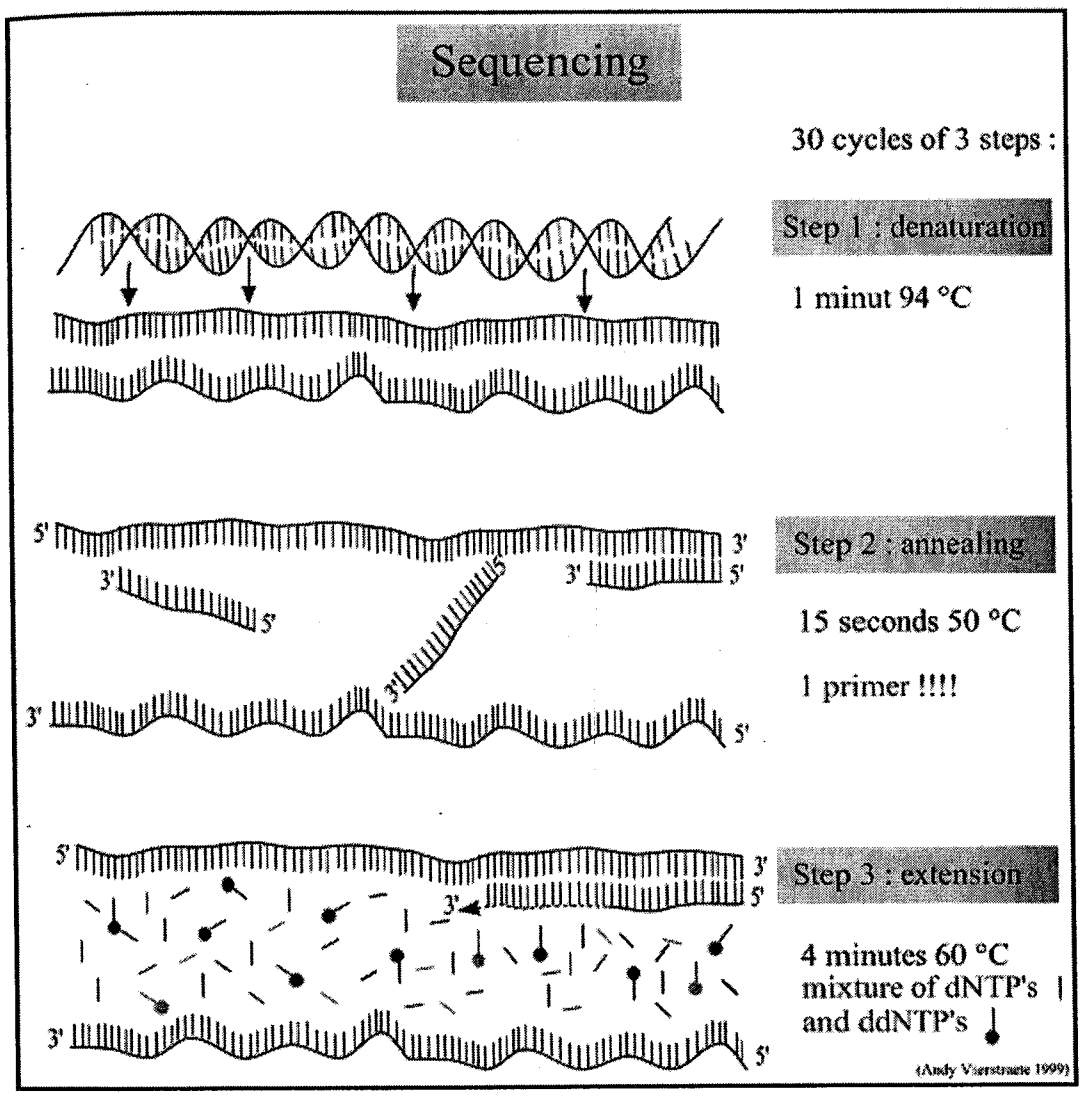
   There are three major steps in a sequencing reaction (like in PCR), which are repeated for 30 or 40 cycles.

   1. **Denaturation** at 94°C :
      During the denaturation, the double strand melts open to single stranded DNA, all enzymatic reactions stop (for example : the extension from a previous cycle).

   2. **Annealing** at 50°C :
      In sequencing reactions, only one primer is used, so there is only one strand copied (in PCR : two primers are used, so two strands are copied). The primer is jiggling around, caused by the Brownian motion. Ionic bonds are constantly formed and broken between the

single stranded primer and the single stranded template. The more stable bonds last a little bit longer (primers that fit exactly) and on that little piece of double stranded DNA (template and primer), the polymerase can attach and starts copying the template. Once there are a few bases built in, the ionic bond is so strong between the template and the primer, that it does not break anymore.

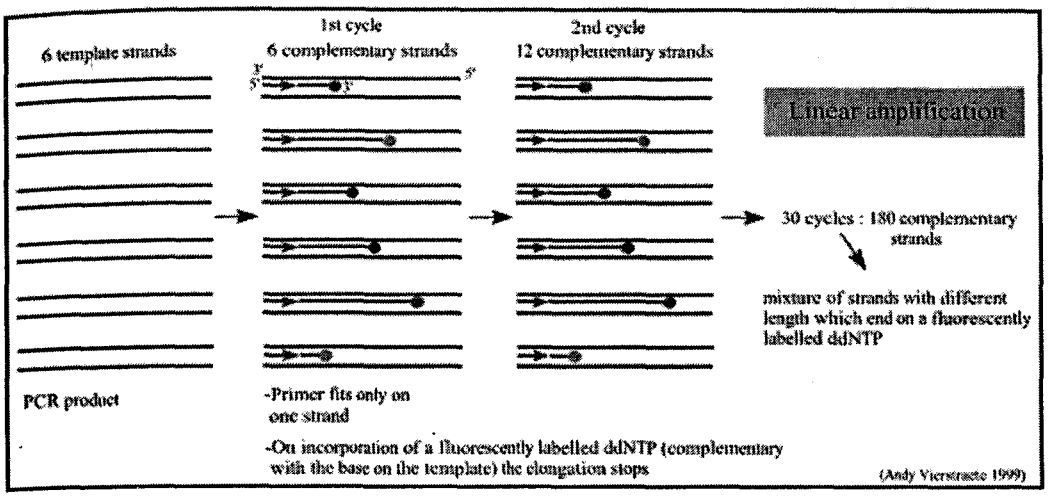
3. **Extension** at 60°C :
   This is the ideal working temperature for the polymerase (normally it is 72 °C, but because it has to incorporate ddNTP's which are chemically modified with a fluorescent label, the temperature is lowered so it has time to incorporate the 'strange' molecules. The primers, where there are a few bases built in, already have a stronger ionic attraction to the template than the forces breaking these attractions. Primers that are on positions with no exact match, come loose again and don't give an extension of the fragment.
   The bases (complementary to the template) are coupled to the primer on the 3'side (adding dNTP's or ddNTP's from 5' to 3', reading from the template from 3' to 5' side, bases are added complementary to the template)
   When a ddNTP is incorporated, the extension reaction stops because a ddNTP contains a H-atom on the 3rd carbon atom (dNTP's contain a OH-atom on that position). Since the ddNTP's are fluorescently labeled, it is possible to detect the color of the last base of this fragment on an automated sequencer.



Source: http://users.ugent.be/~avierstr/principles/seq.html

Figure 1: The different steps in sequencing

Because only one primer is used, only one strand is copied during sequencing, there is a **linear** increase of the number of copies of one strand of the gene. Therefore, there has to be a large amount of copies of the gene in the starting mixture for sequencing. Suppose there are 1000 copies of the wanted gene before the cycling starts, after one cycle, there will be 2000 copies : the 1000 original templates and 1000 complementary strands with each one fluorescent label on the last base, after two cycles, there will be 2000 complementary strands, three cycles will result in 3000 complementary strands and so on.

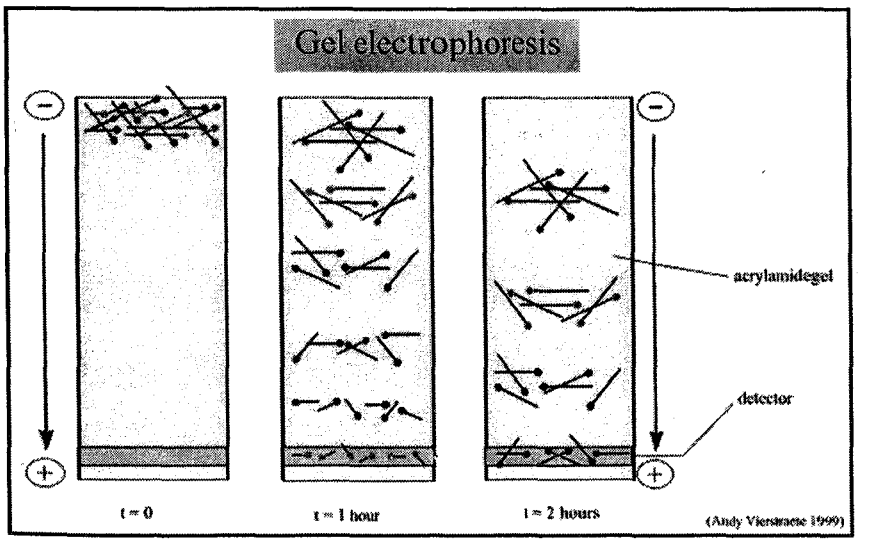


Source: http://users.ugent.be/~avierstr/principles/seq.html

Figure 2 : The linear amplification of the gene in sequencing.

2. **Separation of the molecules :**

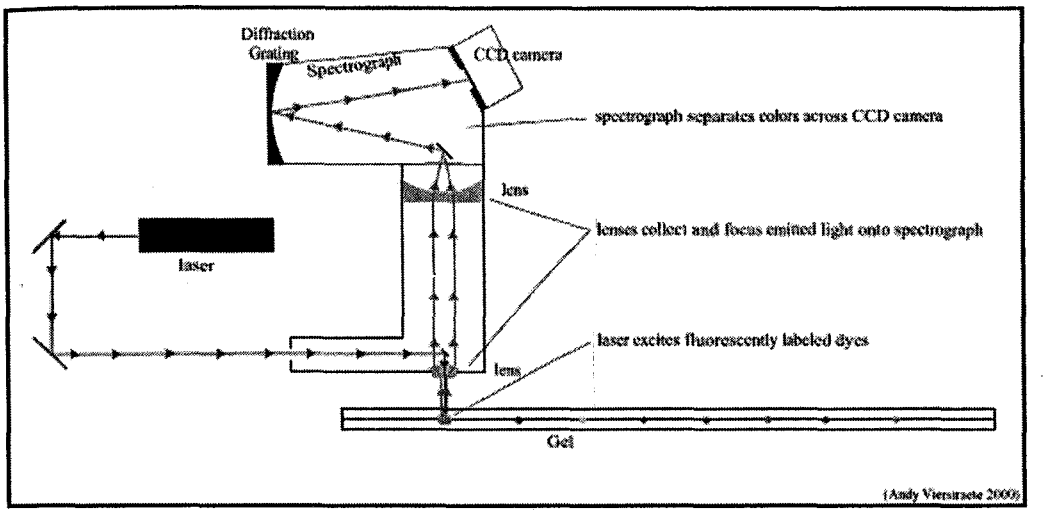
After the sequencing reactions, the mixture of strands, all of different length and all ending on a fluorescently labelled ddNTP have to be separated; This is done on an acrylamide gel, which is capable of separating a molecule of 30 bases from one of 31 bases, but also a molecule of 750 bases from one of 751 bases. All this is done with gel electrophoresis. DNA has a negative charge and migrates to the positive side. Smaller fragments migrate faster, so the DNA molecules are separated on their size.



Source: http://users.ugent.be/~avierstr/principles/seq.html

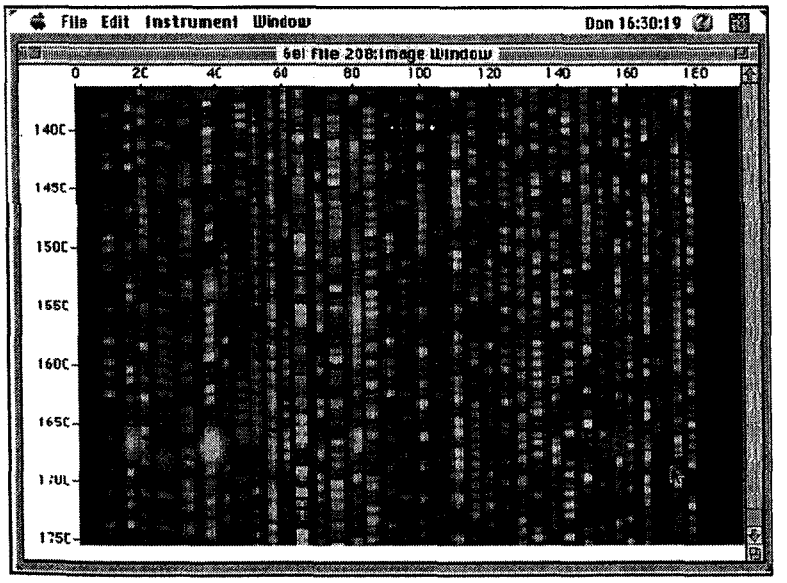
Figure 3 : The separation of the molecules with

Detection on an automated sequencer :
The fluorescently labelled fragments that migrate trough the gel, are passing a laser beam at the bottom of the gel. The laser exites the fluorescent molecule, which sends out light of a distinct color. That light is collected and focused by lenses into a spectrograph. Based on the wavelength, the spectrograph separates the light across a CCD camera (charge coupled device). Each base has its own color, so the sequencer can detect the order of the bases in the sequenced gene.



Source: http://users.ugent.be/~avierstr/principles/seq.html

Figure 4 : The scanning and detection system on the ABI Prism 377 sequencer.

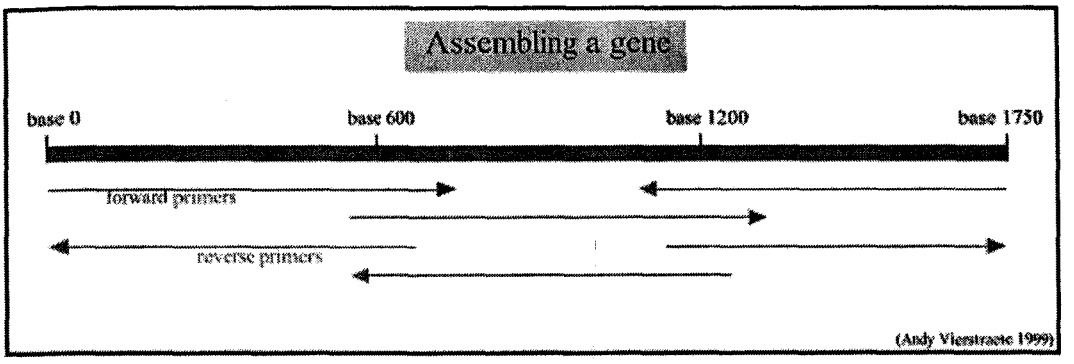


Source: http://users.ugent.be/~avierstr/principles/seq.html

Figure 5 : A snapshot of the detection of the molecules on the sequencer.

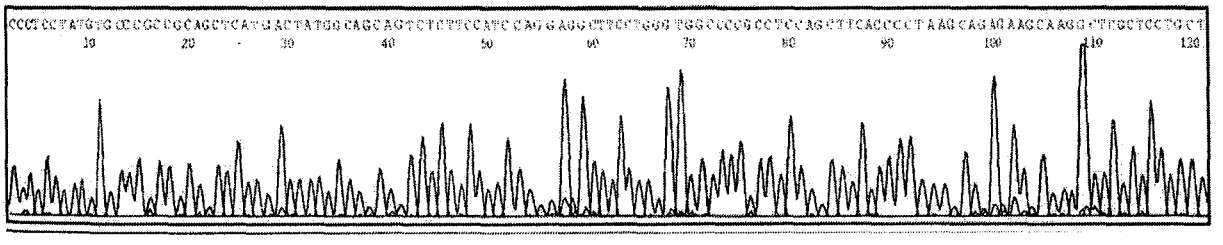
3. **Assembling of the sequenced parts of a gene :**

For publication purposes, each sequence of a gene has to be confirmed in both directions. To accomplish this, the gene has to be sequenced with forward and reverse primers. Since it is only possible to sequence a part of 750 till 800 bases in one run, a gene of, for example 1800 bases, has to be sequenced with internal primers. When all these fragments are sequenced, a computer program tries to fit the different parts together and assembles the total gene sequence.



Source: http://users.ugent.be/~avierstr/principles/seq.html

Figure 6 : The assemblage of the gene.



Souce: Wilairat Leeanansaksiri-Experiment on DNA sequencing (unpublish data)

Figure 7. DNA sequencing result from computer analysis

Topic 5. Southern Blotting Technique

Southern Blotting

- A technique described by Southern (1975)
- A technique used to transfer DNA fragments from agarose gel to solid support. The attached DNA is hybridized to radiolabeled DNA or RNA, and autoradiography is used to located the positions of bands complementary to the probe.

Procedure: 1. DNA extraction and digestion with restriction enzymes
Or PCR fragments
2. agarose gel electrophoresis
3. transfer DNA from agarose to solid support e.g. nitrocellulose or nylon membrane
4. UV crosslinking
5. prehybridization
6. hybridization of radilabled probes to immobilized nucleic acids

Restriction Enzyme : is a protein

1. Type I:
   - bind to the recognition site but cut at random sites
   - carry modification (methylation) and ATP-dependent restriction (cleavage) activity in the same protein
2. Type II : most used in molecular biology
   - restriction endonuclease; cut at recognition site
   - methylase that modifies the same recognition sequence
   - table of recognition sequences
3. Type III (same as type I)
   - bind to the recognition site but cut at random sites
   - carry modification (methylation) and ATP-dependent restriction (cleavage) activity in the same protein

Transfer of DNA from agarose gels to solid supports

1. Capillary transfer
2. Electrophoretic transfer
3. Vacuum transfer

Transfer solution: General protocol

1. Denature the DNA by soaking the gel in 1.5 M NaCl/0.5 N NaOH for 45 minutes
2. Rinse the gel briefly in deionized water
3. Denatralize by soaking the gel in 1 M Tris (pH 7.4)/1.5 M NaCl for 30 minutes
4. 20X SSC

Membrane : Advantages and disadvantages of nitrocellulose and nylon membranes

Making a probe:

5'-end labelled with ^{32}P-ATP using T4 polynucleotide kinase

function: Bacteriophage T4 polynucleotide kinase catalyzes the transfer of the γ phosphate of ATP to a 5'terminus of DNA ot RNA

Hybridization

Prehybridization solution:

6XSSC
5X Denhardt's reagent
0.5% SDS
100 μg/ml denatured, fragmented salmon sperm DNA
0.01% polyvinylpyrrolidone
0.1% Ficoll
0.01%BSA

Hybridization solution: prehybidization solution plus probe

Washing solution: 2XSSC / 0.1% (w/v) SDS
0.2XSSC/0.1% (w/v) SDS, room temperature and 42°C
0.1XSSC/0.1% (w/v) SDS, 68°C

Autoradiograph : **X-ray film**
Phosphoimager

Topic 6. Genetic Engineering and Cloning Techniques

Clone Any group of cells or individuals that have descended from a single cell by asexual reproduction. Since no genetic recombination occurs with asexual reproduction, all members of a clone are genetically identical.

Cloning refers to the process of making many copies of a particular DNA sequence. "to clone a gene" means to transfer the gene of interest to a system in which billions of copies can easily be made

---

DNA polymerase uses a single strand template to make a complementary strand. However, it requires a **primer** and can only add nucleotides to the 3' end of the primer.
DNA ligase ties together adjacent nucleotides in a DNA strand if they are bound to the same complementary strand.
Reverse transcriptase is a DNA polymerase that uses RNA as a template to make a DNA copy.
Transduction refers to the transfer of a bacterial gene from one bacterium to another by a phange; a phange carrying host as well as its own genes is called transducing phange. Also describes the acquisition and transfer of eukaryotic cellular sequences by retroviruses.

Transfection of eukaryotic cells is the acquisition of new genetic markers by incorporation of added DNA

Transformation of bacteria describes the acquisition of new genetic markers by incorporation of added DNA
Transformation of eukaryotic cells refers to their conversion to a state of unrestrained growth in culture, resembling or identical with the tumorigenic condition

---

Restriction Endonucleases : Type II

- Restriction endonucleases (**restriction enzymes**) break double-stranded DNA at specific nucleotide sequences (**restriction sites**).
- The most useful are those that cleave at **palindromic** sequences. These are DNA sequences that read the same on both strands but in opposite directions. For example, 5'-AGCT-3' on one strand and 3'-TCGA-5' on the complementary strand.
- Found in bacteria along with methylases, the restriction modification system.
- There are hundreds of known restriction enzymes, each specific for sequences consisting of 4 to 8 nucleotides. Restriction enzymes (R.E.) may recognize 4, 6 and rarely 8 bp (4(4), 4(6) and 4(8)bases)
- R.E. that recognize identical sequences but cut at different sites are isoschizomers.
- R.E. may cut and leave overhangs (sticky ends) or may be blunt cutters (blunt ends).
- The entire DNA of a human would be broken into millions of **restriction fragments** by some restriction enzymes, but the sequences cleaved would be identical. All the ends of the fragments would be identical also.
- *DNA cloning: vector and the DNA of interested is cut with R.E. and ligated with*

DNA ligase = *recombinant DNA.*

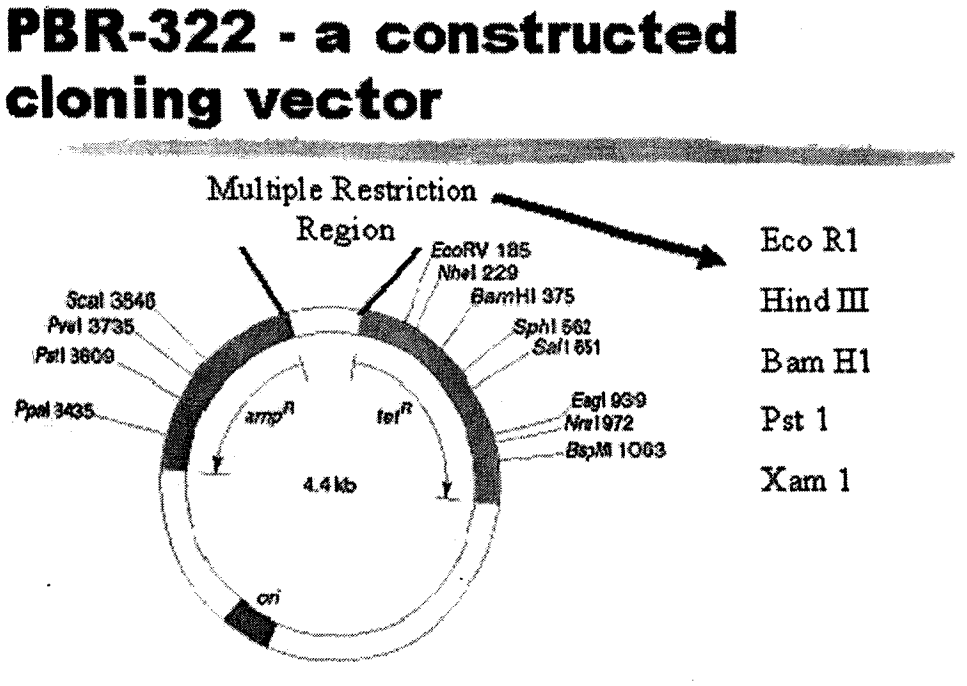
Vectors

Plasmids:

Plasmid vectors are modified versions of naturally occurring plasmids such as ColE1 from *E.coli*. pBR322, one of the first plasmid vectors, contains the ColE1 origin of replication (*ori*), and unique restriction sites in each of two selectable markers (*Pst*I in *ampR* and *Hin*dIII in *tetR*). More sophisticated vectors, like pUC19, have an *ampR* gene and a *lacZ* coding sequence containing a multiple cloning site (or polylinker) which provides a selection of different restriction enzyme sites for cloning; the *lacZ* sequence allows blue/white selection.

pBR322 tetracycline and ampicillin markers. PUC series retain ampicillin resistance and have cloning sites clustered in multiple cloning sites (MCS). The MCS lies within DNA sequence coding for NH2 terminal portion of beta-gal. Beta-gal activity is seen by blue coloration of colony (galactose+indigo dye resulting from X-gal cleavage by beta-gal).

Plasmids can accommodate a max of 10kb inserts. They cannot replicate with inserts greater than 10kb.
Directional cloning can be achieved by selective R.E. usage. This also prevents vector religation.



Source: Progema Product catalog 2004

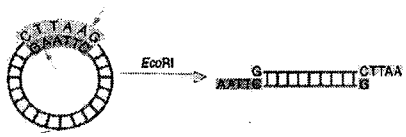
Phage vectors:

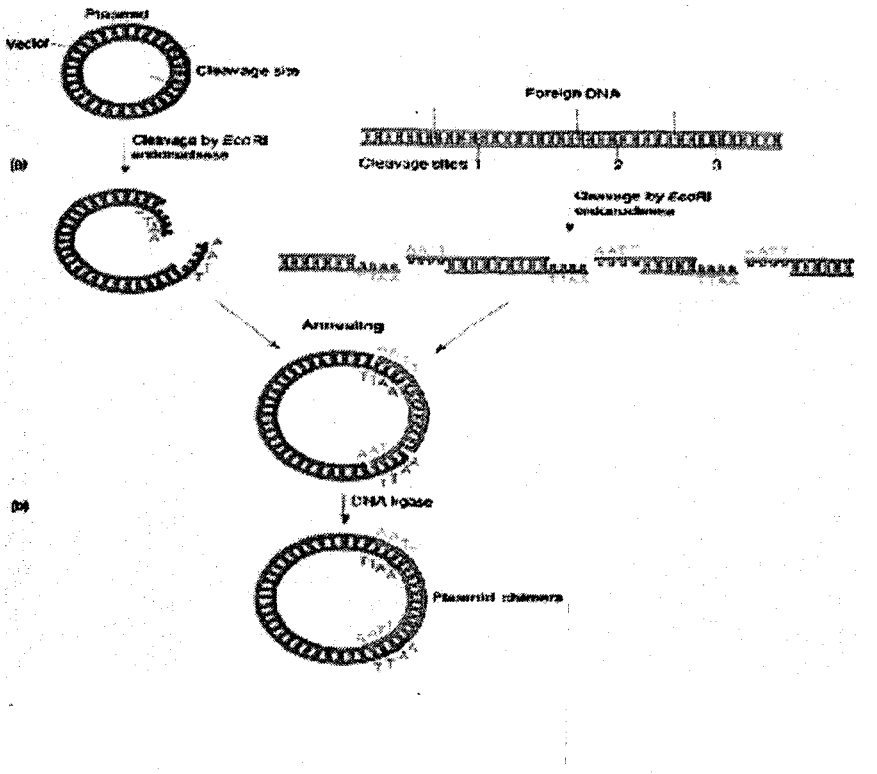
- Bacteriophages infect cells much more efficiently. Lambda phage retains genes needed for phage replication. Phage vectors accommodate no less that 12kb and up to 20kb of foreign DNA.
- Because infection is very efficient lambda vectors are often used for making DNA libraries.
- Libraries are made by partial R.E. digestion or by mechanical shearing (i.e. ultrasound).

Cosmids: Vectors useful for cloning large DNA frag. (40-50kb). They behave as plasmids and phages. They can be packaged into lambda phage heads (cos), and they contain plasmid origin of replication sites, so they can replicate in bacteria (mid).
M13 phage vectors: They have beta-gal site interrupted by MCS. The genome of M13 is single stranded, but after infection of E. coli, it is double stranded. This form is used for cloning. Recombinant DNA can be recovered as single stranded DNA.
Phagemids: pBluescript has MCS inserted into lac Z gene, origin of replication, and MCS flanked by RNA polymerase promoters (T7 RNA pol at one end and T3 pol at the other end).
Yeast Artificial Chromosomes (YACs) are a more recent generation of cloning vector that possess a yeast origin of replication, a centromere and telomeres; they can carry up to 2000kb of foreign DNA. YACs are often used in genome projects to create a library of overlapping clones representing a part of, or the whole of, a chromsome. However, YACs can be difficult to work with and the DNA can become rearranged.
Viral vectors: Retrovirus, Lentivirus, Adenovirus etc.

1) Tools required for gene cloning.
a) Restriction enzymes: These are enzymes that digest or "cut" DNA at specific DNA sequences, often generating sticky ends.
b) Ligase: This the enzyme that links Okazaki fragments during DNA replication. It can also be used to link DNA in a test tube during gene cloning.
c) Bacterial plasmid: a plasmid is a small circular piece of DNA that replicates many times in a bacterial cell. It is used to make many copies of a gene during gene cloning. In this context the bacterial plasmid is called a cloning vector.

2) A typical gene cloning experiment involves a number of steps.
a) Isolation of vector DNA and subject gene DNA.
b) Insertion of the gene into the vector
c) Introduction of the vector into bacterial cells.
d) Cloning of cells
e) Identification of a clone of cells carrying the gene of interest.





Source: Gene VI, Benjamin Lewin 1997

3) Often a gene of interest is recovered from a library of genes.
a) genomic library – contains all of the DNA sequences in a genome.
b) cDNA library – contains copies of all of the expressed genes in a particular tissue.
i) cDNA is synthesized using reverse transcriptase and inserted into a cloning vector.

Host cell transformation with Recombinant DNA.

- Transformation : plasmid, competent cell (COMPETENT =capable of accepting DNA from the external environment).
- Bacteria (*E. coli*) is incubated with high calcium salt concentration to make a leaky plasma membrane
- Use of high voltage (electroporation)
- Coat tiny metal pellets with DNA
- Use temperature sensitive bacteria

PCR also generates a DNA fragment for cloning
- Cloning of cDNAs
- Taq Polymerase (*Themus aquaticus*)

cDNA cloning
- cDNA library (oligo dT primer), reverse transcriptase
- RT-PCR

Identifying a clone
- Probes: polynucleotide probe and antibodies
- What do we do when we do not know the exact DNA sequence?
- When we know the protein sequence, can we design a polynucleotide probe?

Expressing Cloned Genes

- Bacterial expression vectors have 2 elements: strong promoters and ribosome binding sites near AUG
- What is an inducible promoter: lac promoter stimulated by IPTG. Temperature sensitive promoters (lambda repressor is active at 32oC)

Expression vectors producing fusion proteins: most vectors produce fusion proteins. How can a fusion protein help us purify a protein product?

Expression vector

1. Prokaryotic system
2. Eukaryotic system

Exercise 1. Design cloning method and strategy
2. Methods of the confirmation of accomplishment

Topic 7. DNA Microarray Technique

PRINCIPLE:

DNA microarray is a high throughput gene expression analysis technology. This technique may be defined as a miniaturized, systematic immobilization of nucleic acid fragments procured from individual genes on a solid support, which by specific hybridization, enable simultaneous analysis of thousands of genes in parallel (1). They have entered the front line of investigational medicine tomeet the demand for high-throughput survey as we enter the post-genomic era. Microarrays represent a bioinformatics-based intersection between biology and computers that enables gene analysis in human tissues on a genome-wide scale (2). Segments of DNA that serve as probes for detection are arranged regularly on nylon or a glass support, which forms the so-called 'gene chip', or microarray. Hybridizing these arrays with labeled nucleic acids from tissue samples allows quantitative measurement and comparison of genetic information.

DNA microarrays owe their power as an experimental tool to the specificity and affinity of complementary base pairing. Many thousands of oligonucleotides or cDNA clones can be spotted onto a single glass slide microarray, and indeed most of the genome can now be interrogated in a single microarray. Thus, an expression level "snapshot" of cellular activity provides an unprecedented tool for exploring the behavior of the genome under almost any conditions of cell culture, and is amenable to repetitive analyses of frozen human tissue.

THE MICROARRAY SUBSTRATE

In principle, membrane-based arrays can be probed by radiolabeled mRNA to measure gene expression; however, slidebased arrays have proven to be smaller, more convenient, and to facilitate higher throughput. There are two important strategies in the fabrication of slide-based microarrays. One strategy is commonly referred to as "in situ oligo synthesis." In this approach, pioneered by the Affymetrix® Corporation, sequences of fifteen to twenty-five nucleotides can be accurately and efficiently synthesized. In an alternative approach, developed at Rosetta Inpharmatics and Agilent Technologies, an inkjet printer, rather than photolithography, is used to apply sequential rounds of synthesis, using standard phosphoramidite chemistry, and allows the construction of oligonucleotides of sixty to eighty nucleotides in length. This latter style of in situ oligo array can permit competitive hybridization of two samples in the same manner as do the batch-synthesized arrays described below. In the other major strategy, batches of "bio-ink" are synthesized in large quantities, and then printed on a substrate, usually a treated glass microscope slide (Figure 1), through any of a variety of techniques, including both contact and inkjet printing. The bio-ink can be cDNAs from large expressed sequence tag (EST) libraries or long oligonucleotide chains produced in largescale oligo synthesizers. Whereas in situ oligo arrays offer some flexibility in the choice and arrangement of features to be printed, bio-ink arrays benefit from the economy inherent in synthesizing reagents on a large scale.

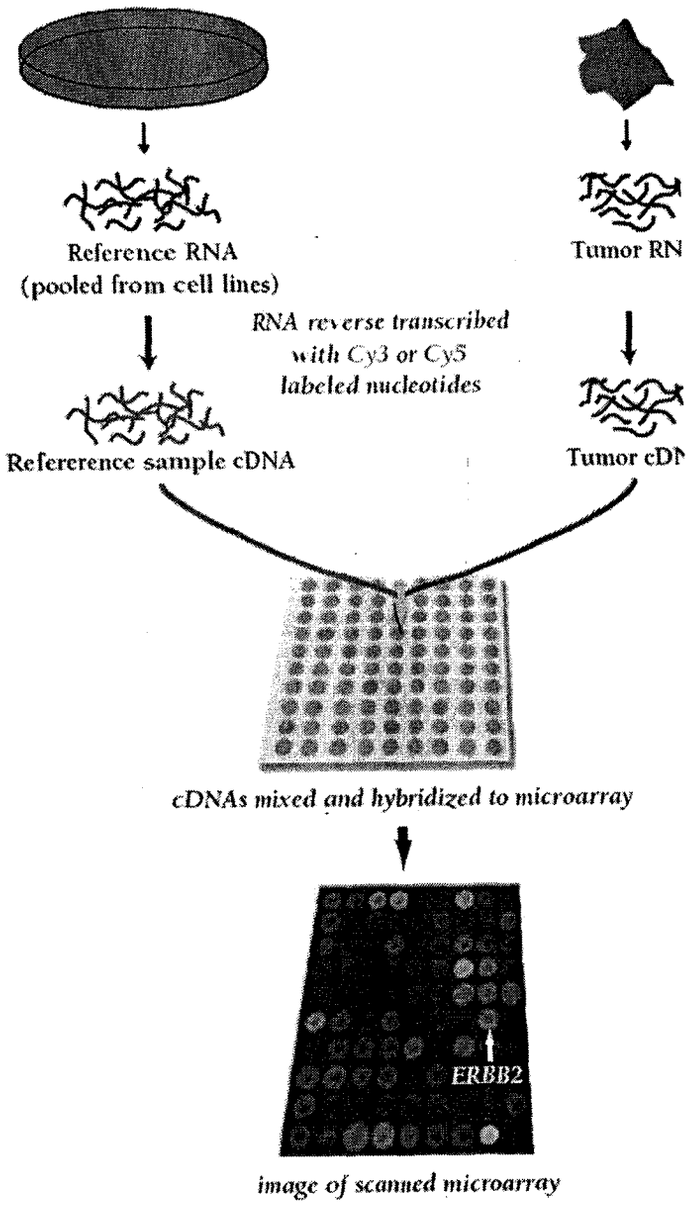
PROBE PREPARATION

Either total RNA or mRNA can be isolated from snap-frozen tumors and used for microarray hybridization. Standard protocols require relatively large amounts of RNA (2–4 micrograms of mRNA or 50 micrograms of total RNA) to achieve adequate signal for weakly expressed transcripts. Because these amounts of RNA probes require larger starting quantities of human tissue than is often available, RNA amplification strategies have become important. These methods permit the use of much smaller quantities of total RNA while still maintaining excellent fidelity with respect to the original tissue sample. When arrays are used to explore the biology of individual tissue components (such as epithelial cells or stromal fibroblasts), laser capture microdissection can be used to purify cell populations *(2)*. The isolated RNA is labeled with fluorescent dyes. By convention, the experimental sample is labeled with Cy 5, a red dye that fluoresces at 635 nm, by means of a reverse transcription reaction. The Cy 5–labeled cDNA sample is mixed with a reference cDNA that has been labeled with Cy 3, a green dye that fluoresces at 532 nm, and the mixture is hybridized to the array (Figure 2). An optical scanner measures fluorescence at the two specified wavelengths, and the ratio of signal intensities from the experimental and reference RNA represents the relative abundance of transcripts present in the sample. This use of ratios permits the canceling out of systematic errors, such as unknown quantities of DNA spotted onto the array and differing hybridization kinetics. Reference RNA, now commercially available, is composed from RNA pooled from ten cell lines that express the majority of human genes and provide a renewable resource. All expression levels in the sample are reported relative to the nominal level provided by the reference RNA, which permits samples using the same reference RNA to be compared.

DATA RETRIEVAL AND ANALYSIS

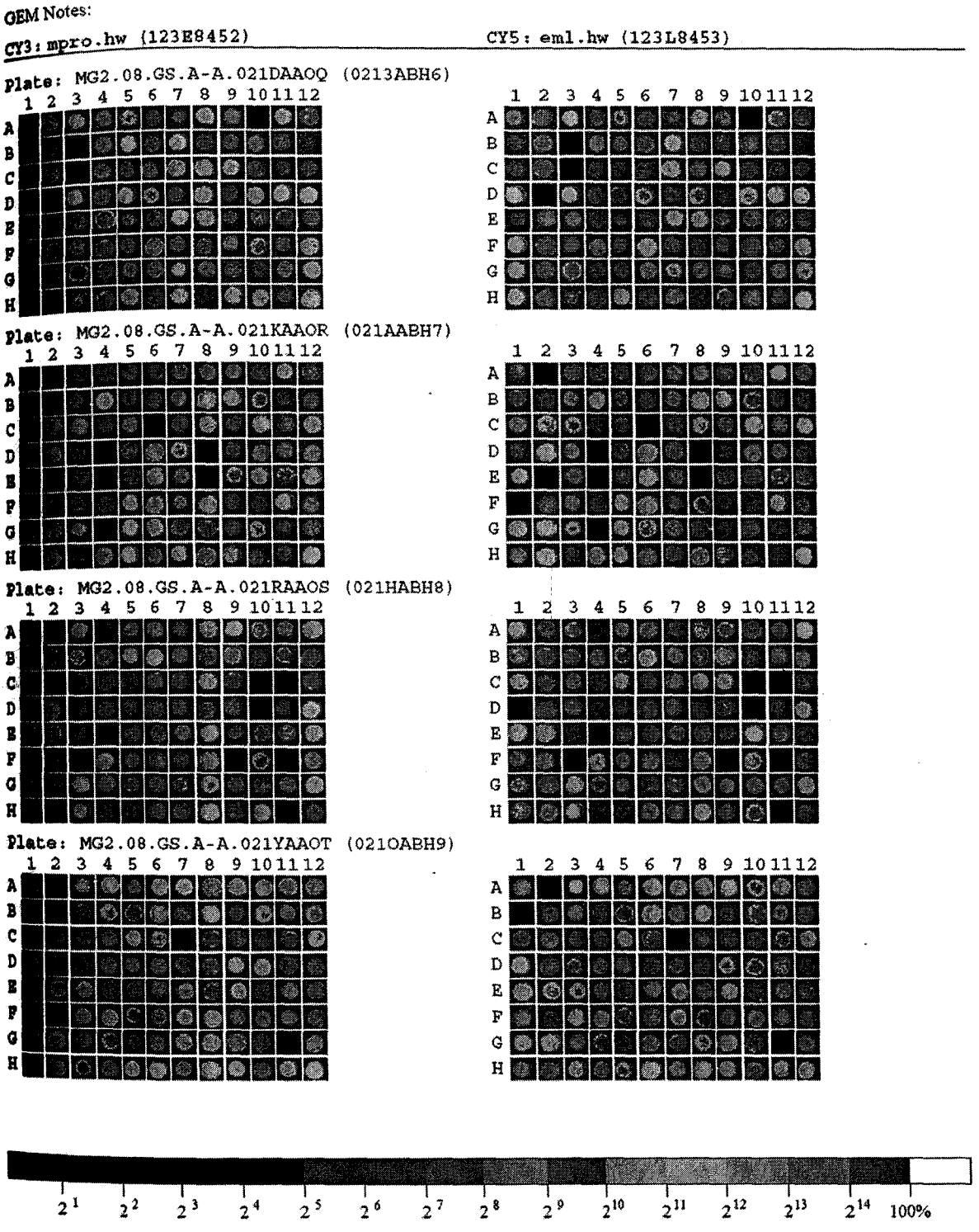
Because microarrays measure the transcription of genes on a genome-wide basis, they detail the cell's metabolic status in terms of which genes are induced or repressed. More importantly, such widespread expression data provide information on relationships among genes that may not have been previously suspected. When microarrays are used to study human cancer, they reveal information about the state of a tissue, even including interactions between malignant cells and their microenvironment. Examples of such interactions include tumor angiogenesis, which is an endothelial response to an epithelial malignancy, or the infiltration of tumor-associated lymphocytes. The complexity of interactions between different cell types in a tissue is thus measurable with microarrays; the analysis of such data, however, presents a formidable problem. The output of experimental information from cDNA microarrays is a plethora of fluorescence intensities. These raw data must be stored in a retrievable format, analyzed, and optimally, subjected to a visualization method so that researchers can use their own intrinsic neural networks (i.e., their brains) to interpret correlations among numerous diverse but interrelated data points. Web-based formats such as the Stanford Microarray Database (SMD; http://genome-www5.stanford.edu/MicroArray/SMD) provide such a repository for data storage and analysis. SMD provides a Web-based link between specific data (gene spots) and publicly accessible databases, such as UniGene, dbEST, and Swiss-Prot, for valuable

information about the clones used in array preparation (http://genome-www5.stanford.edu/cgi-bin/SMD /source/sourceSearch). Moreover, SMD is a resource for public dissemination of expression data for all published studies. One of the most commonly used methods for microarray data analysis is hierarchical clustering *(6)*, a method for organizing genes and experimental samples according to similarity of gene expression profile. In this method, a mathematical vector for the expression profile of each gene in a sample is determined and compared to the vectors from all of the genes expressed in that sample. Genes with similar expression patterns can then be clustered near one another along one axis, and similarly, experimental samples can be ordered according to their overall similarity in gene expression patterns along a second axis. The result is that genes whose expression patterns are most similar are clustered together by rows, and experimental samples whose expression patterns are most similar are clustered by columns Dendrograms linking genes or experimental samples can then be generated to show degrees of similarity. In this way, useful relationships between coexpressed genes or similar experimental samples may be discovered. Besides hierarchical clustering, algorithms that detect pattern similarity include k means clustering and self-organizing maps. Singular value decomposition (principal component analysis) is used to identify patterns that contribute to the overall gene expression matrix *(7)*. Neural networks may apply analytic techniques to a training set of data and test the result on a separate validation set of samples. Hierarchical clustering that is performed on entire data sets without expectation or previous information regarding results is known as "unsupervised clustering," and offers the virtue that no assumptions are made. Gene expression patterns can also be determined using supervised methods, provided that one has prior information about the genes or tissue samples so that predictions of clustering can be made. Information such as clinical outcome, pathologic subtype, or sample similarity may be used to "supervise" an analysis in order to generate a pertinent gene list that may then be tested against a new set of unknown samples.



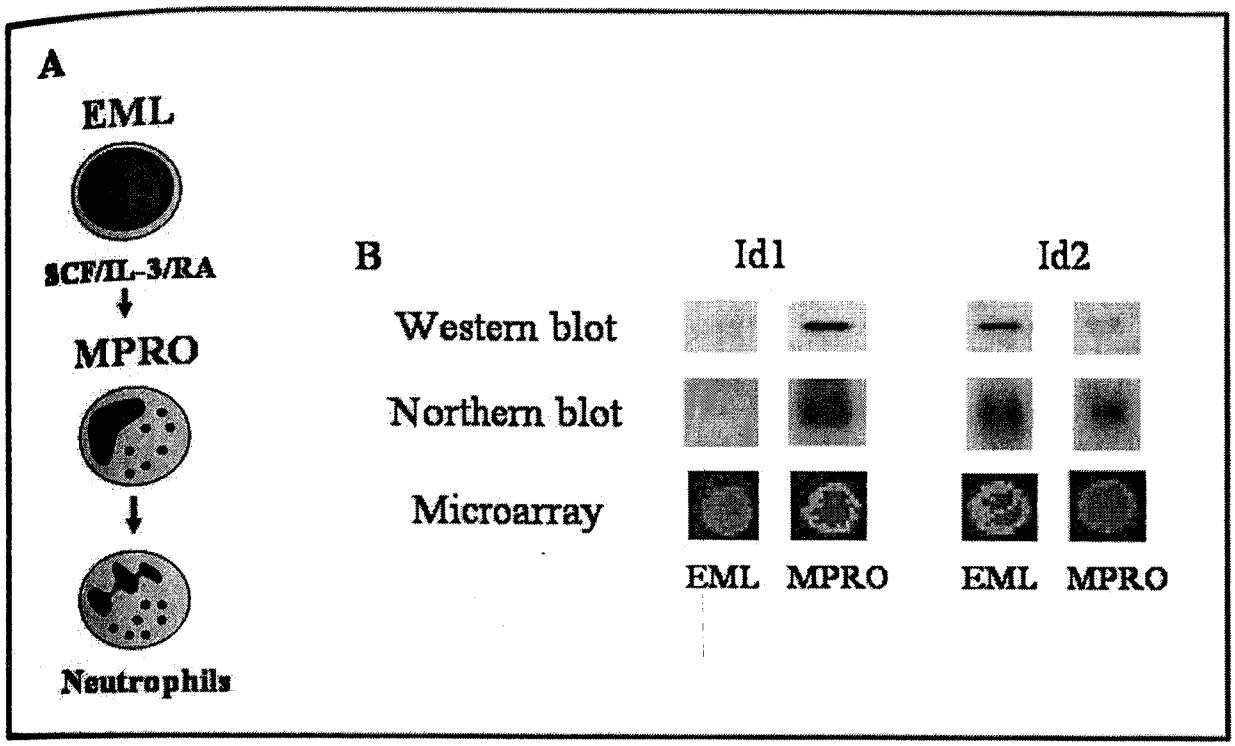
Source: Molecular Intervention, April 2002 Volume 2, Issue 2 page 101-109.

Figure 1. Schematic of microarray technique. RNA from a tumor sample and reference RNA (made commercially from pooled cell cultures torepresent the majority of known genes) are reverse transcribed and labeledwith different fluorescent dyes. The mixture is hybridized overnight to a microarray. The hybridized microarray is then scanned at two wavelengths and the intensities of red and green fluorescence are measured at each spot on the microarray. The red-to-green ratio reveals the abundance of RNA expressed by the tumor sample relative to the reference sample for every one of the 42,000 cDNA clones on the array. This technique provides a comparative measure of the global gene expression of the tumor sample.



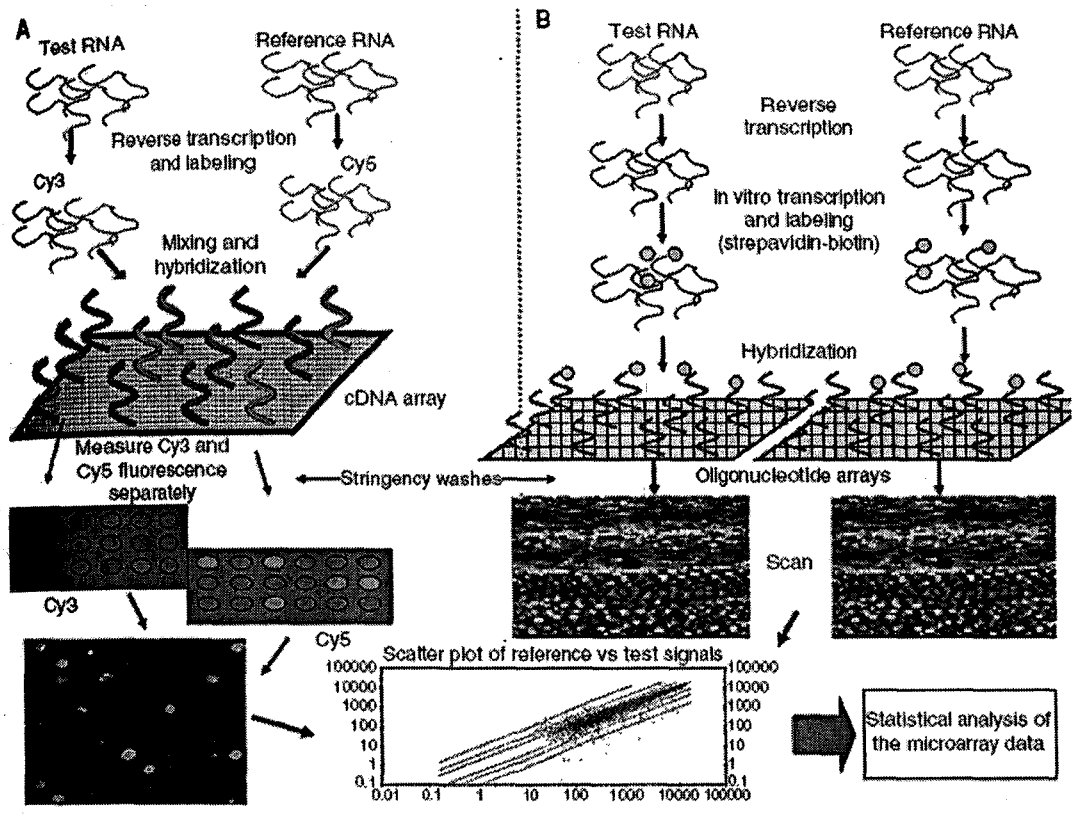
Source: Wilairat Leeanansaksiri- experiment on DNA microarray technique

Figure 2. DNA Microarray result of gene expression in stem cells image using Affimetrix GEM2 Chip.



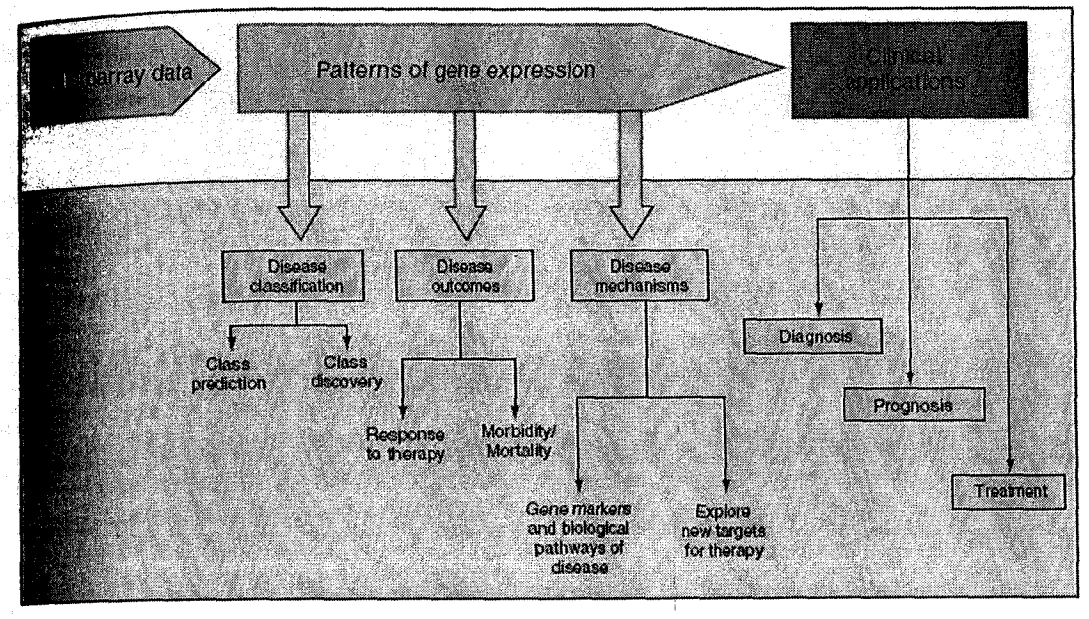
Source: Wilairat Leeanansaksiri-Experiment on DNA microarray technique.

Figure 2. DNA microarray result of stem cells. A. Stem cell development into myeloid cell lineages B. Microarray data and verification of microarray result by western blot and northern blot.



Source: Journal of Experimental Therapeutics and Oncology 3: 297–304, 2003

Figure 3. Scheme of the various steps involved in microarray experiment using different types of array platforms. RNA from test and reference samples are reverse transcribed, labeled by methods appropriate for the type of array used, and hybridized to microarray slides. In case of oligonucleotide arrays, another step wherein the cDNA is subjected to in vitro transcription to obtain the cRNA is performed. Following stringency washes, signals are detected by imaging techniques. (A) Representation of two-color experiment: cDNA arrays are most commonly used. More recently, long oligonucleotide (50–80 base pairs) has been used to conduct two-color array platform; (B) Representation of one-color experiment: oligonucleotide arrays are most commonly used.



Source: Journal of Experimental Therapeutics and Oncology 3: 297–304, 2003

Figure 4. Clinical relevance of microarray gene-expression profiling.

Part II

Principles of Techniques in RNA Study

Topic 8. RNA Extraction and Purification

Overview of cellular eukaryotic and prokaryotic RNA species

1. Structure of RNA

a) usually single-stranded
b) contains ribose instead of deoxyribose
c) contains uracil instead of thymine
d) can loop back and form partially double-stranded regions (especially in rRNA and tRNA)
e) dsRNA adopts A-conformation

2. DNA makes RNA makes protein (central dogma)

3. Three types of RNA:

RNA species	Eukaryote (amount, %, in cell)	Prokaryote (amount , % in *E.Coli* cell)
a) ribosomal (r)RNA	80-90	16
b) transfer (t)RNA		3
c) messenger (m)RNA	1-5	1

Most Important consideration of RNA Extraction:

- Rapid and simple preparation of full-length RNA
- Efficiently inhibit the endogenous ribonucleases (both endonucleases and exonucleases)
- Avoid contamination of ribonucleases of hands (ware gloves), glassware and solutions
- be careful of DNA contamination

RNase Inhibitors : RNasin, vanadyl-ribonucleoside complexesguanidine hydrochloride, guanidinium isothiocyanate, heparin and dextran sulphate etc.

Protein degradation and removal: Proteinase K, phenol and chloroform.

Why should we need to purify RNA ?

For cDNA libraies, cDNA cloning, Northern blotting, in vitro translation, analysis of RNA structure and synthesis which is useful in gene expression studies, and other RNA studies.

Ribonuclease-free conditions :

- separate the reagents and utensils to be used specifically for RNA work e.g. label "RNase-free" with a color tape
- Solution must be treated with diethylpyrocarbonate (DEPC) and effectively autoclaved before use
- Plasticware, glassware, utensile must be nuclease-free: treated with DEPC, autoclave, bake glassware at 250-300°C in hot air oven for at lease 4 hours

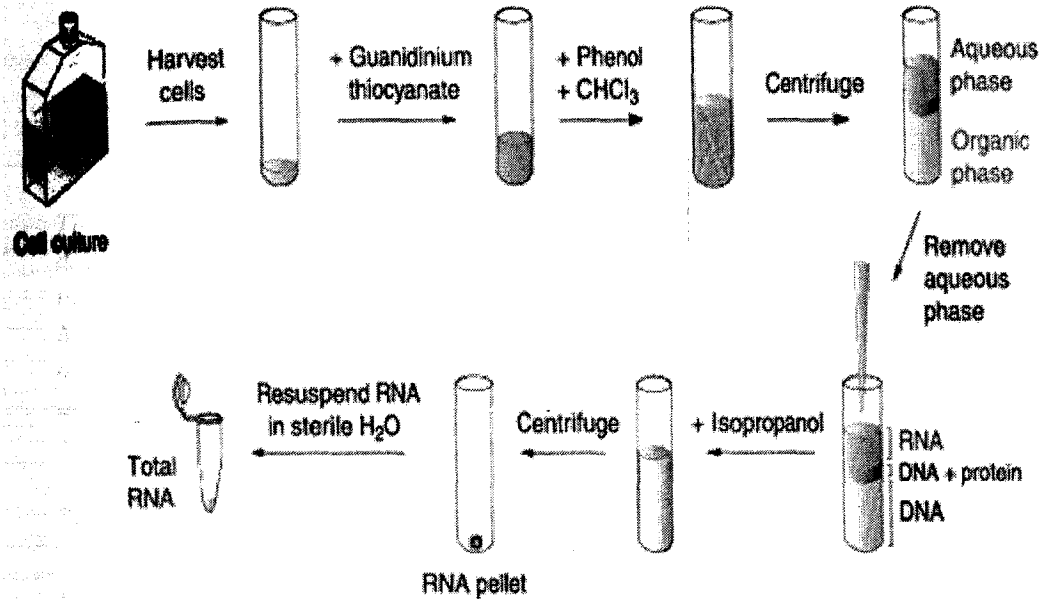
Note: Now many things, especially, plasticware for RNA works usually nuclease-free from the companies (so check before ordering)

RNA Extraction and Purification

1. Total RNA
2. Cytoplasmic RNA
3. Nuclear RNA
4. Messenger RNA (mRNA)

1. Total RNA Extraction

A. Guanidinenium-CsCl method
B. Phenol-chloroform method



(Company : Invitrogen Life Technologies)

Isolation of RNA from Small Quantities of Tissue (1 to 10 mg) or Cells (100 to 10,000) Samples

1. Add 800 μl TRIZOL to the sample. Homogenize cells by pipetting repeatedly. Add 200 μg glycogen directly to the TRIZOL reagent. If processing tissue, pulverize in liquid nitrogen first and then add 800 μl TRIZOL containing 200 μg glycogen (final concentration 250 μg/ml) followed by vigorous vortexing or power homogenization.
2. Place at room temperature, cap vial, and vortex at high speed for 10 sec. Make sure the TRIZOL reagent wets the side of the vial in order to solubilize any sample that may be remaining on the walls.
3. Shear the genomic DNA in the sample by passing twice through a 26-gauge needle connected to a 1 ml syringe. Using the syringe, transfer the sample to a sterile 1.5 ml microcentrifuge tube.
4. Add 160 μl of chloroform (or 49:1 chlorform:isoamyl alcohol) to each sample and vortex up to 30 s.
5. Spin at maximum speed in the microcentrifuge 5 min to separate the phases.

6. Transfer the upper aqueous phase to a fresh tube and add 400 μl ice-cold isopropanol. Allow the samples to precipitate at -20°C 1 h - overnight.
7. Pellet the RNA by centrifugation at maximum speed in the microfuge 15 min at room temperature.
8. Decant the supernatant.
9. Wash the pellet in 200 μl of 70% ethanol and spin again 10 min at maximum speed.
10. Decant the supernatant, removing as much as possible without disturbing the pellet.
11. Dry under vacuum and centrifugation 3 - 5 min.
12. Resolubilize the pellet in 30 - 50 μl RNAse-free deionized water. NOTE: If tissue is high in RNAses (e.g., adrenal gland, pancreas,) resuspend in 100% deionized formamide. (Alma M. Bracete, Donna K. Fox, and Domenica Simms, 1998, Isolation and Long Term Storage of RNA from Ribonuclease-rich Pancreas Tissue, *FOCUS*® 20:3, p. 82)
13. Be sure to vortex or pipette up and down the sample to ascertain that pellet is resolubilized fully.
14. Store at -70°C.

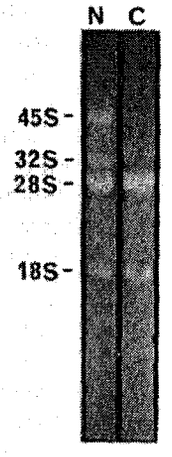
The glycogen remains in the aqueous phase and is co-precipitated with the RNA. It does not inhibit first-strand synthesis at concentrations up to 4 mg/ml and does not inhibit PCR. We obtained nanogram amounts of total RNA from as little as 100 cells. Microgram amounts of total RNA were isolated from 1 to 10 mg of tissue. No difference in yield of RNA was found when the volume of TRIZOL Reagent was decreased to 0.4 ml. However, decreasing the concentration of glycogen below 250 µg/ml TRIZOL resulted in lower and more variable yields of RNA.

2. Cytoplasmic RNA purification

RNA prepared in this way is an excellent template for the preparation of cDNA libraries, for cell-free translation and for primer extension and nuclease-S1 protection assays.

3. Nuclear RNA

Nuclear RNA is useful for studies examining precusor transcripts and RNA splicing intermediates

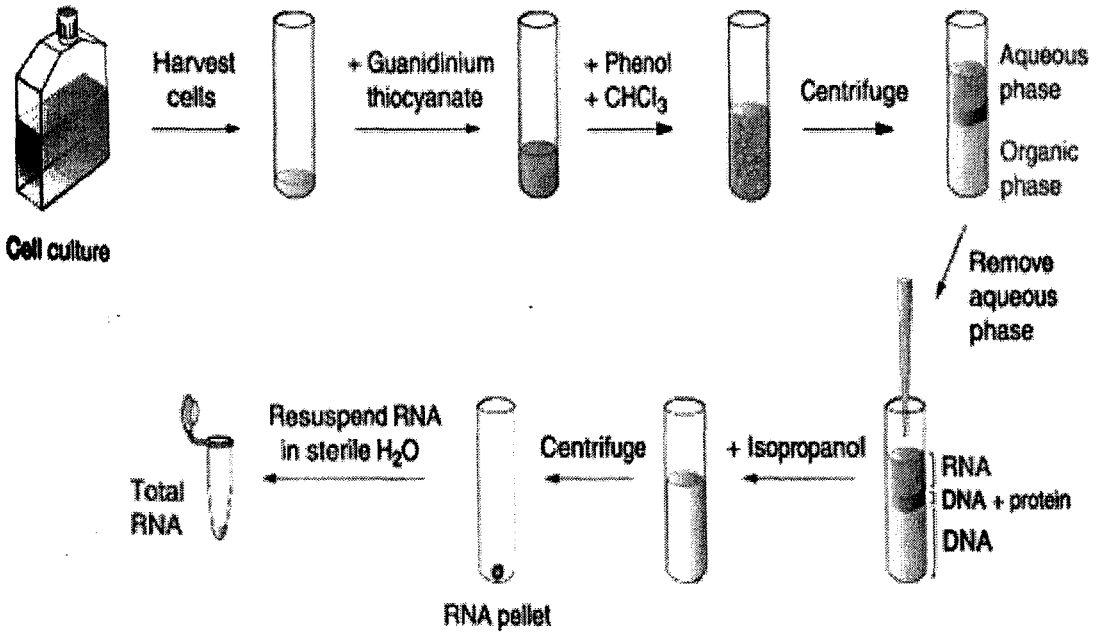


4. Purification of mRNA (sizes of less than 0.5 Kb to over 10 Kb)

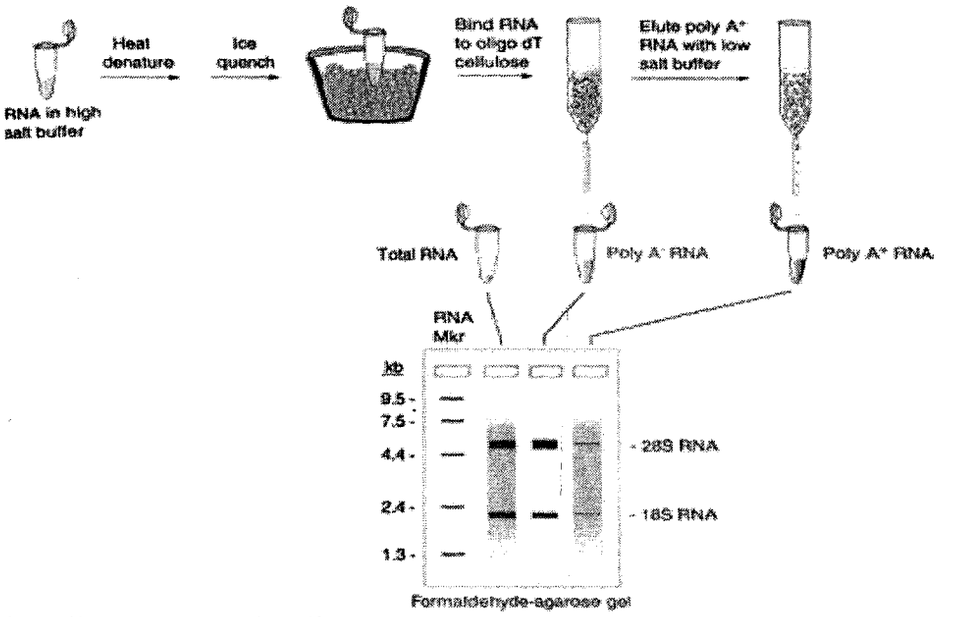
Since RNA is single-stranded, any hydrolysis event that breaks the phosphate backbone will result in cleavage of the molecule into subfragments. Two factors contribute to the biochemical **instability of RNA**.

1. Endoribonucleases (**RNases**) are very stable enzymes that cannot be easily inactivated . In fact, human hands are a rich source of RNase and it is therefore necessary to wear clean latex gloves during RNA isolation procedures and to use RNase-free labware.
2. RNA is thermodynamically less stable than DNA because of the 2' hydroxyl group on the ribose ring that **promotes hydrophilic attack** on the 5'-3' phosphodiester bond to form a **2'-3' cyclic phosphate**. This cyclic phosphate intermediate is stabilized by Mg++, a component of many biochemical reactions.

It is critical that the RNA be isolated intact and pure so that it can function as a **faithful template** for first strand synthesis. One method is to purify mRNA from tissue culture cells using guanidinium thiocyanate and oligo dT cellulose. Total RNA remains in the aqueous phase under acidic pH conditions following phenol:chloroform extraction. After precipitation with isopropanol, the RNA solution is adjusted to high salt (0.5M NaCl) and loaded onto an oligo-dT cellulose column.



Source: http://www.protocol-online.org



Source: http://www.protocol-online.org

Isolation of poly A+ RNA from Total RNA by Oligo(deoxythymidine)cellulose Chromatography

Total RNA is first ioslated from the tissues or cells and then mRNA is isolated by PolyA+ selection using oligo(dT) cellulose. This is necessary for sll tissue sources rich in RNase (and some cell lines)

Protocol

General

Homogenise tissues and isolate total RNA ~ 2 days depending on method

Isolate poly A+ RNA ~ 1day

Recover mRNA, quantitate and formaldehyde-agarose gel check

Reagents All reagents must be made with sterile MilliQ water. Use

DEPC with care - it may inhibit subsequent enzyme reaction

Sterile MilliQ water 10 M NaOH Proteinase K at 10 mg / ml in sterile water Oligo(dT) cellulose (Boehringer or equivalent) STE solution 100 mM NaCl, 10 mM TrisCl pH 7.4 and 1 mM EDTA pH 8.0 20% SDS in sterile water 5 M NaCl 1 M TrisCl. pH 7.4 (DEPC is unstable in Tris buffers) 500 mM EDTA pH 8.0 Binding

Methods

In advance

1 Prepare total RNA from tissues / cell lines

2 Prepare the oligo(dT)cellulose

Wash 1 g of oligo(dT)cellulose in 50 ml of 400 mM sodium hydroxide in sterile water for 30 minutes at room temperature on a rotating wheel. Neutralise by washing twice in 50 ml of 500 mM TrisCl pH 7.2, followed by 6 changes of sterile water. Wash in 50 ml of binding buffer to equilibrate the salt concentration, recover by centrifugation and then resuspend in 50 ml of binding buffer

Day '1'

5 Resuspend total RNA in up to 40 ml of 100 M NaCl, 10 mM TrisCl pH 7.4, 1 mM EDTA pH 8.0 and 0.5% SDS with 300 mg / ml proteinase K in a sterile 50 ml tube. Incubate for 30 - 60 minutes at 65oC 6 Adjust the NaCl concentration to 500 mM and add an appropriate amount of oligo(dT) cellulose in binding buffer 1 and the RNA sample allowed to bind for 2 - 4 hours at room temperature in a total volume of 50 ml on a rotating wheel 7 To remove unbound (mainly ribosomal) RNA, spin the RNA - oligo(dT)cellulose mix at 5 000 rpm for 5 minutes, decant the supernatant and resuspend in 25 ml of washing buffer warmed to 37oC. Mix on the rotating wheel for 15 minutes 8 Recover the RNA - oligo(dT)cellulose mix and wash again 25 ml of washing buffer warmed to 37oC. Mix on the rotating wheel for 15 minutes 9 Spin the RNA - oligo(dT)cellulose mix at 5 000 rpm for 5 minutes, decant the supernatant and recover bound poly(A+)RNA by eluting twice with 5 ml of 0.5% SDS in sterile water heated to 65oC. 10 Extract the eluate once with acid-phenol-chloroform. This removes any enzymatically active proteinase K and also any oligo(dT) cellulose that has been carried into the supernatant 11 Precipitate the RNA with one volume of isopropanol and 0.1 volumes of lithium chloride 2 and recover by centrifugation at 1200g for 10 minutes 12 Wash the pellet in 70% ethanol and then resuspend in 50 - 500 ml of 2 mM DTT, 1 u / ml RNasin in sterile Milli Q water 13 Quantitate by UV absorbance at 260 nm and check the ratio of UV absorbance at 260 and 280 nm

Day '2'

15 Check an aliquot of the RNA by ethidium-agarose-formaldehyde gel electrophoresis 16 Store the RNA at -70oC . 17 Regenerate the oligo(dT) cellulose. Wash it well in several volumes of elution buffer, re-treat with 400 M NaOH as in 4, spin down and resuspend in 500 mM TrisCl pH 7.4. Wash for 15 minutes, spin down, check the pH of the supernatant and resuspend again in 500 mM pH 7.4 if it is still alkaline. Resuspend the neutralised oligo(dT)cellulose in sterile water, wash to remove the salt, spin down and resuspend in absolute ethanol. Store at -20oC protected from light

Notes

1 Pancreas total RNA is best proteinase-K digested before oligo(dT) selection to remove any residual undenatured RNaseA. Other total RNAs can be safely resuspended directly in binding buffer without proteinase digestion. 2 The whole procedure can be performed in Econo columns. However, it is slower and the

resulting mRNA has more rRNA contamination in my hands than using the 50 ml tubes 5 Recover very small amounts of PolyA+ with 20 mg of tRNA and collect in the ultracentrifuge 6 To remove any DNA, resuspend the pellet at 12 in 10 mM TrisCl pH 7.4, 1 mM EDTA, 10 mM MgCl2, and add 1 unit of RQ1 RNase-free DNaseI (Promega) per 10 ml volume. Incubate at 37oC for 30 minutes. Stop the reaction with 10 mM EDTA and 0.2% SDS. Extract once against acid-phenol, re precipitate and recover by centrifugation.

mRNA Purification

- **I. Prepare oligo-dT cellulose**
  - use 40 mg oligo-dT cellulose / 1 mg total RNA
  - swell oligo dT-cellulose in elution buffer
  - wash oligo dT-cellulose 4 x with elution buffer (30 sec. full speed spin in between)
  - equilibrate oligo-dT cellulose with 2 to 3 washing steps using 1x binding buffer
- **II. mRNA purification**
  - bring 1 mg total RNA with H_2O to 600 µl
  - incubate 4 min at 65 °C
  - add 600 µl 2x binding buffer
  - add to 40 mg 1x binding buffer equilibrated oligo-dT cellulose
  - incubate 15 min at RT on a rolling incubator or vortex several times in between
  - spin oligo-dT cellulose down, discard supernatant
  - wash 2 x with 1x binding buffer
  - wash 2 x with wash buffer
  - elute with 250 µl elution buffer at 37 °C for 5 min
  - spin and keep supernatant (for second round of purification or precipitation)
  - elute oligo-dT cellulose again with 250 µl elution buffer
  - spin and keep supernatant (for second round of purification or precipitation)
  - combine eluate and add H_2O to 600 µl
  - repeat mRNA purification by starting again with 4 min incubation at 65 °C
- **III. Recover mRNA**
  - add **50 µl 4 M NaCl** and precipitate with **2 vol. cold EtOH**
  - incubate 1 h at -20 °C
  - spin 10 min full speed, wash with 70 % EtOH, air dry and dissolve in **20 µl H_2O**
  - read A_{260} of **1 µl** (100 to 250 fold diluted)
  - run **1 µg** (and 10 µg total RNA as comparison) on a 1.2 % agarose Formaldehyde gel

recovery = 1-2 % of the total RNA (10-20 µg mRNA / 2 mg total RNA)

Buffers:
[all buffers should be made with DEPC treated/autoclaved components!]

- **oligo dT cellulose (type7, Pharmacia)**

- **2 x binding buffer:**

1 M NaCl; 20 mM Tris pH 7.5; 2 mM EDTA; 0.1 % SDS
[**50 ml**: 12.5 ml 4 M NaCl; 1 ml 1 M Tris; 200 µl 0.5 M EDTA; 500 µl 10 % SDS]

1 x binding buffer:
0.5 M NaCl; 10 mM Tris pH 7.5; 1 mM EDTA; 0.05 % SDS
[**50 ml:** 6.25 ml 4 M NaCl; 500 µl 1 M Tris; 100 µl 0.5 M EDTA; 250 µl 10 % SDS]

wash buffer:
0.2 M NaCl; 10 mM Tris pH 7.5; 1 mM EDTA; 0.05 % SDS
[**50 ml:** 2.5 ml 4 M NaCl; 500 µl 1 M Tris; 100 µl 0.5 M EDTA; 250 µl 10 % SDS]

elution buffer:
10 mM Tris pH 7.5; 1 mM EDTA; 0.05 % SDS
[**50 ml:** 500 µl 1 M Tris; 100 µl 0.5 M EDTA; 250 µl 10 % SDS]

RNA Quantification and Measurement

Quantitation of RNA: A OD260/280 ratio of 1.8-2.0 is appropriate for pure RNA
RNA 1 OD = 40 µg/ml
(DNA 1 OD = 50 µg/ml)
Determination of DNA contamination: by gel electrophoresis

Precipitation of RNA

- precipitated in the presence of alcohol and salt
- standard method : one-ten volume of 3 M sodium acetate (pH 5.2) and 2.0-2.5 vol. Of ethanol, followed by vigorous mixing
- precipitation at -20°C
- time of precipitation and centrifugation depends on the amount of RNA to be precipitated
    1. RNA concentration in ethanol of greater than 10 µg/ml
       Precipitate for at least 20 min at –20 °C/centrifuge for 10 min
    2. RNA concentration in ethanol of less than 10 µg/ml
       Precipitate for at least 20 min at –20 °C/centrifuge for 30 min
    3. RNA concentration in ethanol of less than 10 ng/ml
       Precipitate for overnight at –20 °C to improve recovery

- Generally, RNA is centrifuged in 1.5 ml tubes in a microfuge at maximumspeed. If samples larger than 1.5 ml are to be centrifuged, several options are available
- Free nucleotides (e.g. redioactively labelled NTPs Are efficiently (>90%) removed by precipitation in the presence of 2.5 M ammonium acetate (0.5 vol. Of 7.5 M ammonium acetate) and 2.0-2.5 vol. Of ethanol

Storage of RNA

- stored at –70 °C in the presence of ethanol and salt is stable for several years
- stored at –20 °C is appropriate for a very short-term storage
- the best way to store is at –70 °C with aliquots

Topic 9. RNase Protection Assay (RPA) Technique

Mapping the transcription initiation site
4 methods

1. RNase Protection Assay (RPA) : RNA probe and DNA target
2. S1 Nuclease Analysis : DNA probe with RNA target
3. Primer Extension : DNA primer as a probe with RNA target
4. Rapid Amplification of cDNA Ends (RACE) : DNA primer and RNA target before performing PCR

RNase Protection Assay (RPA)

The ribonuclease protection assay (RPA) is a sensitive method used to detect and quantify specific mRNA transcripts in a complex mixture of total RNA or mRNA molecules.[2] The procedure is simple (Figure 4). An RNA probe is synthesized through an *in vitro* transcription reaction. These probes must be complementary to the gene sequence of interest and incorporate either radioactive or biotinylated nucleotides. The labeled probe is then incubated with a sample of total RNA or mRNA to facilitate hybridization of the complementary region of interest to the labeled probe. After hybridization, the mixture of single-stranded RNA and double-stranded probe:target hybrid is treated with ribonuclease (RNase), which digests all single-stranded RNA but no double-stranded RNA molecules. As a result, only the double-stranded gene target of interest remains. Usually, the sample is electrophoresed on a denaturing TBE-urea polyacylamide gel and detected by methods specific to the label on the probe.

RPA methods offer distinct advantages over other RNA detection methods such as Northern blotting and RT-PCR. In contrast to Northern blotting, the probe and target mRNA in an RPA reaction are hybridized in solution, where target availability is maximal, enabling detection of rare messages.[3-4] RPA methods also enable the researcher to probe for multiple transcripts in one sample of RNA. Quantitative RT-PCR which otherwise rivals RPA in sensitivity is not adaptable to this "multiplexing" technique. Northern blots can be probed for multiple transcripts, although multiple rounds of stripping and re-probing may be necessary,[5] and they are in general ten-fold less sensitive than RPAs.[3] Because of their high sensitivity and resolution, RPAs are also well suited for mapping internal and external boundaries in mRNA (refer to the SuperSignal RPA III Kit (Pierce) for more information on mRNA mapping).

The SuperSignal RPA III Kit (Product # 89832,) offered by Pierce combines patented procedures to facilitate simultaneous RNase inactivation and RNA precipitation with patented SuperSignal Detection Technology to achieve an efficient and sensitive chemiluminescent RPA. The SuperSignal RPA III Kit uses biotinylated rather than radiolabeled RNA probes.

Probe Preparation for RPA

Probe Design

Probes used in RPAs must exhibit specific characteristics. First, RPA probes must be synthesized to include ~10% extra sequence on either side of the binding sequence of interest. This is an important control to ensure that the RPA, specifically the RNase, is working correctly. When a probe with extra sequence is hybridized to its complementary strand, the extra sequence does not hybridize, remains single-stranded and is, therefore, digested along with the non-complementary RNA. When the undigested, labeled probe is run as a control, it appears slightly larger than the protected, digested fragment providing an indication of the RNase's effectiveness when electrophoresed (Figure 5, Step 1-3). Second, probes need to be labeled for visualization after electrophoresis. Probes used in RPAs can be labeled with a radioactive nucleotide, such as [a-^{32}P]CTP, or a biotinylated nucleotide, such as Biotin-N^4-CTP.

Radiolabeled Probes

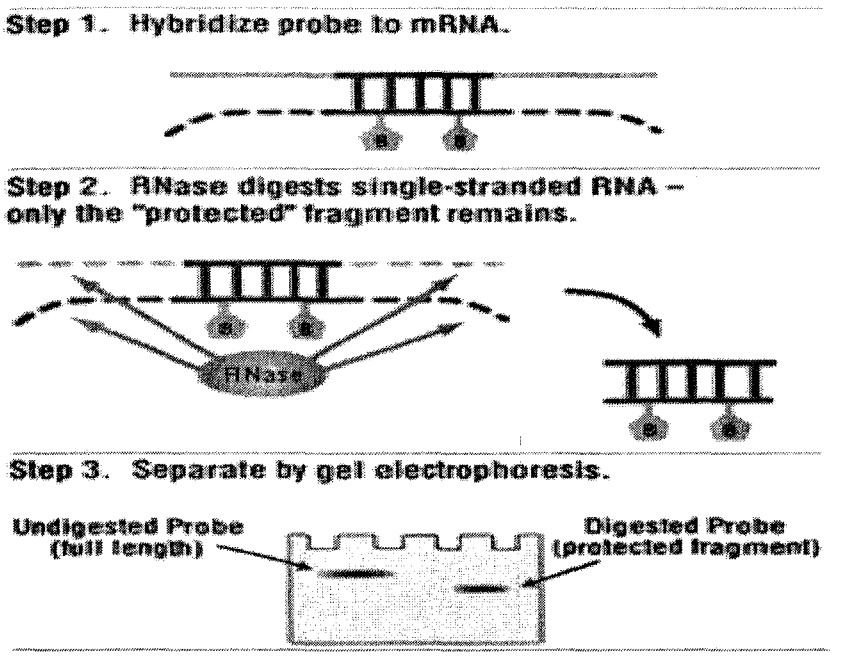
Radioactive probes of high specific activity are prepared by *in vitro* transcription reactions using T7, T3 or SP6 RNA polymerase. In these reactions, all of the non-labeled CTP is completely replaced with [a-^{32}P]CTP. Reactions consist of the ribonucleotides, RNA polymerase, antisense DNA template, a transcription buffer and nuclease-free water. After synthesis, the excess original template is digested with DNase.[1] Gel purification is necessary when working with radioactive probes in order to eliminate unincorporated nucleotides that can cause smearing in the final RPA gel analysis. Radioactive probes can be purified by running the reaction on a TBE-urea polyacrylamide gel, excising the probe bands from the gel, and eluting the nucleic acids from the excised gel plugs with probe elution buffer. Purified radioactive probes can by quantified by liquid scintillation counting.

Biotin-labeled Probes

Biotinylated probes are prepared by similar in vitro transcription reactions, with several exceptions. First, instead of the complete substitution of non-labeled CTP, ~60% of the non-labeled CTP is mixed with ~40% Biotin-N^4-CTP. Second, biotinylated probes are commonly precipitated with LiCl, collected by centrifugation, and washed once in ice cold 70% EtOH. Gel purification of biotinylated probes is not necessary because any unincorporated nucleotides remaining after DNase treatment are not retained on the nylon membrane during subsequent steps. The precipitated probes are re-suspended in diethyl pyrocarbonate (DEPC)-treated water and quantified spectrophotometrically at 260 nm.

Although radiolabeled probes currently are used more commonly for RPAs, they exhibit some disadvantages when compared to biotinylated probes. Radioactive probes must be re-synthesized weekly to guard against short isotope half-life as well as radiolysis. Radioactive probes require lengthy exposure times, often overnight, in order to visualize results. Since RPAs frequently must be optimized for the researcher's specific target RNA and probe, lengthy exposure times common to radioactive RPAs significantly increase the amount of time necessary to generate

valid RPA results. Biotinylated probes, on the other hand, can be used without gel purification and are stable for years, as opposed to weeks.



Source: Pierce product protocol

Figure 1. RPA procedure using the SuperSignal RPAIII Kit

Hybridization and RNase Digestion
Hybridization of labeled probe to the complementary sequence of interest is accomplished by combining a molar excess of probe with the target RNA. A molar excess of probe compared to target is necessary for quantitative detection of mRNA. Each probe/RNA reaction is hybridized 4 to 16 hours at a temperature that is optimal for the specified probe composition in question, commonly between 42°C and 56°C. Following hybridization, each reaction is digested using either RNase T1 or RNase A/T1.
(back to top)

Electrophoresis
After RNase inactivation and RNA precipitation, the RNA pellet is re-suspended in loading buffer, vortexed and heated. Reactions are loaded onto a TBE-urea gel (~6% polyacrylamide to resolve 100-500 nt fragments) and electrophoresed.
(back to top)

Transfer and Detection
With radioactive detection, gels are dried and then exposed to X-ray film for as long as necessary to visualize the results. The exposure time required depends on the specific activity of the labeled probe, as well as the abundance of target. Normal times range from 4 hours to overnight.

For chemiluminescent detection, the gel is transferred to a positively charged nylon

membrane. This transfer can be done in a number of ways including diffusion, capillary, heat accelerated convection, vacuum blotting and electroelution. The transfer method most commonly used for RPA is electroelution or electrophoretic transfer because of its speed and transfer efficiency. Transferred RNA can be cross-linked to the membrane using an UV cross-linker.

The blot is first incubated in protein blocking buffer and then probed with streptavidin-HRP conjugate diluted in blocking buffer. Streptavidin-HRP binds to the biotin label, and the bound horseradish peroxidase (HRP) catalyzes the production of light upon addition of the luminol-based detection reagent. After incubation with the streptavidin-HRP, the membrane is washed and incubated with highly sensitive, chemiluminescent SuperSignal Substrate. Finally, blots are covered with plastic wrap and exposed to X-ray film or a CCD camera. Typical exposures range from 1-60 minutes.

Advantages:

1. The most sensitive of the quantitative methods for start-site locallization
2. RNA-RNA hybrids are thermodynamically stable, much more than RNA-DNA hybrids in S1 nuclease analysis
3. Use to study gene expression levels

Disadvantages:

1. Time consuming than primer extension
2. Optimization of hybridization and digestion must be determined which can be difficult to establish for a new gene

Methods :

1. RNA sample preparation : can use both total or mRNA
2. RNA probe preparation : *in vitro* transcription
3. RPA reaction
4. Acrylamide gel electrophoresis
5. Phosphoimager or autoradiograph

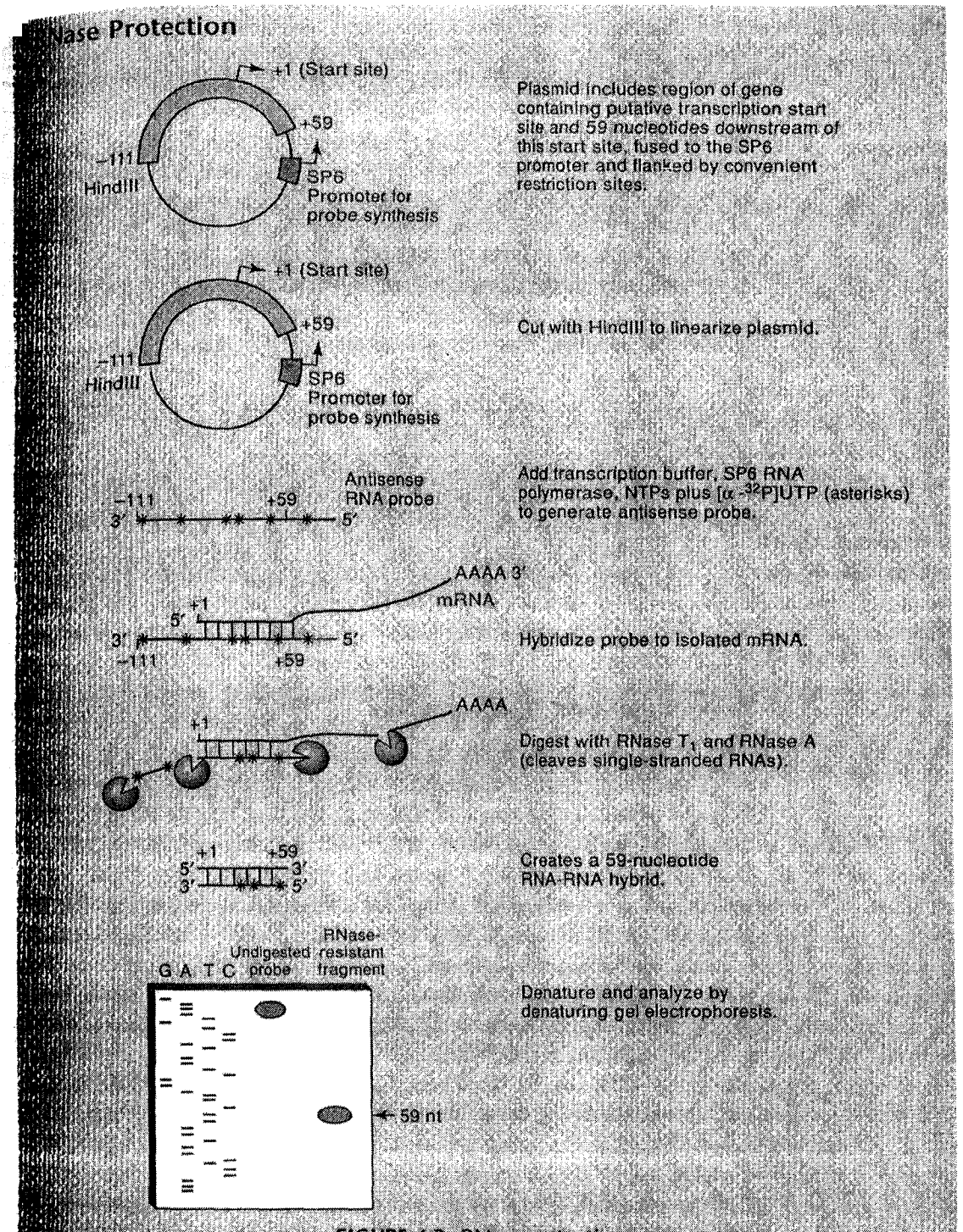


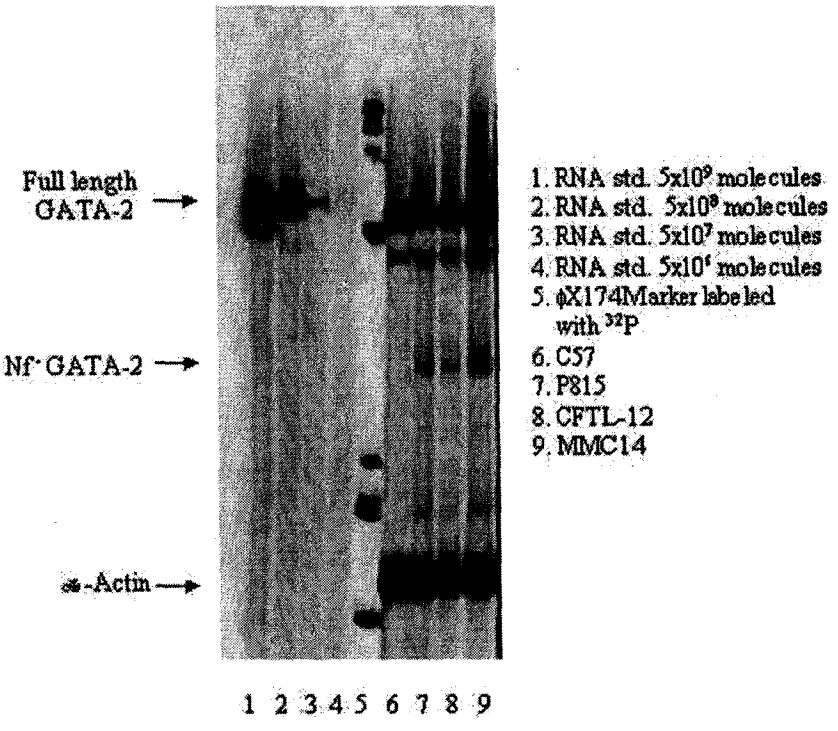
FIGURE 4.3. RNase protection.

RNase protection procedure (Fig. 4.3) begins with an RNA probe that is uniformly labeled incorporation of one [α-^{32}P]NTP, usually [α-^{32}P]UTP. The RNA probe is synthesized by bacteriophage RNA polymerase (SP6, T7, or T3), which initiates transcription from specific promoters that have been engineered into a number of common plasmid vectors. For site mapping, the plasmid template contains a genomic DNA fragment spanning the region thought to contain the transcription start site for the gene of interest. This genomic frag-

Source: Short Protocols in Molecular Biology, 3rd Edition, Fred Ausubel et al 1995
Figure2. Principle of RPA

analysis

See attached protocol for more detail

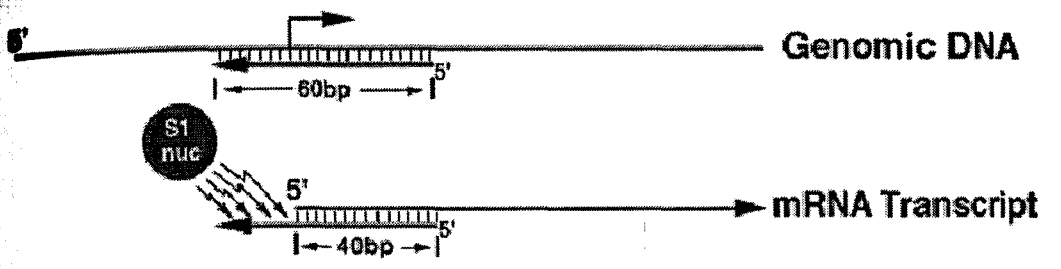


Source: Wilairat Leeanansaksiri-Experiment on quantitative RPA

Figure 3. Quantitative RPA with internal β-actin control

Topic 10. S1 Nuclease Analysis Technique

S1 Nuclease Analysis



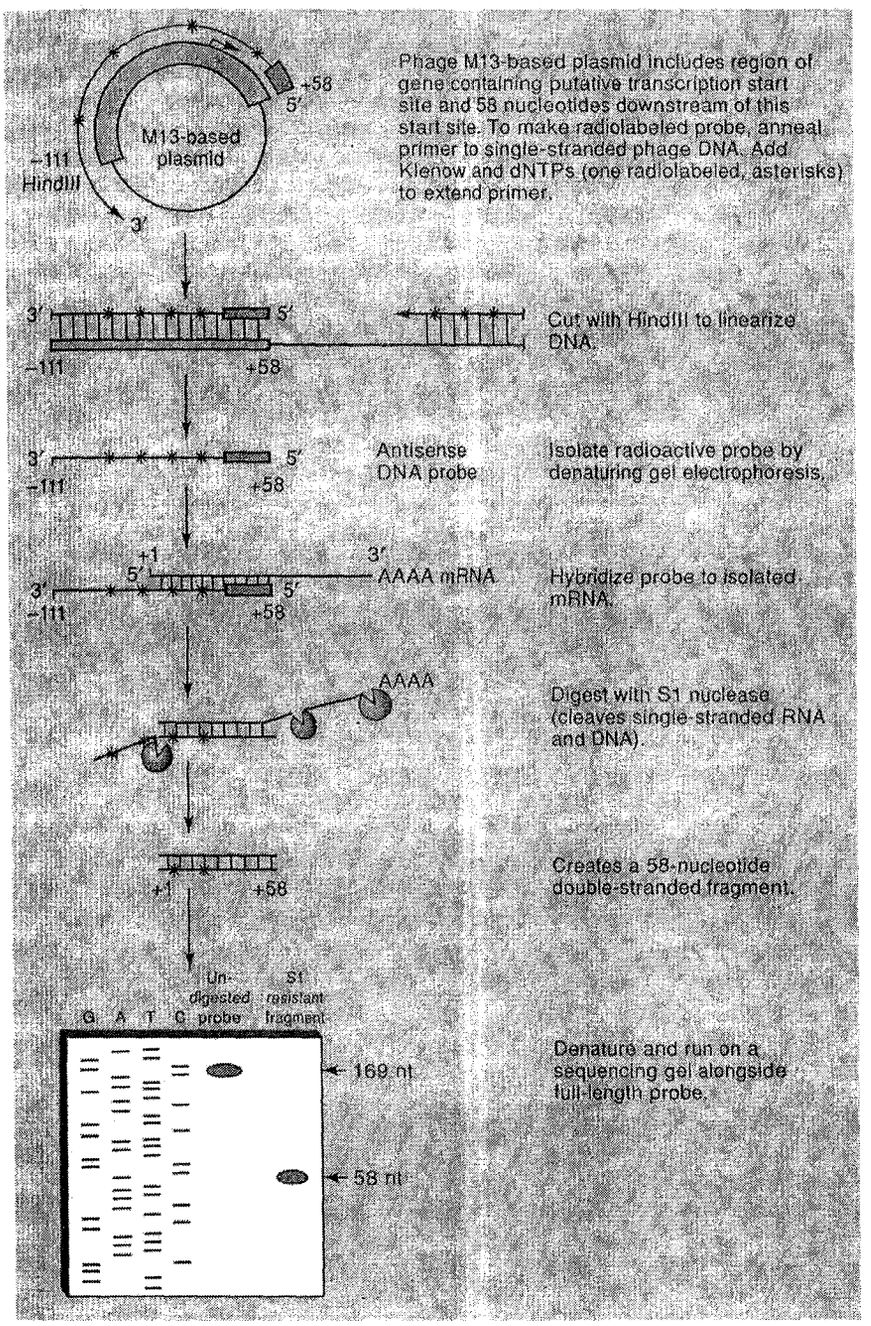
Experimental design considerations for promoter studies
- Used to quantitate transcript levels by assessing the number of 5' ends of the RNA of interest.
- 60 bp oligonucleotide should be designed that span the promoter initiation site in an anti-sense direction [the 5' 40bp of this oligo should hybridize to the mRNA of interest while the remaining 20bp should be complimentary to the 20bp sequence immediately upstream from transcription initiation site (which will not be contained in the mRNA) – this allows the detection of a specific cleaved product upon digestion with S1 nuclease following hybridization with RNA].
- Method typically works well using total RNA (typically prepared using tri-reagent).

Protocol
I. Kinase Oligonucleotide
1ul Oligonucleotide (100 ng/ul)
1.5ul Polynucleotide Kinase Buffer (10X)
1ul Polynucleotide Kinase
6ul [gamma-32P]ATP
5.5ul H 2 O
15ul Total
- 37oC, 1 HR.
- Add 135ul TE
- Run over G25 Spin Column
II. Hybridization
- ETOH Precipitate 50ug Total RNA/sample.
- Wash with 800ul 70% EtOH.
- Dry briefly and suspend in 15ul 2X Hyb. Buff [suspend by mixing with finger a few times followed by letting soln sit for a few minutes (repeat 3-4 times)].

- Add 15ul Formamid and 3ul of labeled oligonucleotide.
- Heat to 68oC for 10 min.
- Transfer directly to 37oC water bath and incubate overnight with glass plate (or cover) over water bath
(covering the water bath prevents evaporation of sample which condenses on top of eppy tube)
III. S1 RXN.
- Make up S1 digestion mix (500 units/ml; 0.5ul S1/1ml of 1X S1 Buffer [BRL S1 nuclease,
20000u/20ul, CAT# B001SA]).
- Add 300ul S1 RXN. mix per sample,
- Mix by vortexing on low speed for a couple seconds and spin briefly (10 sec on high speed).
- Incubate 30 min. at 37oC.
- Phenol/CHCl3 extract(300ul).
- Add1ml ETOH.
- Spin 30 min on high.
- Wash with 900ul 70% ETOH.
- Dry Briefly.
- Suspend in 10ul Sequencing Stop Buffer.
- Heat @ 68oC for 10 min and load 4ul onto 8% sequencing gel [as a control, load a small amount of the
original undigested probe in a separate lane (for control, dilute 1ul of original probe in 100ul TE, mix,
and transfer 1ul of this solution into a new tube with 8ul of sequencing loading dye – load 4ul of this
onto gel)].

Buffers for S1 Nuclease Analysis

10X Polynucleotide Kinase Buffer (-ATP) 10ml
0.5 M Tris (pH 8.0) 2.5ml 2M Tris (pH 8.0)
70 mM MgCl2 700ul MgCl2
10 mM DTT 100ul 1M DTT
H2O 6.7ml H2O
2X Hyb Buffer 10 ml
40 mM Pipes, 6.4 800 ul 1M Pipes, 6.4
1 mM EDTA 40 ul 0.5M EDTA
0.4 M NaCl 1.6 ml 5M NaCl
0.1% SDS 200ul of 10% SDS
H2O 7.36 ml H2O
S1 Buffer 50 ml
0.28 M NaCl 2.8 ml 5M NaCl
50 mM NaAc, 4.6 2.6 ml 1M NaAc, 4.6
4.5 mM ZnSO4 2.25 ml 0.1 M ZnSO4
H2O 42.4 ml H2O



Advantages:

1. background signals are often diminished

Disadvantages:

1. Less sensitive than RPA
2. Less stable of RNA-DNA hybrids than RNA-RNA hybrids in RPA
3. Probe preparation from single-stranded M13 plasmids or phagemids is more complicated and variable than those used to prepare RPA probe

Method:

1. RNA sample preparation
7. DNA probe preparation
8. S1 nuclease reaction
9. Acrylamide gel electrophoresis
10. Phosphoimager or autoradiograph
11. Data analysis

See attached protocol for more detail

Dot and slot hybridization : An excess of radiolabeled probe is hybridized to RNA that has been immobilized on a solid support (Kafatos et al. 1979; Thomas 1980; White and Bancroft 1982). Densitometric tracings of the resulting autoradiographs can allow comparative estimates of the amount of the target sequence in various preparations of RNA.

Now: less popular
: new techniques have been developed

Topic 11. Phage Display Technique

Definition of Phage and Phage Display

Phage

Phage or bacteriophage is a virus that infected only bacteria, this virus is not infectious to humans. Phage are made up of two components : genetic material and a protein coat. Phage genome cause infected bacteria to make more phage. The bacteria used "gene 3" to produce protein which assembles on one end of the phage. Each phage particle surrounded by a protein coat. The phage genome contained a recipe for making all of these protein. The protein coat protects DNA when the phage go from cell to cell.(1)

Phage are tailed , cubic , filamentous or pleomorphic and contain single-strand or double-strand DNA or RNA. They are classified into 13 families 1 genera. Tailed phage are more numerous than other type, are enormously diversified and seem to be the oldest of all phage group.

Bacteriophage occur in over 140 bacterial genera and many different habitats. Infection result in phage multiplication or the establishment of lysogenic or carries states. Bacterial gene may be transmitted in the process. Some phage (e.g. T4 , T7 , MS2 , fd and λ) are famous experimental models. Phage research has led to major advances in virology , genetics and molecular biology.

Phage taxonomy

Early attempts at classification by serology , host rang and inactivation tests show that phages were highly diversity, but these attempts proved premature. Phage were grouped into six basic types on the basic of morphology and mature of nucleic acid.

Phage families and genera

1. **Tailed phage** : with approximate 4905 observations, tailed phage comprise 96% of phage and are the largest virus group known. They contain a single molecule of dsDNA and are characterized by a tubular protein tail, a specialized structure for the transfer of phage DNA into host bacteria. sample of tailed phages e.g. T4 , T7 , λ etc.
2. **Cubic, filamentous and pleomorphic phage** : this group includes 10 small phage families that correspond to approximately 4% of phages, differ greatly in nucleic acid nature and particle structure, and some time have a single member. Filamentous phage have according to present knowledge, helical symmetry. Sample of this phage group e.g. PM2, PRD1 , fd, TTV1 , MVL2 SSV1 etc.
   (Ackermann, H.W., 2003)

Phage display

Phage display is a powerful tool that extends the range of modern combinatorial screening techniques, allowing the discovery and characterization of proteins that interact with a desired target. Phage display is a system in which a protein is displayed on the surface of a phage as a fusion with one of the coat proteins of the virus. The DNA that encodes this protein is housed within the virion. By cloning large numbers of DNA sequences into the phage, display libraries are produced with a repertoire of many billions of unique displayed proteins.(2)

In phage display new genetic material is inserted into the phage gene. The bacteria process the new gene so that the new protein and peptide is made. This protein and peptide is expose on the phage surface.

Seven Step of phage display (1)

1. **Gene Insertion** : phage display begins by inserting a diverse set of gene into the phage genome. Each phage receives a different gene. Generally, inserts the diverse of gene into the phage's existing gene 3 because it is the site that peptide can be expressed on coat protein of phage.
2. **Protein display** : the modified gene 3 contains and added segment : an antibody , small protein or peptide which is expressed on the surface of the phage. Each phage receive only one gene, so each expresses a single protein or peptide. Collectively, the population of phage can display a billion or more protein or peptide, each tied to its own gene.
3. **Library creation** : A collection of phage displaying a population of relates but diverse protein or peptide is called ''library". The related protein keep most of the physical and chemical properties of their parent protein. These protein are initial leads to providing medical therapies that may cure certain disease.
4. **Target Exposure** : The library is exposed to an immobilize target which is disease causing molecule, such as a receptor or enzyme.
5. **Binding** : When the library is exposed to a target. Some member of library will bind to the target through an interaction between the displayed molecule and the target itself. Huge genetic diversity in the displayed protein increases the likelihood that some phage will bind well or very well to the disease target to interest. Drugs usually act by binding disease causing molecules and changing their behavior. The protein isolated by phage display are drugs candidates because of their tight and specific binding to disease targets.
6. **Isolate** : After give the phage a chance to bind to the target, the immobilized target were wash to remove phage that did not bind out of the reaction.
7. **Amplification** : Phage display allows rapid discovery of drug candidates because only bound phage from the library that were captured to use for amplifying. Replicating this phage in bacteria increases the amount of this phage several million fold overnight, providing enough material for sequencing. Sequencing of the phage DNA can be told: what is the active compound. Good candidate will undergo additional testing and development, usually recover 10 to

1,000 candidate binders from an initial diversity of more than 1 billion. Developing extensive capability to select and tests active compound through automation.

Proteins that are created and isolated by phage display process have a specific reaction with a known disease target, making this a rapid, effective and focused drugs discovery method. The protein are candidates for effective drug therapies.

Phage display is a fundamental tool in protein engineering. The directed evolution of proteins using display methods can be engineered for specific properties and selectivity. The new methods involved in protein detection, amplification, and directed evolution. This is essential to the proteomics effort aimed at identifying, mapping, and understanding all human proteins in a comprehensive manner. A variety of display approaches are employed for the engineering of optimized human antibodies, as well as protein ligands, for such diverse applications as protein arrays, separations, and drug development. The use of phage display in screening for novel high-affinity ligands and their receptors has been crucial in functional genomics and proteomics. Display methods will make it possible to target essential components and pathways within many different diseases, including cancer, AIDS, cardiovascular disease, and autoimmune disorders.

The phage display process generally consists of the following steps:

(1) generating one or more phage libraries,
(2) selecting binding compounds with high affinity and high specificity to a target from the phage display libraries, and
(3) producing and evaluating the selected binding compounds.

Application (Phage display meeting : Cambridge Healthtech Institute)

Phage display Technologies (3)

1. Peptides from Phage Display Libraries as Tools for Drug Discovery
Hamilton,P.T.(2002)

High-affinity peptide probes were isolated directed to biologically relevant sites on target proteins. These peptides can be used as tools for proteomic analysis and in drug discovery. Peptides were used for format HTS-compatible assays for targets that are difficult to screen by traditional HTS methods and identified active compounds. Target-directed peptides were expressed inside cells. In addition, Protein conformation-dependent peptides were identified for nuclear receptors that can be used in an in vitro assay to predict the biological response produced by compounds.

2. Peptide Phage Display for Vaccine Development on Anti-HIV Antibody
Scott,J.K.(2002)

peptide that binds tightly and specifically and block interaction of human monoclonal antibody b12 which can be neutralizes a broad range of HIV-1 primary isolates. The structure of b12 antibody in complex with peptide reveals that the

peptide, in part, mimics critical aspects of the b12 epitope on gp120. Immunization studies aimed at eliciting b12-like, neutralizing antibody responses with the peptide.

3. Profiling the Cancer Immunome Using Phage cDNA Display

Somers, V.(2002)

"Serological Antigen Selection," as a procedure for the rapid generation of a profile of tumor antigens corresponding to the humeral immune response in cancer patients. Serological Antigen Selection involves the display of cDNA expression libraries on filamentous phage and the selection with serum from cancer patients. Different antigen of cancer patients were identified and analyze detailed serological analysis with sera from colorectal cancer patients and healthy individuals. This analysis showed that six antigens had preferential reactivity after vaccination with autologous tumor cells. This shows the potential for monitoring of tumor cell base vaccination trials.

4. Polyclonal Antibody Libraries for Cancer and Infectious Diseases

Sharon, J.(2002)

A polyclonal antibody library (PCAL) is a standardized mixture of antibodies specific for a multiantigen target, for which the genes are available and therefore the mixture can be perpetuated, amplified, and modified as desired. A PCAL contains both target-specific and cross-reactive antibodies but is designed to recognize the antigenic profile of the target over the antigenic profiles of cross-reactive multiantigens, with a high signal-to-noise ratio. PCALs are generated by positive +/- negative selection from Fab phage display libraries and mass transfer of the selected heavy- and light-chain variable region gene pairs from the phage display vector to a mammalian expression vector to produce full-length glycosylated antibodies. Developing a system for PCAL generation and have selected Fab phage display libraries reactive with human colorectal cancer cells and with the protozoan parasite Cryptosporidium parvum. These libraries were shown to be diverse at the DNA level by fingerprint analysis and at the antigen level by immunoblot analysis and are now being transformed into full-length antibody libraries.

SELECTION OF DRUG-LIKE COMPOUNDS

5. Vascular Antigen Discovery by in vivo and ex vivo Biotinylation and Antibody Phage Technology

Neri, D.(2006)

The antibody-based delivery of therapeutic agents at sites of disease is an attractive avenue for the development of biopharmaceuticals with improved potency and selectivity. The use of in vivo and ex vivo biotinylation procedures, followed by a mass-spectrometry-based comparative proteomic analysis of biotinylated samples, allows the identification of novel markers of pathology, which are easily accessible from the bloodstream and thus ideally suited for antibody-based targeted pharmaceutical applications.

6. Attaching Biologics to the Surface of Medical Devices via Phage Display Technology

Hamilton, P.(2006)

We are developing bifunctional peptide linkers that can be used to immobilize cells or protein therapeutics onto the surface of a medical device. The bifunctional linkers consist of 2 binding domains, each discovered by screening phage displayed peptide libraries. We have used these peptide linkers to rapidly attach cells to metal surfaces for improved bone integration of orthopedic implants. In addition, we have developed linkers that bind and retain Bone Morphogenetic Protein-2 (BMP-2) on a collagen sponge and have shown that these peptide linkers improve bone healing. These peptide linkers offer a robust, biofriendly means of functionalizing the surface of any medical device.

7. CDR Repair , a New Approach to Antibody Humanization

Dennis, M.(2006)

CDR Repair is a new phage display-base approach that enables rapid antibody humanization. This method differs from the traditional approach to humanization which grafts CDRs to the most homologous germline framework or incorporates mouse residues in the framework in order to re-establish antigen binding. CDR Repair-base humanizations utilize a consensus variable domain framework and change are recruited within the CDRs (rather than framework) to restore antigen binding. Method along with several example.

8. Engineering Speificity Changes in Protein-Protein Interactions using Phage Display.

Palzkill, T.G.,(2006)

Beta-lactamase inhibitory protein (BLIP) interacts with several different class A beta-actamase with varying affinities and is therefore a good model protein with which to study and engineer specificity determinants of protein-protein interactions. A phage display system for BLIP was previously developed and used to evolve a BLIP derivative that binds TEM-1 beta-lactamase with a Ki of ~ 10 pmol. This system has been modified for use in a capture phage display method to engineer BLIP derivatives that exhibit large changes in binding affinity and thus specificity for several target beta-lactamases.

9. Re-Engineering a Protein-DNA Interface using Unnatural Nucleotides.

Simon, M.(2006)

The adaptability at a protein-DNA interface was explored using phage display to re-engineer a homeodomain to specifically recognize a chemically synthesized, unnatural nucleotide. The engineered homeodomain preferentially binds to the modified DNA with affinity and specificity similar to those of natural homeodomains. High resolution x-ray structures of the homeodomain bound to modified and unmodified DNA reveal that the engineered homeodimain, dispite these significant perturbations, recognizes its target using interactions similar to wild-type homeodomain.

10. Synthetic Antibodies Specific for Different Lysine-Liked Forms of Polyubiquitin Developed vai Phage Display.

Gordon, N.C.(2006)

Phage display antibody libraries with restricted amino acid diversity in the complementarity determining region have been used to generate high-affinity.

Specific antibodies against validated targets. This approach may be successful used restricted diversity synthetic libraries to develop high-affinity, specific antibodies capable of binding and discriminating between different lysine-linked forms of polyubiquitin.

Advantage

Advantage of phage display

1. Phage display technology, when apply to antibodies also offered the potential of screening large number of antibody clone. (4)
2. M13 and the closely related filamentous bacteriophages fd and f1 are non-lytic, meaning that they do not lyse the host during phage production. This greatly simplifies the intermediate phage purification steps between rounds of panning, as a simple PEG precipitation step is sufficient to separate the phage from almost all contaminating cellular proteins. (5)
3. Multiple antibody phage display library increase probability of technique success in finding a high affinity antibody to the target of interest while minimizing development time through screening library in parallel.
4. There are ability to select an exclusive or nonexclusive library for antibody discovery to the target of interest.
5. A key advantage of T7 phage display for isolating specific RNA-BP cDNAs is the speed of the selection process. Two cycles of selection can be performed per day making it possible to isolate a unique cDNA clone in as little as 2-3 days, a significantly shorter time than is required for other cloning methods.
6. phage display is that selection is performed *in vitro*, where conditions for the RNA-protein interaction can be controlled precisely.

Disadvantage

1. The original of the antibodies heavy and light chain which contribute to the antibodies binding affinity, is lose the cloning process and a probability of restoring these original pairs is extremely low, As a result significant manipulation may be require to achieve affinity level to appropriate for clinical trails. (4)

2. half of the phage particles, which are produced after super-infection, can carry only the helper phage genome instead of the ligand gene, even though they display a specific ligand protein. These phages can compete with those which carry both the ligand gene and its protein during the selection process and consequently they loose the ligand protein during the second round of selection. In any given library, this problem can cause the loss of rare but functionally very important ligands during the sequential selection procedure.

3. reinfection of a super-infected bacteria during or after the super-infection. This can increase the frequency of those phage particles which only carry the helper-phage genome, in the total phage population.

4. since for ethical reasons it is not practical to immunize human and the phage library is derived from human immune tissue.

5. The library is typically nonimmune. Therefore antibodies response cannot be driven toward high specificity and high affinity and is instated a result of random combination and mutagenesis. As a result the antibodies derived from nonimmune human antibody phage display library are typically of substantially lower affinity than those derive from immunized donors.

6. Phage display that used tailed phage e.g. T7, T4, lambda phage which they are all lytic phage, necessitating additional time-consuming purification steps between rounds to avoid panning amplified phage in the presence of cellular proteins (including proteases which can degrade interested target during panning). (5)

The Product for Phage Display Technology

$QP^{Display}$ (by Genetix). identification and picking of single colonies, panning, re-arraying of precious positives and the ability to track plate identities.

Platform (Cat X9500) Is supplied as standard with:

- Organism specific 96 pin picking tool
- Two stacker lanes, configured for standard or deep well plates
- Imaging area for two 22cm trays or ten 10cm round Petri dishes
- Biowork table Full control software
- Data tracking kit

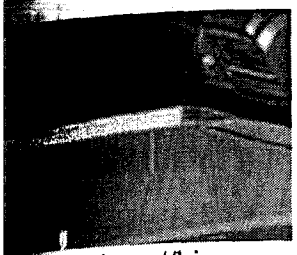
Options:

- Re-arraying & replication kit
- Membrane gridding kit

Workflow

Successful application of Phage Display requires reliability, precision and accuracy. Genetix is recognised as the world leading supplier of colony picking and library re-arraying systems.
$QP^{Display}$ is the dedicated system for management of the crucial later stages of antibody development.
Plate:
Enriched library selections are plated to obtain single colonies

Picking



Source: http://biocompare.com

The highest performance in single colony identification, picking accuracy and throughput.

The **QP**$^{\text{Display}}$ is configured with world-leading picking technologies that have been established through more than fifteen years experience and have delivered proven results when it matters, time and time again.

Individual phagemid colonies or plaques are imaged and picked into bar coded micro-well plates.

The Picking Process

Source: http://biocompare.com

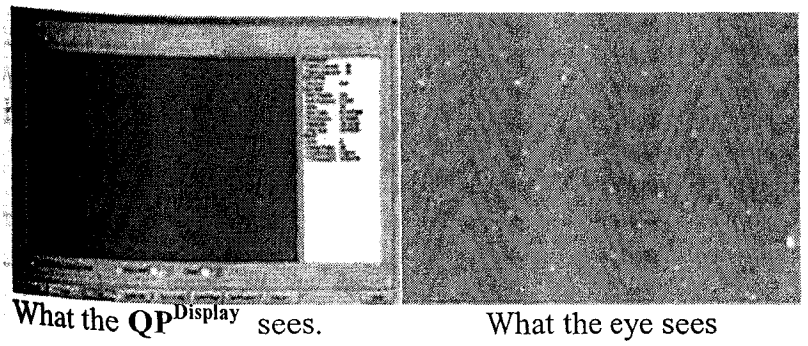
http://biocompare.com

After overnight incubation two 22cm picking trays, or ten 10cm Petri dishes are placed on the instrument light table. Lidded 96- or 384-Well plates are placed into the instrument stackers.

After selecting 'run', trays are automatically imaged and analysed. Colonies are selected on the basis of size, shape, colour and proximity to neighbours.

If desired, the images can be reviewed and colonies can be selected or de-selected on an individual basis.

How the QP$^{\text{Display}}$ **decides what to pick** (Source: http://biocompare.com)



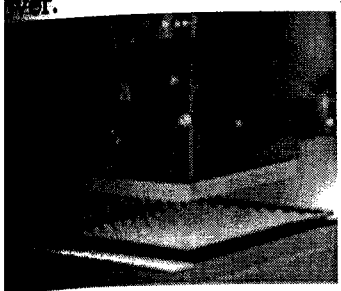
What the **QP**$^{\text{Display}}$ sees. What the eye sees

reen = selected, other colours = discarded, each colour representing the reason why e colony has been discarded.

he Picking Process

ypical Performance

wo 22cm picking trays, each containing 2000 pickable colonies, will be processed in ss than an hour, with an accuracy of better than 99.9% and no demonstrable carry-er.



rce: http://biocompare.com

system cleans and sterilises the picking head using a combination of washes and gen heat.

rce: http://biocompare.com

micro-well plates are then delivered to the picking area with lids automatically oved.

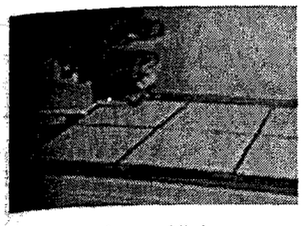
rce: http://biocompare.com

nies are picked using the proprietary 96 pin tool and inoculated into the micro-plates.
s are returned to the stack and the process repeated until all colonies are picked.

Data Tracking

The QP^{Display} utilises an on-board bar code reader to record, confirm and track plate and tray identities.
These are automatically combined with individual picked colony and re-arraying data in a simple XML format to enable information transfer with 100% confidence.

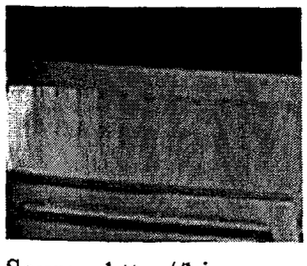
Panning



Source: http://biocompare.com

Delivering both capacity and convenience, membrane gridding is a viable alternative to panning using micro-well plates.
Efficiency of phage screening can be improved using arrays of bacteria spotted onto nylon membranes. These can then be interrogated with the antigen and detected using a second antibody against the antigen and a colormetric of flurometric assay to identify the reacting phage.
Following a simple tool change, six replicanylon membranes can be printed, each arrayed with up to 28,000 duplicate fragments (equivalent to 150 x 384 well micro-titre plates) in an unattended manner.
Results from the subsequent off-line analysis can then be correlated to plate identity through the data tracking module or log files.

Cherry Picking



Source: http://biocompare.com

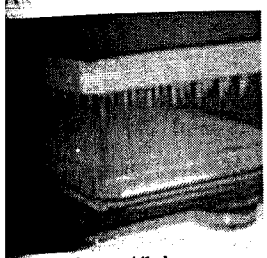
Locate and re-array positives with 100% confidence.
The re-arraying kit enables the automatic location of positive wells within a library of plates, the subsequent 'cherry picking' into new plates and the tracking of all relevant plate information.
Inherent in the phage display cycle is the need to identify and select a subset of the main library. For instance, positives may need to be selected for further rounds of panning or for further validation work such as sequencing. Selection of the subset of clones with unique and interesting sequence data is then simplified using the data tracking module.

Display enables highly efficient 'cherry-picking' using a 96 pin tool configured with ividually addressable pins, an easy-to-use graphical interface, and tracking and ntification of all plates throughout the process.

acked data

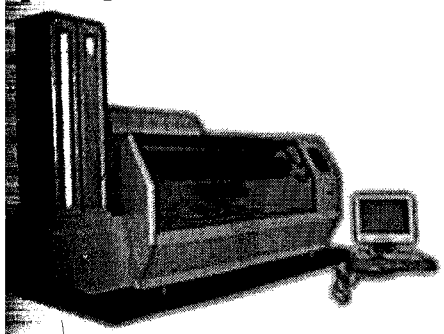
urce plate identity
stination plate identity
ividual well information

quencing



rce: http://biocompare.com
sequencing to determine uniqueness.

tomated delivery and lid management for up to 50 transformant plates and 20 p well plates



urce: http://biocompare.com
y-to-use methods and data tracking via Qsoft
0mm including
work tabl

Topic 12. mRNA Display Technique

Definition of mRNA and mRNA Display

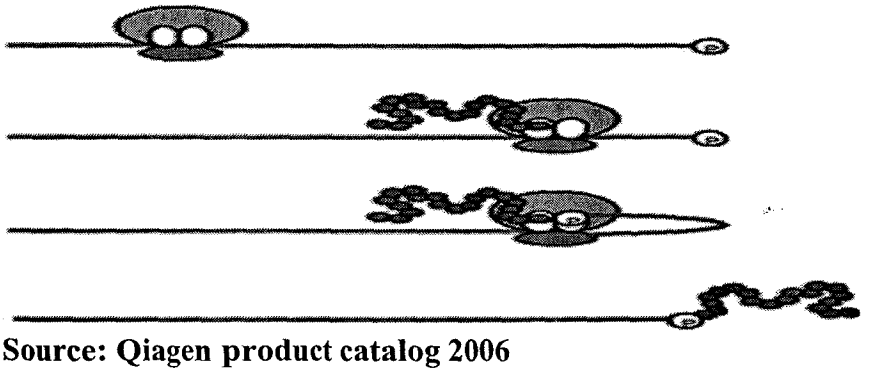
mRNA

mRNA is the template for protein synthesis ; the form of mRNA that carries information from DNA in the nucleus to the ribosome sites of protein synthesis in the cell. mRNA is a long linear polymer of nucleotides found in the nucleus but mainly in the cytoplasm of a cell where it is associated with microsomes ; it transmits genetic information from DNA to the cytoplasm and controls certain chemical processes in the cell ;" ribonucleic acid is the genetic material of some viruses"

The basic function of the nucleotide sequence of mRNA is to determine the amino acid sequence in proteins

mRNA display

mRNA Display is a technique used to perform in vitro protein evolution. A synthetic mRNA with puromycin (is an antibiotic that is a potent inhibitor of translation. at its 3' end forms an mRNA-peptide fusion that is purified by affinity chromatography



Source: Qiagen product catalog 2006

mRNA display process

Step 1. Transcription : A library of dsDNA sequences is transcribed to generate mRNA.

Step 2. Ligation : The mRNA is ligated to a puromycin oligonucleotide and used to program an *in vitro* translation reaction.

Step 3. Translation : cDNA synthesis is performed

Step 4. Purification : the cDNA/mRNA–protein fusion by reverse transcriptase is immobilized using the target of interest.

Step 5. Regenerate : PCR is used to regenerate the full-length DNA construct.

mRNA display Technologies

1. Discovery and Optimization of Peptide Ligands Using mRNA Display Roberts,R.W.,(2006)

mRNA display provides a means to generate and sieve peptide and protein libraries with more than one trillion members, entirely in vitro. There are isolating high affinity RNA-binding peptides. This work has resulted in molecules that bind a common RNA structural element, the GNRA tetraloop. Our recent efforts to extend and generalize this approach to more complicated RNA targets will be presented.

2. Use of DNA Microarrays to Deconvolute the Results of mRNA-Display Genomic Selections

Short III,G.F.,(2006)

Since its inception, mRNA-display has been exploited as a selection platform allowing the derivation of functional peptides and proteins from random sequence libraries. Recently, developing mRNA-display genomic libraries as tools for deciphering protein-protein interactions mediated by PDZ domains. In an effort to facilitate high-throughput proteomic screening, generating a 70-mer DNA microarray containing probes for all known and predicted human PDZ domains based on the SMART database (EMBL). The DNA microarray has been used successfully for the deconvolution of a model PDZ selection using the C-terminus of the NMDA receptor as bait. Instead of relying upon iterative cycles of selection/amplification, as is traditionally done in mRNA-display, we have been able to read-out the selection directly by DNA microarray analysis, thus obviating the need for DNA sequencing. These results suggest that this approach may be a viable tool for the rapid elucidation of candidate proteins from various genomic libraries and as a useful method for proteomic discovery itself.

3. Advancements in Selecting Ultrahigh-Affinity Binders Routinely Using mRNA Display Technology

Hale,S.P.,(2006)

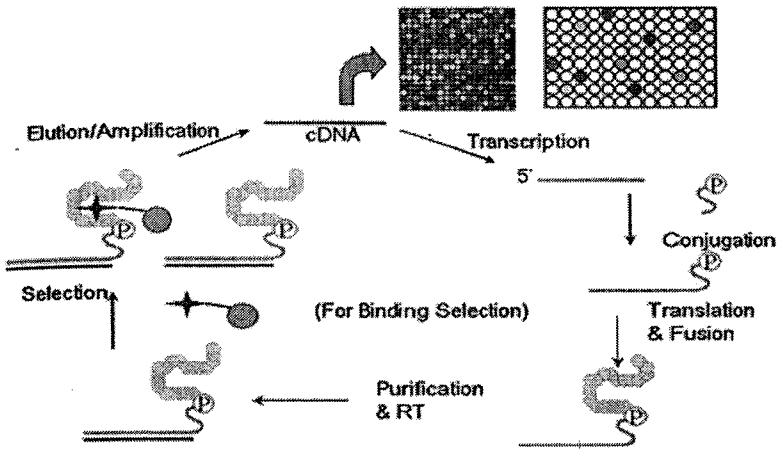
mRNA display allows for the generation of libraries of immense diversity. The potential for molecules of highly specialized function being present in such libraries is great. A thorough understanding of the parameters describing both the desired functionality and the selection system itself allows for the conditional enrichment of highly specialized molecules.

4. Ribosomal Synthesis of Unnatural Peptides

Szostak,J.W.,(2006)

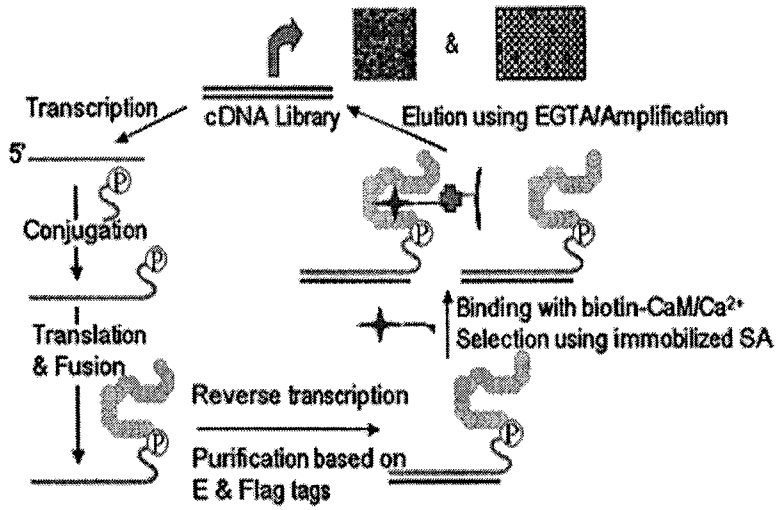
Many pharmacologically interesting molecules are synthesized by large multi-enzyme complexes such as the non-ribosomal peptide synthases. There are modifying the translational apparatus to allow similar highly modified cyclic peptides to be made on ribosomes. This would allow mRNA-display to be used to select for rare molecules with high affinity and specificity for particular targets, from libraries of over 1013 initial structures.

Genomic Selections Using mRNA-Display



Source: http://www.pegsummit.com/06-PGD.asp

Proteomic Selection of CaM-binding Proteins Using mRNA-Protein Fusions



Source: http://www.pegsummit.com/06-PGD.asp

Oligotex mRNA Mini Kit (6)

For miniprep purification of poly A+ mRNA from total RNA and cleanup of in vitro transcripts

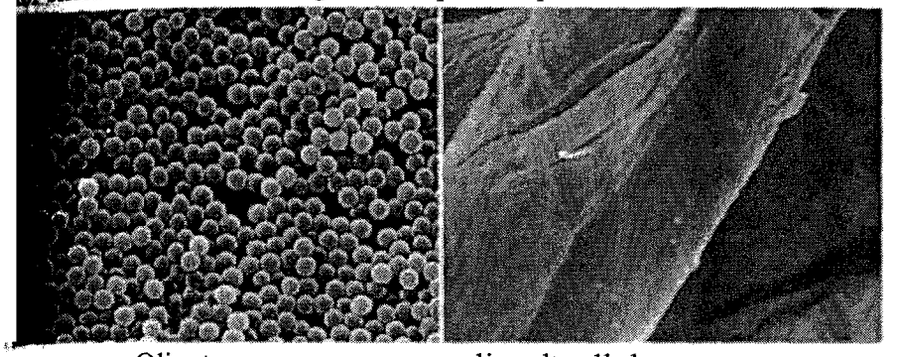
- High recovery of pure mRNA in as little as 30 minutes
- No oligo-dT cellulose or ethanol precipitation
- Flexibility for use with widely varying amounts of starting RNA

Oligotex mRNA Kits contain spin columns and all necessary reagents and buffers for isolation of pure poly A+ mRNA from total RNA preparations or cleanup of in vitro transcripts. Poly A+ mRNA purification takes less than 30 minutes with >90% recovery. Poly A+ mRNA is immediately ready for all standard applications, as well as for sensitive applications such as expression array, differential display, cDNA

ary construction, microinjection, and SAGE technology. Oligotex Suspension is available separately.

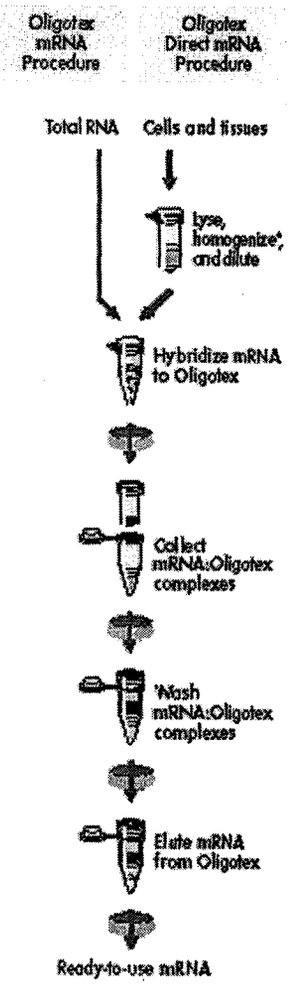
nciple

igotex resin consists of polystyrene–latex particles of uniform size (1.1 μm meter) and a perfect spherical shape with dC10T30 oligonucleotides covalently ked to the surface. The size, composition, and surface structure of Oligotex ticles have been optimized for uniform dispersion and minimal centrifugation time. he particles form a stable suspension that provides a large surface area for rapid and icient binding of polyadenylic acids. The high capacity and accuracy of ybridization provides superior purification of poly A+ mRNA through fast and icient hybridization. mRNA recoveries are consistently greater than 90%. The igotex purification procedure minimizes loss of poly A+ RNA, eliminates the risk degradation by RNases, and requires minimal hands-on time. This makes it ideal r simultaneous handling of multiple samples.



Oligotex oligo-dt cellulose

Source: http://www.giagen.com/product/RNA



Procedure

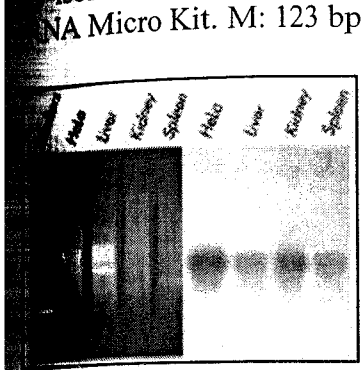
Optimized Oligotex purification procedures are convenient and Total RNA preparations are incubated with Oligotex resin, and Oligotex:mRNA complexes are collected by a brief centrifugation step. Spin columns, supplied in the Oligotex Kits, provide the most convenient handling, and optional batch protocols are also provided for certain applications. After washing, the mRNA is eluted in a small volume of Tris buffer or water. Alcohol precipitation is not required.

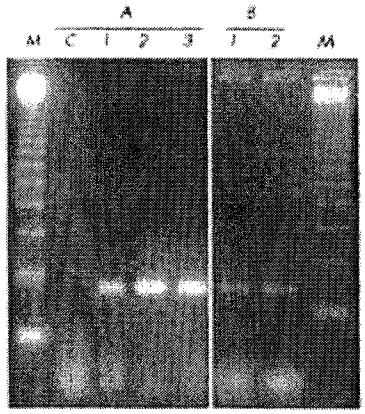
Oligotex resin can replace soluble oligo-dT primers in cDNA synthesis reactions to yield cDNAs immobilized on Oligotex particles. Oligotex:cDNA complexes represent a reusable cDNA pool that can be directly used in subtractive hybridization to enrich low-abundance cDNAs of differentially expressed genes (Hara, et.al.,1991)(Hara, et.al.,1993) (Kuribayashi-Ohta, et.al.,1993). The subtracted cDNA pool can then be used for library screening or construction.

Oligotex has also been used successfully for cleanup of poly-A tailed oligonucleotides (Adum,et.al.,1996)

CR of Beta-Actin mRNA

rse transcription was oligo-dT primed and PCR carried out using actin-specific ers. A: Poly A+ mRNA was isolated from 0.25, 0.5, or 1 µg mouse-liver total A (lanes 1–3), using the Oligotex mRNA Mini Kit. C: negative control. B: mRNA isolated from 103 or 104 HeLa cells (lanes 1 & 2), using the Oligotex Direct NA Micro Kit. M: 123 bp ladder.





ource: http://www.giagen.com/product/RNA

olation of mRNA from Total RNA Formaldehyde agarose gel and northern blot (GAPDH-probed) of poly A+ mRNA isolated from total RNA from various samples using Oligotex mRNA Kits.

Advantage of mRNA display (7)

1. mRNA display solves the "large display object problem" in an elegant fashion, while maintaining the major advantage of ribosome display, which is large diversity not constrained by bacterial transformation.
2. Large diversity for protein selection techniques (>10^12)
3. Covalent linkage between Protein and the RNA that encodes it allows for stability in low salt conditions
4. Completely in-vitro -- no phase of selection needs to go through an organism of any type.
5. m-RNA-display can be easily applied to genomic applications

Disadvantage of mRNA display (Differential display) (8)

1. Not very quantitative
2. Prone to false positive
3. Sensitivity can be an issue
4. Not easily automated or scaled-up

Reference

Hara, E., Kato, T., Nakada, S., Sekiya, S, and Oda, K. (1991) Subtractive cDNA cloning using oligo(dT)30-latex and PCR: isolation of cDNA clones specific to undifferentiated human embryonal carcinoma cells. Nucleic Acids Res. 19, 7097.

Hara, E. et al. (1993) DNA-DNA subtractive cDNA cloning using oligo(dT)30-latex and PCR: identification of cellular genes which are overexpressed in senescent human diploid fibroblasts. Anal. Biochem. 214, 58.

Kuribayashi-Ohta, K., Tamatsukuri, S., Hikata, M., Miyamoto, C., and Furuichi, Y. (1993) Application of oligo(dT)30-latex for rapid purification of poly(A)+ mRNA and for hybrid subtraction with the in situ reverse transcribed cDNA. Biochim. Biophys. Acta 1156, 204.

Adam, G.I.R., Cui, H., Miller, S.J., Flam, F., and Ohlsson, R. (1996) Allele-specific in situ hybridization (ASISH) analysis: a novel technique which resolves differential allelic usage of H19 within the same cell lineage during human placental development. Development 122, 839.

Ackermann, H.W., (2003) Bacteriophage. The desk encyclopedia of microbiology 135-146

Hamilton,P.T.,(2002) Peptides from Phage Display Libraries as Tools for Drug Discovery. (online) http://www.healthtech.com/2002/pgd/index.htm

Scott, J.K.,(2002) Peptide Phage Display for Vaccine Development on Anti-HIV Antibody.(online) http://www.healthtech.com/2002/pgd/index.htm

Somers, V.(2002) Profiling the Cancer Immunome Using Phage cDNA Display (online) http://www.healthtech.com/2002/pgd/index.htm

Sharon, J.(2002) Polyclonal Antibody Libraries for Cancer and Infectious Diseases (online) http://www.healthtech.com/2002/pgd/index.htm

Neri, D. (2006) Vascular Antigen Discovery by in vivo and ex vivo Biotinylation and Antibody Phage Technology. (online) http://www.pegsummit.com/06-PGD.asp

Dennis, M.(2006) CDR Repair , a New Approach to Antibody Humanization (online) http://www.pegsummit.com/06-PGD.asp

Palzkill, T.G. (2006) Engineering Speificity Changes in Protein-Protein Interactions using Phage Display. (online) http://www.pegsummit.com/06-PGD.asp

Simon, M.(2006) Re-Engineering a Protein-DNA Interface using Unnatural Nucleotides (online) http://www.pegsummit.com/06-PGD.asp

Gordon, N.C.(2006)Synthetic Antibodies Specific for Different Lysine-Liked Forms of Polyubiquitin Developed vai Phage Display. (online) http://www.pegsummit.com/06-PGD.asp

Roberts, R.W. (2006) Discovery and Optimization of Peptide Ligands Using mRNA Display.(online) http://www.pegsummit.com/06-PGD.asp

Short, G.F.,(2006) Use of DNA Microarrays to Deconvolute the Results of mRNA-Display Genomic Selections.(online) http://www.pegsummit.com/06-PGD.asp

Hale, S.P., (2006) Advancements in Selecting Ultrahigh-Affinity Binders Routinely Using mRNA Display Technology.(online) http://www.pegsummit.com/06-PGD.asp

Szostak, J.W.,(2006) *Ribosomal Synthesis of Unnatural Peptides.(online)* http://www.pegsummit.com/06-PGD.asp

1. http://www.dyax.com/discovery/phagedisplay.html phage display discovery tools.

2. http://www.cf.ac.uk/phrmy/PCB/PagePhageDisplay.htm

3. http://www.healthtech.com/2002/pgd/index.htm

4. http://www.amgen.com/pdfs/abgenix/2001-02-05.pdf

5. http://www.neb.com/nebecomm/tech_reference/drug_discovery/phdFaq.asp

6.http://www.giagen.com/product/RNAstabilizationPurification/oligotexsystem/oligitex mini.aspx

7.http://genetics.mgh.harvard.edu/szostakweb/research/protein/framecontent/proteincontent.html

8. http://www.genetix.com

Topic 13. RNA interference (RNAi) Technique

History of RNAi

Before the discovery of RNAi, co-suppression of gene expression was discovered in plants. In 1990, Napoli C, *et al* tried to deepen the purple color of the petunias by injecting the gene responsible into the petunias but were surprised at the result (1). Instead of a darker flower, the petunias were either variegated or completely white. This phenomenon was termed co-suppression, since both the expression of the existing gene (the initial purple color), and the introduced gene (to deepen the purple) were suppressed. Co-suppression has since been found in many other plant species (2, 3) and also in fungi (4). It is now known that double stranded RNA is responsible for this effect. Thereafter, antisense RNA was introduced to suppress protein synthesis and create a knockout organism to study the gene function. Antisense RNA methods have also been used for commercial food production such as Flavr Savr tomato. This tomato was developed by Calgene Inc. of Davis, California in 1988 (5) and was approved by the U.S. Food and Drug Administration (FDA) in 1994 (6). The tomato was the first whole food created by biotechnology that was evaluated by the FDA. One of the problems associated with tomato farming is that the fruit must be picked while still green in order to be shipped to market without being crushed. The enzyme that causes softening in tomatoes is polygalacturonase (PG). This enzyme breaks down pectin as the tomato ripens, leading to a softer fruit. Calgene suppressed the expression of the gene encoding PG by introducing a gene encoding the antisense strand of the mRNA. When the introduced gene was expressed, the antisense strand bound to the PG mRNA, suppressing the translation of the enzyme. The Flavr Savr tomatoes therefore had low PG levels and remained firmer when ripe. This meant the Flavr Savr tomatoes can ripen on the vine and then be shipped to market. Although the Flavr Savr tomatoes were approved for sale in the U.S., production problems and consumer wariness stopped the production of this fruit in 1997 (http://www.organicconsumers.org).

By an accident while studying of antisense using RNAs synthesized *in vitro*, one should use sense RNA of the same coding region as a control. Surprisingly, preparations of sense RNA often turn out to be as effective an inhibitor as antisense RNA. It seems that the preparations of sense RNA often are contaminated with hybrids: sense and antisense strands that form a double helix of double-stranded RNA (dsRNA). This dsRNA corresponding to a particular gene suppose to be a powerful suppressant of that gene. Guo and Kemphues (7) proved this evidence by injection of sense RNA as well as antisense RNA which resulting in a specific and reproducible phenotype. This phenomenon was later separated from classical antisense RNA and called in term of "RNA interference" or RNAi (also called post-transcriptional gene silencing or PTGS) by Craig M, *et al.* (8), however, leaving the mystery of RNAi mechanism behind.

One proven hypothesis that the preparations of sense and antisense RNA generated by *in vitro* transcription reactions using plasmid templates might be contaminated with small amounts of RNA of the opposite polarity by infidelity of the used viral RNA polymereases in the reaction was observed by Fire A, *et al.* (9). Triggering the sequence-specific degradation of targeted endogenous mRNAs in *Caenorhabditis elegans* by dsRNA and purified single-stranded RNA. ssRNAs were

less efficiency than those of dsRNAs in silencing specific gene espression. Later on, a flooding number of RNAi silencing papers in *C. elegans* gene has been reported. However, not only gene function has been elucidated in worms using RNAi silencing, but also in mammalian in recent years including in the developing of antiviral, anti-cancer and metabolic diseases therapeutics.

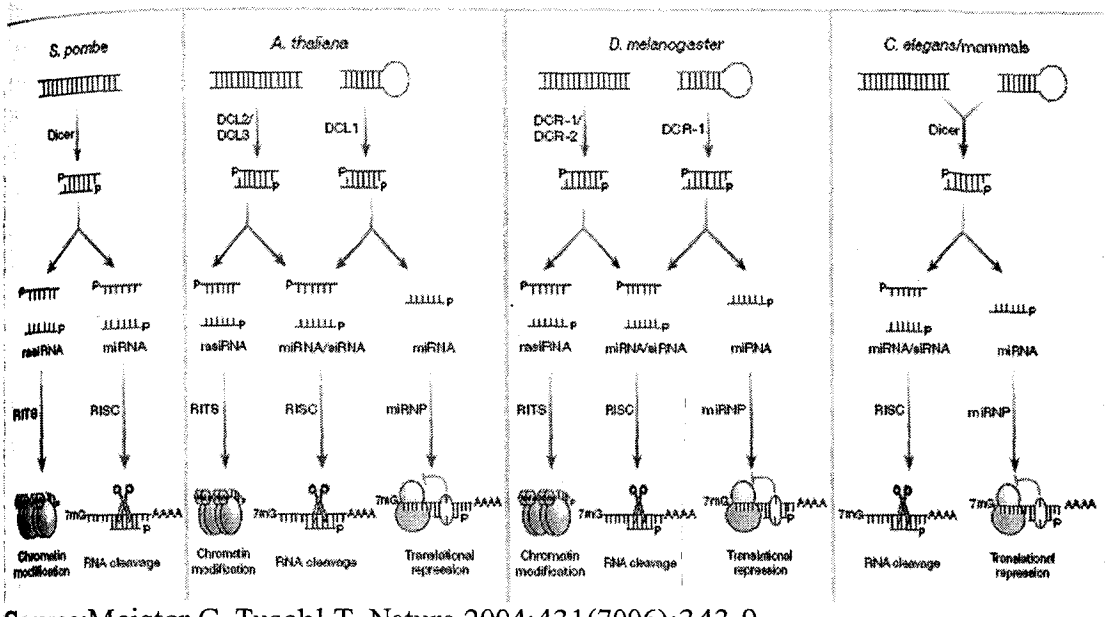
The RNAi pathway

There are three types of naturally occurring small RNA according to their origin or function: short interfering RNAs (siRNAs), repeat-associated short interfering RNAs (rasiRNAs) and microRNAs (miRNAs). RNA silencing associated function in at least four different mechanisms: 1. endonucleolytic cleavage of the cognate mRNAs, 2. translational repression, 3. transcriptional repression through the modification of DNA and/or histone, and 4. DNA elimination through the modification of histone (10). Silencing pathways of these RNAs share a common set of proteins that produce or amplify small RNAs and couple small RNAs to specific regulatory outcomes. Naturally, dsRNA can be produced by RNA-templated RNA polymerization (for example, from viruses) or by hybridization of overlapping transcripts (for example, from repetitive sequences such as transgene arrays or transposons). Such dsRNAs give rise to siRNAs or rasiRNAs, which generally guide mRNA degradation and/or chromatin modification. In addition, endogenous transcripts that contain complementary or near-complementary 20- to 50-base-pair inverted repeats fold back on themselves to form dsRNA hairpins. These dsRNAs are processed into miRNAs that mediate translational repression, although they may also guide mRNA degradation (11). miRNAs differ from siRNAs in their biogenesis, not in their functions. Moreover, miRNAs are most likely to target at 3' UTR of the target mRNA. Finally, artificial introduction of long dsRNAs or siRNAs has been adopted as a tool to inactivate gene expression, both in cultured cells and in living organisms (11).

The maturation of small RNAs is a stepwise process catalysed by dsRNA-specific RNase-III-type endonucleases, termed Drosha and Dicer, which contain catalytic RNase III and dsRNA-binding domains (dsRBDs). Drosha is specifically required for the processing of miRNA precursors, but not for the processing of long dsRNA. miRNAs are transcribed as long primary transcripts, which are first processed by Drosha in the nucleus. When Drosha excises the fold-back miRNA precursor, a 5' phosphate and a 2-nucleotide 3' overhang remain at the base of the stem. The miRNA precursor is then exported to the cytoplasm by means of the nuclear export receptor, exportin-5. Because exportin-5 lacks an obvious single-stranded or double-stranded RBD, it is not known whether the miRNA precursor binds directly to exportin-5 or to an RNA-binding adaptor protein. Once it is in the cytoplasm, the miRNA precursor is further processed by Dicer. Processing of dsRNAs by Dicer yields RNA duplexes of about 21 nucleotides in length, which — like Drosha-processing products — have 5' phosphates and 2-nucleotide 3' overhangs. Several organisms contain more than one Dicer gene, with each Dicer preferentially processing dsRNAs that come from a specific source (Fig. 1) (11).

Four Dicer-like (DCL) proteins (DCL1 to DCL4) have been identified in *Arabidopsis thaliana*, three of which are involved in processing dsRNAs that come from different sources. DCL1 processes miRNA precursors, requiring two more proteins to do so: HEN1 and the dsRBD protein HYL1. DCL2 is required for the production of siRNAs from plant viruses, and DCL3, which also cooperates with

HEN1, is involved in the production of rasiRNAs. In *C. elegans*, only one Dicer (DCR-1) has been identified. This cooperates with the dsRBD protein RDE-4 during RNAi, although RDE-4 is not required for miRNA function (11).



Source:Meister G, Tuschl T. Nature 2004;431(7006):343-9.

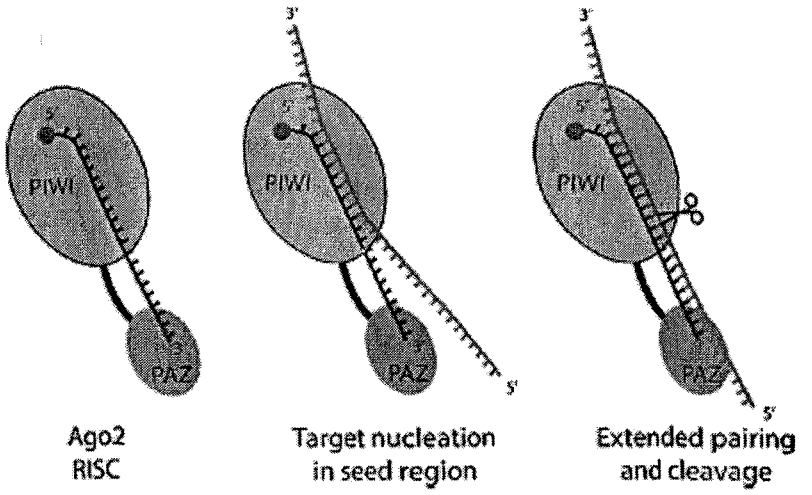
Figure 1 RNA silencing pathways in different organisms. Long dsRNA and miRNA precursors are processed to siRNA/miRNA duplexes by the RNase-III-like enzyme Dicer. The short dsRNAs are subsequently unwound and assembled into effector complexes: RISC, RITS (RNA-induced transcriptional silencing) or miRNP. RISC mediates mRNA-target degradation, miRNPs guide translational repression of target mRNAs, and the RITS complex guides the condensation of heterochromatin. In animals, siRNAs guide cleavage of complementary target RNAs, whereas miRNAs mediate translational repression of mRNA targets. rasiRNAs guide chromatin modification. S. pombe, C. elegans and mammals carry only one Dicer gene. In D. melanogaster and A. thaliana, specialized Dicer or DLC proteins preferentially process long dsRNA or miRNA precursors. 7mG, 7-methyl guanine; AAAA, poly-adenosine tail; Me, methyl group; P, 5' phosphate.

The siRNA- and miRNA-duplex-containing ribonucleoprotein particles (RNPs) are subsequently rearranged into the RNA induced silencing complex (RISC). Members of the argonaute protein family are core components of these RISC or RISC-like complexes. They are genetically required for RNA silencing in each organism where their function has been studied, but the exact role of this family has generally not been determined. Lots of proteins acting in the RNAi pathway share a common domain of activity, that is the DEA(H/D)-box helicase, but no specific biochemical function in RNAi has been ascribed to these helicases. The main catalytic activity of the RISC complex is to cut target RNA. In humans and mice, this is performed by an enzyme called Ago2 that belongs to the argonaute protein family (12). Actually in human Ago1–Ago4 all associate with miRNAs and siRNAs, but only Ago2-containing RNPs exhibit RISC activity (13). This protein family is characterized by two domains: the PAZ domain and the PIWI domain. Additionally,

RISC components include the RNA binding protein VIG, the Drosophila homolog of the Fragile X protein, dFXR, helicase proteins, and Tudor-SN. This last protein has five staphylococcal nuclease (SNase) domains and a Tudor domain. The presence of SNase domains made it an obvious candidate for Slicer. Several lines of evidence, however, are inconsistent with this. Many essential catalytic residues are absent in the SNase domains of Tudor-SN. While the protein still exhibits some nuclease activity, the chemistry of the cleavage reaction differs from that observed with Slicer. Specifically, products of the Slicer reaction have 5' phosphate and 3' OH moieties. The scissile bond has been mapped to the center of the siRNA, indicating an endonucleolytic reaction. SNase, however, is an exonuclease that produces 5' OH and 3' phosphate products. While Tudor-SN may have a role in degrading Slicer products, it is not Slicer itself (14).

Mechanistic studies on RISC have recently reached an apex with the crystal structure of AfPiwi complexed with a dsRNA. The AfPiwi protein is not a perfect model, since it lacks the PAZ domain, and its cellular role is a mystery. Nevertheless, this structure provided molecular details to a number of experimental observations. For example, microRNA/target pairs have demonstrated the importance of nucleotides 2–8 in the microRNA, termed the 'seed' region, for target recognition. The first nucleotide does not contribute to target recognition, and nucleotides 9 to the 3' terminus have reduced importance. The AfPiwi structure shows that the first nucleotide of the siRNA does not pair with the target, but is sequestered in a binding pocket. Not only is base pairing unnecessary, but also a strong pairing may distort RNA binding and reduce Slicer activity. Interestingly, a strong base pair at the 5' terminus of the siRNA should not occur in any case. Rules that govern strand incorporation into RISC are based on low pairing energy at the 5' end of the incorporated (guide) strand, compared to the discarded (passenger) strand. This means that an effective siRNA (or microRNA) will begin with an A or U, thus will not prevent proper binding to Argonaute. Of course, another way to achieve specific strand loading is to have a mismatched G or C at the 50 end, which would strongly base pair to a matched target. This did not reduce Slicer catalytic rate, however (15).

For slicer catalysis action, the siRNA guide strand is bound at the 5' end by the PiWi domain and at the 3' OH terminal end, or duplexes with a 3' overhang by the PAZ domain. The 5' phosphate is coordinated by conserved basic residues. mRNA targets are initially bound by the seed region of the siRNA and pairing is extended to the 3' end. The catalytic engine is the RNaseH fold in the Piwi domain. Typical RNase H endonucleases cleave the RNA strand of a RNA/DNA duplex, in a cation dependent manner, generating 5' phosphate 3' OH products. When presented with a long RNA substrate duplexed with a short DNA oligonucleotide, however, the enzyme cleaves the RNA in the center of the oligonucleotide. This is essentially a Slicer activity, though the siRNA guide takes the place of the DNA oligonucleotide (15). The model of slicing as shown in figure 2.



Source: Bartel DP. Cell 2004;116(2):281-97.

Figure 2 Model for slicer catalysis.

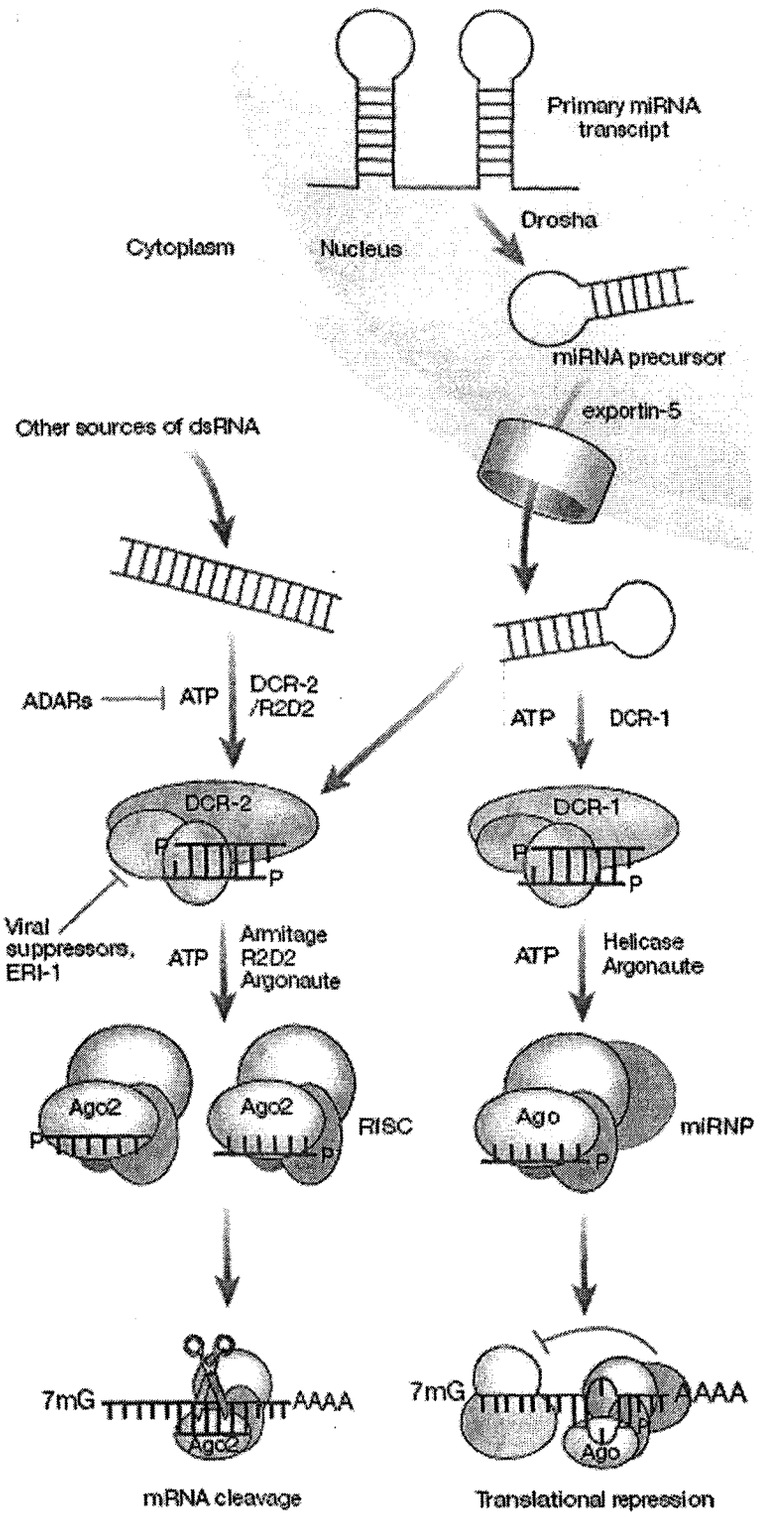
This model also explains binding of mRNA targets to miRNA RISC. The 5' seed region of the miRNA is essential for binding affinity, as has been observed for bonafide, and artificial mRNA targets (16). Since Slicer activity is not required, the 3' region of the miRNA is relatively unimportant. However, it is not clear about the mechanism of translational suppression. Target degradation does occur, but this does not appear to be the primary cause of gene silencing. While the miRNA and RNAi pathway share the same core machinery, some specialization may exist. For example, in Drosophila Ago1 preferentially binds miRNAs and Ago2 siRNAs (15). Animal miRNAs typically, but not always, mediate translational repression rather than cleavage. In contrast, most plant miRNAs direct target RNA cleavage. In all of the well-studied examples of translational repression by animal miRNAs or by siRNAs, translational repression is a response to the binding of multiple small RNA-programmed RISCs to sites in the 3' UTR of the mRNA target. In fact, the presence of multiple candidate sites in 3' UTR sequences is a useful predictor of an mRNA being regulated by miRNA. Complementarity between a small RNA and its RNA target helps determine if the small RNA directs target cleavage or represses mRNA translation, because an A-form helix must be formed with the target RNA at the center of an siRNA guide strand for cleavage to occur (14). Overview siRNA/miRNA silencing is shown in figure 3.

A third silencing is mediated by rasiRNAs (24-29 nucleotides in length) which found highly enriched in testis of *Drosophila*. Unlike siRNA, rasiRNAs do not arise from double-stranded precursor but rather from antisense strand. Moreover, rasiRNAs lack the 2',3' hydroxy termini as a main characterer of animal siRNA and miRNA and not requite either Dicer-1 or Dicer-2 which makes miRNA and siRNA, respecitvely. They form a complex similar to the RISC complex except the main protein involved is Piwi, a member of the Argonaute family of proteins. It seems like rasiRNAs function to inhibit bits of selfish DNA that copy themselves and multiply within the genome. These "selfish genes" are called retrotransposons (17). RasiRNAs match repetitive sequence elements in sense and antisense orientation and function in the establishment of heterochromatin in repetitive elements leading to transcriptional

silencing. RasiRNAs not only found in *Drosophila* but also in plants, *Trypanosoma brucei*, and fission yeast. Moreover, in fission yeast, rasiRNAs can suppress the transcription of repetitive transposable elements at the level of transcription (10, 18). The biochemical basis of rasiRNA-induced heterochromatin formation is most comprehensively studied in fission yeast. The heterochromatin DNA in fission yeast consists of the simple transposon-derived tandem array that surround the central core centromeric region of each chromosome. In mutants deficient in Ago1, Dicer, and RdRP genes, the centromeric outer-transposon repeats are de-repressed (10).

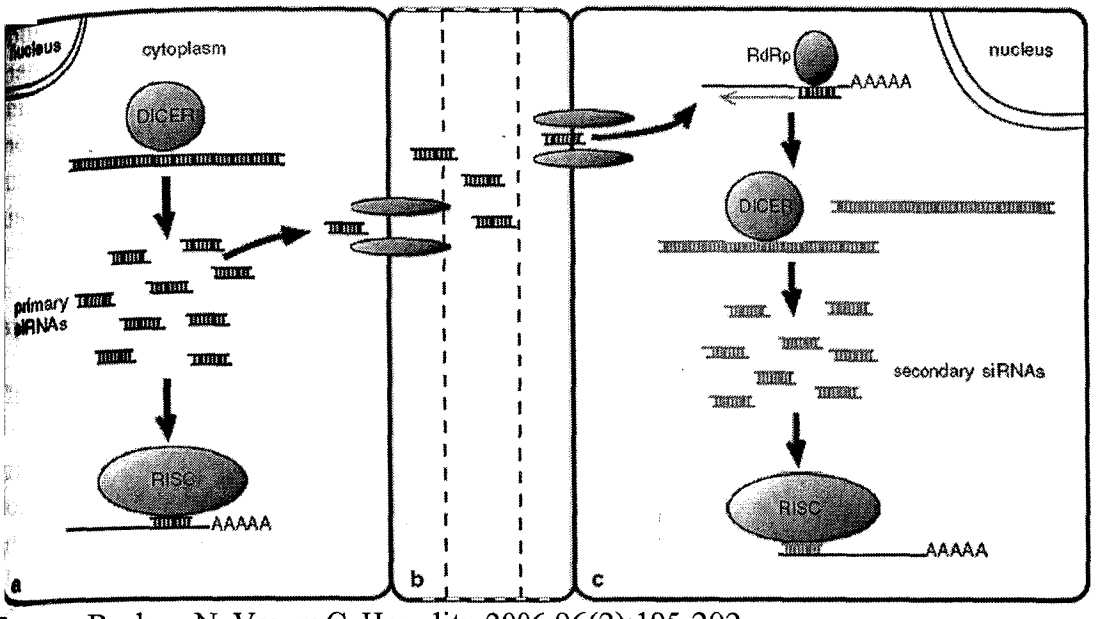
Recently, a new small noncoding RNA, Piwi-interacting RNAs (piRNAs) has been identified (19). PiRNAs are 25-31 nucleotide in length and produced from endogenous mammalian cells. They form complex with Piwi designated piRNA complex (piRC). PiRNAs are believed not generated from double strand DNA as same as miRNAs and siRNAs. So far the mechanism is not clear, however, piRNA is believed to be linked to transcriptional silencing.

Gene silencing is thought to be a naturally fist-line immune defense against viruses and transposable elements. siRNA has to be amplified in order to destroy large amounts of targeted RNA very efficiently (Figure 4). Initial studies performed in plants and nematodes have reported that RNA-dependent RNA polymerases or RdRp could be involved in the generation of secondary siRNAs homologous to the full-length of the target RNA (20, 21). Two models of action have been proposed. First, RdRp is able to generate newly synthesized siRNAs. Second, RdRp uses siRNA as a primer for extension, and synthesizes new long dsRNAs. These newly synthesized RNAs are able to reinitiate RNAi and thus amplify the system. A difference in the length of siRNAs has been observed: primary siRNAs are 21–22 nucleotides long, whereas siRNAs resulting from an amplification step are 24–26 long. These discoveries highlighted an unexpected complexity of small RNA molecules involved in the RNAi pathway (12). Transitivity has been found only in plants and nematodes which might be some organisms lack of RdRp and may have been rarely detected due to its high specificity for certain target genes, or for any steps in the pathway.



Source:Meister G, Tuschl T. Nature 2004;431(7006):343-9.

Figure 3 Model of small-RNA-guided post-transcriptional regulation of gene expression.



Source: Buchon N, Vaury C. Heredity 2006;96(2):195-202.

Figure 4 Mechanisms for transitivity and systemy. In cell (a), double-stranded RNAs are recognized by the RNAi pathway and cleaved by a Dicer-like protein to form siRNAs. These siRNAs seek out and destroy homologous target RNAs in the cell where they have been generated. This RNA silencing can spread over 2–5 cells (indicated as (b) cells) by diffusing with the help of a transmembrane protein called SID. This phenomenon is called systemy. As illustrated (but still controversial), the spreading molecules may be siRNAs themselves. A mobile signal allows the RNAi-silencing pathway to be activated in neighbouring cells (indicated as (c)) and to target the same mRNA. In some organisms, an enzyme called RdRp generates new siRNAs able to reinitiate RNAi and amplify the system (see cell (c)). This phenomenon is called transitivity.

RNAi in plant models works as a real antiviral defence system targeting viruses and endogenous repeated sequences. This was confirmed when a subset of mutations selected as defective for RNAi were shown to mobilize families of transposable elements. In *D. melanogaster*, mutations in spindle E, a gene involved in RNAi, led to a derepression of retrotransposons in the germline, and the argonaute protein piwi is required for this repression in the male germline. Moreover, it has recently been shown that gypsy and ZAM, two endogenous retroviruses of *D. melanogaster*, are regulated by a piwi dependent mechanism in somatic. In *Trypanosoma brucei*, Ago1 deficiency leads to an increase in retroposon transcript abundance. In fungi, such as *Neurospora crassa*, transposons are known to be repressed by repeat induced point mutation (RIP). RNAi in these organisms is able to target transposon RNAs and act in concert with RIP to silence them (12). The existence of such an RNA-based defence mechanism against invading elements could be particularly important for organisms such as plants and invertebrate animals which lack protein-based adaptive immunity. Nevertheless, evidence has recently emerged that mammals have also developed post-transcriptional silencing to silence or destroy foreign genetic material detected in a cell. In preimplantation mouse embryos,

bition of the RNAi pathway resulted in a 50% increase in both murine ogenous retrovirus-L (MuERV-L) and intracisternal A particle (IAP) transcript ndance (22). In human cells, RNA-silencing limits the replication of at least one mmalian virus, PFV. Furthermore, natural HIV-1 infection also provokes nucleic-based immunity (23). Therefore, if this ancestral genomic defence is a general chanism targeting nucleic invaders from plant to mammalian genomes, it can be pected that numerous miRNA encoded by all the organisms remain to be covered (12).

pplication of RNAi

enome wide RNAi screens

After the discovery of RNAi in worm, many researchers began to study gene nctions in worm, flies and especially mammalian cells. RNAi is used to study of ene discovery and functional annotation in various processes, including early mbryonic development, lethality, sterility, genome instability, and longevity in *Drosophila* (24) and genes expressing small interference in mammalian (25). Some of uccessfully mammalian gene knockdowns by siRNA are shown in Table 1. The fficiency of transfection depends not only on the cell type, but also on the passage number and the CONFLUENCY of the cells. Most groups use Oligofectamine™ as a cationic lipid carrier for nucleic acids because of its low toxicity and its ability to transfect at low cell confluency. Depending on the cell type used, non-cationic lipid-based carriers (such as Transit-TKOTM) might be preferable. Many other transfection reagents are available, and as different reagents can be cell specific, it is always prudent to test which one works best for a given cell line. Several groups have used multiple lipid transfections for siRNA-mediated knockdowns. siRNA can also be electroporated into cells that are difficult to transfect. Although this method gives high siRNA transfection efficiency (>95%), 50% or more cells can die during the electroporation. If the cells are adherent, dead cells can simply be washed off the plate on the following day (26). Both liposome-based and electroporation induce transient silencing only. Study in mammalian, such as mice, function of genes can be done by produce transgenic. Another approach using viral gene transfer could induce silencing locally or systemic.

ble 1 Use of small interfering RNAs in mammalian cells

ted	Cell type	Location on mRNA*	Effective/ tested siRNAs	Time (hours)	%KD	Verification‡	No. of tfxs	Tfx method§	Phenotype	References
P135	CHO	8–26	1/1	48	85–90	W, IF	2	Lfm	Cell growth, microtubule organization	100
110	HeLa	397–418, 696–917	2/2	48	97–99	W, IF	3	Ofm	Centrosome amplification	101
iple	HeLa, 3T3, rat FB	See reference	21/22	40–48	See reference	W, IF	1	Ofm	Apoptosis, arrest, loss of stress fibres, aberrant spindles, blebbing; 13 out of 21 genes were essential for growth	60
F1	HeLa, Caco2	NR	2/2	72	97–99	W	1	Ofm	Exocytosis defect	102
	HeLa	677–697, 236–256	2/2	48	97–99	W	1	Ofm	ND	103
	HeLa	135–155	1/1	48–72	NR	W, IF	1	Ofm	Checkpoint-arrest defect	104
	HL-5, 293	NR	1/1	24–28	NR	NR	1	Ofm	Inhibition of pro-ANP processing	105
	HeLa	65–87, 461–482	2/2	96	NR	NR	3	Ofm	Inhibition of GAG synthesis	106
	T cells, HeLa	See reference	1/5	36	85–90	N, F	1	E, Ofm	Inhibition of CD4, CD8; primary T cells	53
	HeLa, Cos-1, 293, HaCaT	See reference	6/16	96–120	90	N	1	Lpf	Inhibition of TF activity	53
	293T	413–433	1/1	72	NR	N, W	1	Ofm	Arrest of HIV budding	107
	HeLa	89–109	1/1	48–72	NR	IF, W	1	Ofm	Aberrant chromosome partioning	108
	HeLa	NR	1/1	48	NR	W, IF	2	Ofm	Defective DNA-damage response	109
	HeLa, 293, MEF	NR	1/1	24	>85	N, F	1	Lfm	GFP silencing	54
	HeLa, cos, 293, 3T3	Distributed	6/7	40–48	>85	W, IF	1–3	Ofm, Lfm	Luciferase, lamin silencing	52

Source: McManus MT, Sharp PA. Nat Rev Genet 2002;3(10):737-47.

*Location of the siRNA target site along the mRNA is expressed as the number of nucleotides downstream from the start codon. ‡Knockdown (KD) verification was carried out by western blot (W), northern blot (N), immunofluorescence (IF) or flow cytometry (F). Transfection (Tfx) method§ included Oligofectamine (Ofm), Lipofectamine (Lfm), Lipofectin (Lpf) and electroporation (E). ANP, atrial natriuretic peptide; Cdc14A, cell division cycle 14 homologue A; CEP135, centrosomal protein of 135 kDa; CP110, centriolar protein of 110 kDa; EGFP, enhanced green fluorescent protein; GAG, glycosaminoglycan; GAGII, beta 3GalT6; HMps1 kinase, human mono polar spindle-kinase; HTF, human tissue factor; LMNA, lamin A/C; Luc, luciferase; ND, not determined; NR, not reported; PSKH1, protein serine kinase H1; Rad17, RAD17 homologue; RalA, v-ral simian leukaemia viral oncogene homologue A; RASSF1, Ras association (RalGDS/AF-6) domain family 1; ShcA, src homology 2 domain-containing transforming protein C1; Tsg101, tumour susceptibility gene 101.

Combined RNAi and ES cells for prospect future therapy

RNAi allows researchers to test the role of many genes in human embryonic stem (ES) cells without the need to generate mutant mice. Most of the genes so far shown to regulate the undifferentiated state of ES cells were chosen because of their expression patterns in the early embryo or from functional cDNA overexpression screens. With new technologies, it is possible to identify other candidate regulators of ES cells and study them with loss-offunction screens using RNAi. The availability of new ES cell lines expressing reporter genes under the control of promoters of ES cell-specific genes will allow researchers to monitor the undifferentiated state of ES cells. For example, human embryonic stem (hES) cell lines have been generated that express green fluorescent protein under the control of the Oct4 promoter (13, 27). Oct4 is an important marker of the undifferentiated state and play a major role in regulation of pluripotency in ES cells (27). Microarray analysis of ES cells suggests that some genes may have an important role in determining the stem cell state, because they are upregulated in these cells when compared to somatic cells (13).

Involvement of RNAi in ES cells may help to explore: understanding the basic biology of ES cells and cellular differentiation; modeling disease states *in vitro*; validating new drugs and assessing their toxicity; directing differentiation of cell types of interest from ES cells; controlling the cell cycle and immune repertoire of ES-derived cells to be transplanted and targeting infectious agents. Diseases that involve the loss or damage of a single or very few types of cells are the most attractive candidates for ES cell therapies. Parkinson's disease, lower motor neuron loss and spinal cord injuries and type I diabetes mellitus are all potentially treatable by these therapies . In addition to the conventional cell-replacement approaches aimed at repairing damaged tissues, the combination of ES cell and RNAi technologies may result in novel therapies for infectious diseases such as HIV, tuberculosis or malaria. One such strategy for combating the HIV virus has already been reported (28). It involves isolation of hematopoietic stem cells from an infected individual and treating them with a lentivirus that leads to expression of a shRNA targeted against either viral RNA or against the cellular receptor targeted by HIV. These stem cell populations are then expanded ex vivo and reintroduced into the patient. As hematopoietic stem cells give rise to the cells comprising the immune system, it is hoped that such a procedure will confer HIV resistance to the immune system (the main target of HIV). Alternatively, hematopoietic progenitors derived from ES cells carrying RNAi vectors that target HIV infection may be used. This approach would circumvent the need to extract hematopoietic stem cells from the patient, taking advantage of the fact that ES cells can be grown in very large numbers (13).

Although the potential of RNAi and stem cell-based gene therapies is extremely promising, there are issues of safety and efficiency that must be addressed before any potential therapy can be applied in humans. The combinatorial therapeutic use of RNAi and ES cells, while it may yield great benefit, also compounds the limitations and potential negative side effects which both tools may illicit. Currently, one of the most pertinent limitations involving the use of ES cells is the lack of knowledge regarding the details of ES cell developmental biology. The range of cell types that can currently be derived from ES cells is fairly limited. As a result, the disease targets of potential stem cell-based therapies are restricted to those affecting the small subset of cell types that can be derived. It is likely that in the near future the

number of cell types that can be derived from ES cells will greatly increase, but the goal of complete ES cell-based organ replacement may be far off. Most of the differentiation protocols to obtain a cell type of interest yield a heterogeneous population that contains other cell types as well. It will also be important to achieve cell purity before ES-derived cells can be transplanted into patients. Another major caveat of ES cell-based therapies is the possibility of tumor formation. If a transplant happens to contain contaminating undifferentiated ES cells, these could lead to the formation of teratocarcinomas. Downregulation through RNAi of specific genes involved in regulation of the cell cycle may be a way to avoid these tumors. Another limitation is the immune response following engraftment of an ES-derived transplant not immunologically matched to the patient. The use of immunosuppressive drugs can prolong the survival of allogenic ES cell progeny. This is not an ideal method due to the fact that the ability of the body to heal is compromised when the immune system is suppressed. Another strategy to enhance the compatibility of the graft is to decrease ES cell expression of cell surface proteins that activate host immune responses (i.e. major histocompatibility complex and costimulatory molecules) or to increase ES cell expression of immune-inhibitory antigens (i.e. Killer cell immunoglobulin-like receptors). These approaches in isolation will not likely permit long-term engraftment of ES cell-derived cells. Another way to prevent immune rejection is through the production of patient-specific ES cells by somatic cell nuclear transfer and thus has become a very popular topic of study. The use of RNAi itself also presents hurdles that must be overcome. The two main hurdles are effective RNA delivery and specificity of gene silencing. shRNA delivery for some gene therapies, such as those used in two clinical trials recently approved by the FDA (for agerelated macular degeneration), could be as simple as the injection of naked RNA. Many other techniques for RNAi delivery have been formulated including liposomal carriers, aerosolized vapors and viral vectors, but like any other potential therapeutic treatment these methods must be carefully evaluated for both efficiency as well as any possible off-target effects (13).

Using siRNAs as small molecule drugs

To use siRNAs or their small RNA precursors as drugs, they must be efficiently delivered into the cytoplasm of cells, the site of precursor RNA processing by the enzyme Dicer and uptake of the siRNA guide strand into RISC. Because Dicer may direct endogenously processed siRNAs and microRNAs to the RISC, small double-stranded RNAs that are between 25 and 30 nucleotides in length and require Dicer processingappear to be more efficient at inducing RNAi than smaller siRNAs designed for direct incorporation into RISC. Once within the RISC, the durability and efficiency of silencing depend on many factors, most of which are poorly understood, but may include sequence preferences for RISC binding along the siRNA, accessibility of the target site on the mRNA, and thermodynamic considerations. The most important factor that determines durability of silencing in cells appears to be the rate of cell division, which leads to dilution of the activated RISC as it is divided between daughter cells. In rapidly dividing tumor cells, used for most of the early RNAi *in vitro* experiments, silencing peaks 2–3 days after transfecting siRNAs and begins to wane around day 5 and disappears by day 7. However, in terminally differentiated non-dividing cells, such as macrophages, silencing lasts for as long as the cells can be cultured (several weeks). There is also some suggestion that the persistence of the activated RISC for any siRNA may depend on the presence of the

get mRNA. For therapeutic purposes, the rate of division of the target cell will be important determinant of the dosing interval (29).

Diseases that are intractable or poorly responsive to current therapy are in the high rank list of target, especially cancer, neurodegenerative disease, viral infection, and macular degeneration, and these are the disease models that have been most studied so far.

Designing siRNAs for therapy

Optimal siRNAs favor the guide strand entering Ago 2 to the exclusion of the passenger strand, thereby directing RISC to bind to the target mRNA, where potent cleavage of the target ensues. For effective RNAi, the 5' end of the guide strand requires sufficient complementarity to bind to the target and the orientation of the 3' end of the guide strand needs to favor catalysis of the target by Ago 2. Nonspecific cleavage of mRNA – unintended mRNA targets, or 'off-targets' – would be undesirable in RNAi therapeutics. mRNA targets in hereditary diseases or somatic gene mutations might differ from wild-type mRNA by a single nucleotide change. Single nucleotide polymorphisms offer another potential site for RNAi therapy. Therefore, siRNA designs will need to account for such small sequence changes. Furthermore, in general, mammalian cells resist uptake of double-stranded RNA, thereby requiring formulations or conjugations of siRNAs to facilitate their cellular uptake (30).

Maximizing guide strand selection is an initial, critical goal in siRNA design. Based on the previously discussed thermodynamic properties of asymmetric siRNAs, guide strands of siRNA can be predictably recruited into RISC. Inactive siRNAs, or less than optimally effective siRNAs, can achieve better RNAi by introducing nucleotide mismatches (thermodynamically unstable nucleotide pairings) (30). An example for huntingtin mRNA is provided (Figure 2). Strategically placed nucleotide mismatches can increase the RNAi by the guide strand, as noted by the increase in cleaving its target of huntingtin mRNA. siRNA can be functionally asymmetric, with only one of the two strands able to trigger RNAi (31). Improved RNAi activity would reduce therapeutic dose requirements, because more of the active guide strand than passenger strand enters RISC. Passenger strand effects have been implicated to have off-target effects; limiting passenger strand loading into RISC mitigates this concern. Use of asymmetric siRNA can be extended to design shRNAs for viral delivery of siRNA (30).

Selection of siRNA duplexes from the target mRNA sequence

The targeted region is selected from a given cDNA sequence beginning 50 to 100 nt downstream of the start codon. Initially, 5' or 3' UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. More recently, however, we have targeted 3'-UTRs and have not experienced any problems in knocking down the targeted genes. In order to design a siRNA duplex, we search for the 23-nt sequence motif AA(N19)TT (N, any nucleotide) and select hits with approx. 50% G/C-content (30% to 70% has also worked in our hands). If no suitable sequences are found, the search is extended using the motif NA(N21). The sequence of the sense siRNA corresponds to (N19)TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, we convert the 3' end of the sense siRNA to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense

antisense 3' overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3'-most nucleotide residue of the antisense siRNA, can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) should always be complementary to the targeted sequence. For simplifying chemical synthesis, we always use TT. Should an siRNA be expressed from pol III expression vectors, it is preferable that the first transcribed nucleotide is a purine. Upgraded selection rules suggest to bias the stability of the siRNA duplex in a manner that the 5' portion of the antisense siRNA is paired less stably to the sense siRNA than its 3' portion (http://www.rockefeller.edu/labheads/tuschl/sirna.html). Considerations for generating an effective and specific siRNA is summarized in table 2 (32). BLAST-search of genome sequence databases should be performed in order to ensure that only one gene is targeted.

Table 2. Criteria for specific and effective siRNA.

Criteria	Probable reason
Biophysical, thermodynamic and structural considerations	
Overall low to medium G+C content (30-50%)	Facilitates interaction with RISC and unwinding
Low internal stability at the 5' antisense strand	Promotes antisense-strand selection by RISC
High internal stability at the 5' sense strand	Blocks sense-strand selection by RISC
Absence of internal repeats or palindromes	Increases the concentration of functional, stable hairpins
A-form helix between siRNAs and target mRNA	Enhances RNA–RNA interactions and promotes cleavage
Base preferences at specific positions in the sense strand	
Presence of an A at position 3 and 19' of sense strand	Promotes antisense-strand selection by RISC
Absence of a G or C at position 19 of sense strand	Promotes antisense-strand selection by RISC
Presence of a U at position 10 of sense strand	Promotes RISC mediated cleavage of mRNA and dissociation of the RISC–siRNA complex
Absence of a G at position 13 of sense strand	Promotes efficient unwinding
Enhancing specificity of siRNA-mediated gene silencing	
Perform stringent homology searches	Minimizes potential nonspecific gene silencing
Avoid low-stringency sequence interactions between siRNA and 3'UTR	Minimizes potential nonspecific gene silencing

Source: Mittal V. Nat Rev Genet 2004;5(5):355-65.

Potential risks of RNAi

Possible effect of RNAi can change unintended profile of mRNA which induce nonspecific but sequence-dependent effects (33). Sequence complementary can also leading to off-target of unexpected even in natural target regulation by miRNAs (34). This can be explained by the design siRNAs, which has base mismatches to induce silencing activity, showed a complementarity to other mismatched seed sequences regulated by miRNAs. Additionally, miRNA targets are well expressed and contain longer 3' UTR regions, while anti-targets have shorter 3' UTRs, which limits miRNA target site density. miRNA anti-targets may also be expressed temporally or spatially to avoid coexpression of the transcript and miRNA (34). Mismatches of up

four nucleotides and G:U wobbles can lead to translational repression of candidate target mRNAs (30). Using different delivery systems can make cells response differently with either upregulation or downregulation which might be the effects of packaging reagents (35). Further studies *in vivo* should help to resolve the extent of off target siRNA effects. Lacking in studies of nonspecific mRNA silencing is the protein profile of the off target mRNAs. A small change in mRNA expression does not necessarily determine a biologically important change in its cognate protein. Minimal, but reproducible, changes in protein profiles might have clinical significance. Forecasting specific untoward effects based on subtle changes in mRNA content at this point is a leap of faith (30).

Topic 14. Gene Therapy Technology

Gene therapy

At the basic level gene therapy can be sescribed as the Intracellular delivery of genetic material to generate a therapeutic effect by correcting an existing abnormality or providing cells with a new function (36). Somatic cell gene therapy is a form of medical treatment that is being developed as a genetic approach to disease management. Three basic gene therapy strategies have been investigated ;

1) *In vivo* gene therapy is done by targeting the gene delivery system to the desired cell type in the patient using either physical means such as tissue injection (brain tumor) or biolistics (dermal DNA vaccination), or potentially in the future, using systemic infusion of cell-specific receptor-mediated DNA carriers (reconstructed liposomes or viruses). Importantly, neither of these gene therapy strategies involve reproductive germline cells and therefore the genetic alteration will NOT be transmitted to the next generation. In many countries, human germline gene therapy is considered unethical or even illegal.

2) In situ delivery is an administering the material directly to the desired tissue. Cystic fibrosis (CF) has been shown some success with this strategy. This disorder result from mutations in the cystic fibrosis transmembrane conductance regulator (CFTR). The CFTR gene has been delivered using lipid and adenoviral vectors to defined sites in the respiratory tract. Initial results showed a minority changed in CFRT-mediated ion transport at these sites.

3) Ex-vivo gene therapy involves removing tissue from the patient, transfecting (or virally infecting) the cells in culture, and then reimplanting the genetically altered cells to the patient. Two types of ex-vivo gene therapies under development are those directed at fibroblasts and hematopoietic stem cells.

Nonviral delivery systems

1. Physical nonviral delivery

Nonviral delivery can be divided by physical and chemical approaches. The fist one is physical methods compose of electroporation, micro injection and bioloistics.

The bombardment of cells/tissues with DNA coated pariticles or microprojectiles is a novel physical approach for transfection. Advancing the technique, an electrical discharge and then high-pressure helium replaced gunpowder ad trhe particle propellant for most particle-mediated devices. Biolistic is used as particle-mediated delivery system to deliver genes into skin *in vivo*. The transfer genes are expressed transiently. Advantages of this technology are the applicability to different cell types and tissues and the ability to deliver large DNA molecues. Additionally, multiple genes could be introduced in to target cells (37). However, it would be at risk of cell damage. Similarly, electroporation appears to be promising for local gene therapy: skin, muscle, brain and liver (38, 39). The intramuscular injection of plasmid DNA, however, has several advantages over viral vectors. First, plasmid DNA vectors are easier to construct and can be prepared as pharmaceutical-grade solutions without the risk of contamination with wild-type infectious particles. Second, previous infection by wild-type adenovirus or AAV may induce a

neutralizing antibody response that could preclude administration of the recombinant virus (39).

2. Chemical nonviral delivery

The vast majority of current chemical nonviral delivery systems are based around polycationic entities which cause compaction of negatively charged nucleic acids accompanied by the formation of nanometric complexes. Typically, cationic liposome and cationic polymer-based are two main categories of polycationic based method. Principally, the complex need to have a net positive charge to enter cells and that endocytosis is triggered y nonspecific interaction between cationic complexes and the anionic cell-surface proteoglycans of adherent cells. Once inside, a proportion of the bound nucleic acids should be able to dissociate and escape from early-endosomes into the cytoplasm either to perform a therapeutic function there as in the case of mRNA, or else traffic into the nucleus to perform a therapeutic function (36).

Viral delivery systems

Most viral systems have been used extensively in gene delivery studies. The most widely studied include retrovirus, adenovirus (types 2 and 5), adeno-associated virus and herpesvirus. Common to all viral vectors is the fact that their genomes have been modified, deleting areas which render them replication incompetent. This results in limiting the virus particle to only a single infectious cycle. Successful gene therapy requires that the:

- Genetic malfunction/nature of a disease is understood
- Therapeutic material can be delivered to the target cells in the affected tissue or organ
- Therapeutic material is active for the intended duration and delivers the intended benefit to the target cells
- Harmful side effects, if any, are manageable

Retrovirus-based vectors

Retroviruses are enveloped RNA-containing eukaryotic viruses which replicate through DN intermediates, facilitated by an RNA-directed DNA polymerase (reverse transcriptase). The reverse transcribed DNA genomes are subsequently integrated into the host chromosomes as proviruses, directed by the virus-encoded integrase protein, and this DNA is transcribed using the host machinery. Retroviral genomes contain three core genes termed gag, pol and env, which are flanked by long terminal repeat (LTR) sequences and a packaging signal which directs the assembly of the genome into the viral particles. The gag gene encodes proteins which form the viral core, while the pol gene encodes reverse transcriptase (RT), the viral integrase (Int) and a viral protease which acts on the gag gene products. The env gene encodes the glycosylated envelope proteins which determine the tropism of the virus. The LTR sequences contain the cis-acting sequences required to regulate viral genome replication, transcription and mediate integration into the host genome. The majority of retroviruses does not cause cancers but simply infect a cell and produce a persistent infection with continual production of virus by infected cells. A number of retroviruses are, however, tumorigenic, such as the well characterized Rous sarcoma virus (RSV), which causes a connective tissue cancer in chickens. The tumorigenicity of oncogenic retrovirues can be related to accessory oncogenes additional to the gag, pol and env genes, which convert the cell to a caner cell, capable of indefinite growth and eventually killing the organism containing it. The human immunodeficiency

ruses (HIV), the causative agents of AIDS, are also retroviruses of the lentivirus subgroup, which contain several additional accessory genes (36).

In terms of somatic gene transfer, replication-defective retroviruses offer a number of advantages over current delivery techniques. As well as providing significantly elevated targeting efficiencies and encoding no cytotoxic or immunogenic viral antigens, they result in efficient and stable integration of the transgene into the host chromosome, enabling a permanency that facilitates long-term expression. The targeting of retroviruses to specific cell populations is also being developed by engineering heterologous protein domains into the envelope glycoproteins modifying their tropism. However, a number of fundamental obstacles limit the effectiveness of retrovirus gene transfer, including their inability to infect somatic mitotic tissues such as muscle, due to their dependence on mitosis to gain entrance to the nucleus of infected cells to facilitate integration. Size limitations (9-10 kb) on the amount of DNA which can be stably incorporated into the retroviral cassettes to allow packaging into the retroviral particles, limit the application of larger genes or large regulatory elements. Additionally, the fairly low tiiters ($<10^9$ infectious units per ml) and the relative instability of the virions further hampers the applicability of retroviruses to gene therapy protocols. Sensitivity of murine retroviruses to inactivation by the human complement-mediated lysis pathway of the immune system has also proved problematic (36). Naldini and coworkers created replication defective versions of HIV that were capable of transducing nondividing cells and achieving stable expression through integration of the provirus into chromosomal DNA (40). This group of retroviruses that can transfect nondividing cells is called lentivures. Although several HIV packaging cell lines have been developed, compared to transient transfection, these typically yield lower vector titers, and particle production may be transient (lasting only 5–10 days) after induction of vesicular stomatitis virus (VSV) G protein expression (41). Two groups of research, Ikeda *et al.* and Ni *et al.*, however, could constructed the lentiviral systems that prolong vectors production with high titer (42, 43). In addition, besides the low copy number production, the fidelity of HIV was significantly lower when vector was produced by transient transfection compared to stable producer cell lines (44).

How are HIV lentiviral vectors made?

To obtain a lentiviral gene therapy vector, a reporter gene or therapeutic gene is cloned into a vector sequence that is flanked by LTRs and the Psi-sequence of HIV. The LTRs are necessary to integrate the therapeutic gene into the genome of the target cell, just as the LTRs in HIV integrate the dsDNA copy of the virus into its host chromosome. The Psi-sequence acts as a signal sequence and is necessary for packaging RNA with the reporter or therapeutic gene in virions. Viral proteins which make virus shells are provided in the packaging cell line, but are not in context of the LTRs and Psi-sequences and so are not packaged into virions. Thus, virus particles are produced that are replication deficient, so are designed to be unable to continue to infect their host after they deliver their therapeutic content (http://biology.kenyon.edu /slonc/gene-web/Lentiviral/Lentivi2.html).

Three Plasmid Expression System

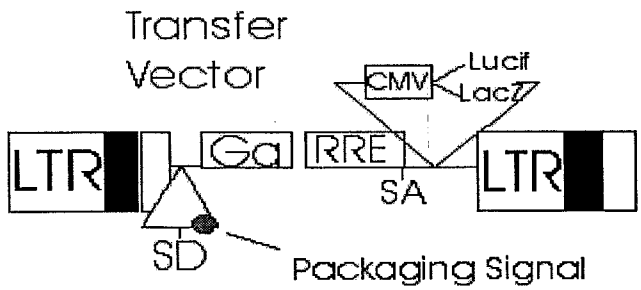
Lentiviral vectors are usually created in a transient transfection system in which a cell line is transfected with three separate plasmid expression systems. These include the transfer vector plasmid (portions of the HIV provirus), the packaging plasmid or construct, and a plasmid with the heterologous envelop gene (ENV) of a different virus. The three plasmid components of the vector are put into a packaging

cell which is then inserted into the HIV shell. The virus portions of the vector contain insert sequences so that the virus cannot replicate inside the cell system.

Lentiviral vectors are usually created in a transient transfection system in which a cell line is transfected with three separate plasmid expression systems. These include the tramsfer vector plasmid (portions of the HIV provirus), the packaging plasmid or construct, and a plasmid with the heterologous envelop gene (ENV) of a different virus. The three plasmid components of the vector are put into a packaging cell which is then inserted into the HIV shell. The virus portions of the vector contain insert sequences so that the virus cannot replicate inside the cell system.

1) Transfer vector plasmid

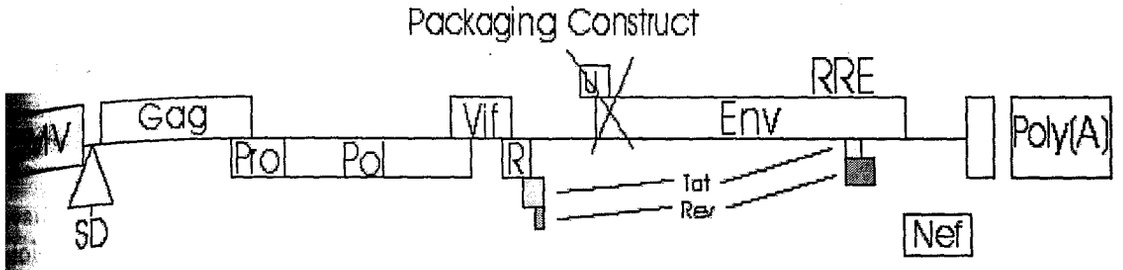
The transfer vector plasmid contains cis-acting genetic sequences necessary for the vector to infect the target cell and for transfer of the therapeutic (or reporter) gene and contains restriction sites for insertion of desired genes. The 3' and 5' LTRs, the original envelop proteins, and gag sequence promoter have been removed.



Source: (http://biology.kenyon.edu /slonc/gene-web/Lentiviral/Lentivi2.html)

2) Packaging plasmid

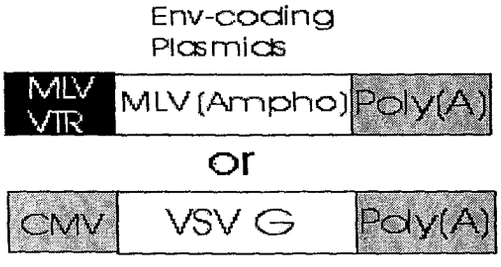
The packaging plasmid is the backbone of the virus system and is also known as pCMVAR9. In this plasmid are found the elements required for vector packaging such as structural proteins, HIV genes (except the gene env which codes for infection of T cells, or the vector would only be able to infect these cells), and the enzymes that generate vector particles. Also contained is the human cytomegalovirus (hCMV) which is responsible for the expression of the virus proteins during translation. The packaging signals and their adjacent signals are removed so the parts responsible for packaging the viral DNA have been separated from the parts that activate them. Thus, the packaging sequences will not be incorporated into the viral genome and the virus will not reproduce after it has infected the host cell. Previous HIV vectors used two plasmids as the packaging plasmid contained the viral envelop gene. However, in the newer, better vectors the packaging plasmid lacks a viral envelop gene because this has been shown to be more desirable in terms of titer (minimum volume needed to cause a particular result in titration), stability, and broad range of target cells.



ource: (http://biology.kenyon.edu /slonc/gene-web/Lentiviral/Lentivi2.html)

Envelop gene of third plasmid

The third plasmid's envelope gene of a different virus specifies what type of ll to target and infect instead of the T cells. Normally HIV can infect only helper T-lls because they use their gp120 protein to bind to the CD4 receptor. However, it is ssible to genetically exchange the CD4 receptor-binding protein for another protein at codes for the different cell type on which gene therapy will be performed. This ves the HIV lentiviral vector a broad range of possible target cells. There are two pes of heterologous envelope proteins. The amphoteric envelop of MLV, another pe of vector, is transcribed first followed by the transcription of the G glycoproteins f the vesicular stomatitis virus, known as VSV-G. Both of these help to provide ability to the vector by bringing together the particles that were made by the ackaging plasmid pCMVAR9.



Source: (http://biology.kenyon.edu /slonc/gene-web/Lentiviral/Lentivi2.html)

Uses for HIV lentiviral vectors

Scientists have recently been using the HIV lentiviral vector to repair neurons. HIV is also being developed as a delivery system to provide successful gene therapy in many diseases such as metabolic diseases, cancer, viral infection, cystic fibrosis, muscular dystrophy, hemophilia, retinitis pigmentosa, and maybe even Alzheimer's disease.

Concerns with using HIV lentiviral vectors

There is still concern with using lentiviral vectors for safety reasons. One concern involves the possibility that the HIV could self-replicate and could be produced during manufacture of the vector in the packaging cell line or in the target cells by a process of recombination. Thus, the person undergoing gene therapy would also be infected with HIV in addition to the new therapeutic gene. A self-replicating

ous vector could cause cancer by inserting itself into the host genome and te a neighboring proto-oncogene, thus causing mutagenesis.

ovirus vectors

Recombinant forms of adenoviruses based on human serotype 5 were indeed able and had been characterized extensively as molecular tools for studying basic gical processes. One could disable the adenovirus by deleting an important diate early gene (i.e., E1) and could incorporate into the vector genome gene cassettes in excess of 6 kb. The E1 deleted adenoviruses were easily grown igh quantities through propagation in an available cell line called 293, which ors a 5' region of the human Ad5 genome (45).

Adonovirus vectors are attractive candidates for gene transfer due to their pendence of host cell replication and low pathogenicity in man. Adenoviruses n relatively stable particles which can be readily obtained at high titers of 10^{11} to plaque forming units (pfu) per ml making systemic clinical administration ible (36). The adenovirus genome is also relatively easy to manipulate, facilitating e of vector generation, and does not normally undergo rearrangement at a high rate 6). However, adenoviral vectors show some limitations with association of bstantial inflammatory responses in the areas of lung where vector-encoded nsgene was expressed. Transgene expression was transient that diminished to detectable level within 2 to 4 weeks (47). Adenovirus can effectively activate otoxic T lymphocytes (CTLs) to vector transduced cells using the transgene, as ell as the products of other viral genes as antigenic targets (47, 48). Transgene pression, therefore, exterminates by CTL response via the tranduced cells destruction and contributes to inflammation. Rhesus monkeys receiving adenoviral gene therapy developed mild to moderate acute hepatitis 1 to 3 weeks which mediated by MHC class I and T-cell response (49). Furthermore, the adenovirus system was the relative inefficiency by which gene transfer was achieved (50).

Adeno-associated virus vectors

Adeno-associated viruses (AAV) belong to the family Parvoviridae and consist of nonevelped icosahedral virions of 18-26 nm diameter, with a linear single-strnaded DNA genome of 4680 nucleotides for the most characterized AAV2 strain. The genome consists of two coding regions, *cap* and *rep*, which are flanked by inverted terminal repeats (ITRs) at either end of the genome and an encapsidation signal. The cap gene encodes the capsid (cap) proteins and rep encodes for proteins involved in replication and integration functions (36). In the absence of helper virus, the wild-type AAV genome integrates in a site-specific manner via its inverted terminal repeats (ITRs) into the q13.3-qter of human chromosome 19. Furthermore, the relatively higher stability, viral titers, and transduction efficiency of AAV add to the desirable features of this type of vector (51). Naturally, AAV is replication incompetent and requires additional genes from a helper virus infection which in nature are generally complemented by adenovirus or herpes simplex virus co-infection (36). There is no consistent indication of AAV infection being associated with human disease, although adverse effects on early embryos have been suggested from *ex vivo* studies (52).

AAV serotype 2 (AAV2) is the first AAV that was vectored for gene transfer applications by replacement of the rep and cap genes with a transgene of interest, retaining the terminal repeats and packaging of the genome (36, 51, 53). When helper plasmids and vector plasmids were co-transfected into permissive cells (usually

human embryonic kidney 293 cells) that were also infected with a helper adenovirus, those cells packaged rAAV2 virions containing only the therapeutic vector genomes. However, the restricted tissue tropism of AAV and its low transduction efficiency have limited its further development as vector (53). Alternative AAV serotypes such as AAV1, 4, 5 and 6 have developed by improving potency and broadened tropism of the AAV vector (53). Recombinant AAV2 is being elucidate as gene therapy for treating cystic fibrosis in clinical trial phase I and II (52).

Application of gene therapy

Cancer Immunotherapy

Gene therapy-induced expression of immunostimulatory molecules such as B7 at tumor cell level may evoke antitumor immune mechanisms by recruiting and enhancing viability of antigen-processing cells and specific tumoricidal lymphocytes (54). In addition, tumor cells can be transduced with cytokine genes (IL-2, 4, and 12) or combine with chimeric anti-tumor antibody to enhance the efficacy of immune response. Zhang X, et al. contructed immunocytokine called IL-2-183B2scFv that composed of the single-chain Fv of a monoclonal antibody, COC183B2, specific for an ovarian carcinoma-associated antigen (OC183B2), fused with the coding sequence of interleukin 2 (IL-2). This immunocytokine, IL-2-183B2scFv, maintaining functions of both the antibody and IL-2 was able to target IL-2 to tumor cells that overexpress OC183B2, and stimulate the proliferation of an IL-2-dependent cell line which can be treat ovarian cancer in the future (55). Another form of cancer vaccine is adoptive immunotherapy which T-cells are stimulated *ex vivo* by activating and expanding autologous tumour-reactive T-cell populations to large numbers of cells that are then transferred back to the patient {Dudley, 2003 #70}. Study in malignant glioma, for example, the vaccination combined with adoptive immunotherapy with *in vitro* activated cells demonstrated tumor response and improved survival despite a condition of advanced disease and immunosuppression resulting from previous treatment or tumor burden {Sloan, 2000 #71}. Furthermore, in previous clinical trials of adoptive immunotherapy using cloned melanoma-reactive CD8+ T cells, the transferred cells had high *in vitro* reactivity against melanoma Ags, yet objective clinical responses were not seen in these patients {Dudley, 2002 #72}.

Suppression of oncegene at transcriptional level by deliver transcriptional repressor acting on promoter of oncogene or post-transcriptional level by deliver ribozyme, antisense, siRNA, dominant negative molecule. For example, the E6 and E7 genes of human papillomaviruses encode proteins, that interfere with the function of the tumor suppressor proteins p53 and Rb. When introduced the antisense RNA transcripts of the E6 and E7 genes of HPV type 16 into human cervical cancer SiHa cells harboring HPV 16, E6 and E7 protein expression was suppressed, and p53 and Rb protein expression increased. The final result was the apoptosis of infected cells (56). Moreover, cancer cells can be replaced with tumor suppressor gene, for instance, Rb and p53 (57), or cell death-inducing gene, such as caspases and TRAIL-mediated apoptosis. Recently, using ES cell-derived astrocytes as gene therapy vectors transferring tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) induced human malignant glioma cells apoptosis (58). A tumor-specific killer, Apoptin, targets asialoglycoprotein receptor (ASGPR) presented only on the surface of hepatocytes can be used as gene inducing apoptosis. *In vivo* study showed that systemic delivery of asialoglycoprotein-Apoptin specifically induces apoptosis in malignant hepatocytes but not in normal cells (59). The activation of an inducible

pase (iCaspase-9) mediates apoptosis of neovascular endothelial cells, and ercomes the prosurvival effect of vascular endothelial growth factor or basic broblast growth factor. Recombinant adenoviral vectors containing iCaspase-9 duced neovascular disruption in transduced human dermal microvascular dothelial cells (60).

Suicide gene therapy is the other way for cancer treatment. A streptolysin O-pressing adenovirus was injected into human cervical cancer cell-derived tumors own in a nude mouse model. The result showed an decrease in average size of mors which was a result of pore-forming toxins genes carried by adenoviral vector 61). Thymidine kinase/ganciclovir (TK/GC) system is another suicide gene therapy udy. The Herpes simplex-derived thymidine kinase (TK) converts ganciclovir (GC) o ganciclovir-monophosphate, which is further phosphorylated by cellular kinases to xic ganciclovir-triphosphate. Mammalian cells lack TK, thus GC causes toxic ffects only in cells transfected with TK. A widely applied cancer chemotherapy gent is 5-flurouracil (5-FU). In mammalian cells, 5-FU is metabolized first into ucleoside fluorouridine by uridine phosphorylase and then phosphorylated into 5-fluoro-2'-uridine-5'-monophosphate (FUMP) by uridine kinase (62). From this system transfected tumor cells appear to be capable of inducing the death of neighboring untransfected cells. This cell kill is called the "bystander effect". This suggests that cell-to-cell contact is necessary for the bystander effect and cells might communicate through gap junctions (62). TK/GC protocol has been used widely in treatment cancer and many type of cancer therapies are under clinical trial such as lymphoma {Cooper, 2006 #68}, prostate cancer {Ayala, 2006 #69}, glioblastoma multiforme {Rainov, 2000 #67}.

rences:

Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone hase Gene into Petunia Results in Reversible Co-Suppression of Homologous es in trans. Plant Cell 1990;2(4):279-289.

Vaucheret H, Nussaume L, Palauqui JC, Quillere I, Elmayan T. A nscriptionally Active State Is Required for Post-Transcriptional Silencing uppression) of Nitrate Reductase Host Genes and Transgenes 105/tpc.9.8.1495. Plant Cell 1997;9(8):1495-1504.

Kneissl ML, Deikman J. The Tomato E8 Gene Influences Ethylene synthesis in Fruit but Not in Flowers. Plant Physiol 1996;112(2):537-547.

Howard K, Foster SG, Cooley RN, Caten CE. Disruption, Replacement, and suppression of Nitrate Assimilation Genes inStagonospora nodorum. Fungal netics and Biology 1999;26(2):152-162.

Sheehy RE, Kramer M, Hiatt WR. Reduction of polygalacturonase activity in mato fruit by antisense RNA. Proc Natl Acad Sci U S A 1988;85(23):8805-8809.

Commission E. Opinion of the scientific committee on food on the evaluation toxiological information related to the safety assessment of genetically modified matoes. In; 2000.

Guo S, Kemphues KJ. par-1, a gene required for establishing polarity in C. legans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell 1995;81(4):611-20.

Rocheleau CE, Downs WD, Lin R, Wittmann C, Bei Y, Cha YH, et al. Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. Cell 1997;90(4):707-16.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 1998;391(6669):806-11.

Kim VN. Small RNAs: classification, biogenesis, and function. Mol Cells 2005;19(1):1-15.

Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature 2004;431(7006):343-9.

Buchon N, Vaury C. RNAi: a defensive RNA-silencing against viruses and transposable elements. Heredity 2006;96(2):195-202.

Heidersbach A, Gaspar-Maia A, McManus MT, Ramalho-Santos M. RNA interference in embryonic stem cells and the prospects for future therapies. Gene Ther 2006;13(6):478-86.

Tomari Y, Zamore PD. Perspective: machines for RNAi. Genes Dev 2005;19(5):517-29.

Hammond SM. Dicing and slicing: the core machinery of the RNA interference pathway. FEBS Lett 2005;579(26):5822-9.

Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116(2):281-97.

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science 2006;313(5785):320-4.

Schramke V, Allshire R. Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science 2003;301(5636):1069-74.

19. Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, et al. Characterization of the piRNA complex from rat testes. Science 2006;313(5785):363-7.
20. Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM. EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol 2000;10(4):169-78.
21. Forrest EC, Cogoni C, Macino G. The RNA-dependent RNA polymerase, QDE-1, is a rate-limiting factor in post-transcriptional gene silencing in Neurospora crassa. Nucleic Acids Res 2004;32(7):2123-8.
22. Svoboda P, Stein P, Anger M, Bernstein E, Hannon GJ, Schultz RM. RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. Dev Biol 2004;269(1):276-85.
23. Lecellier CH, Dunoyer P, Arar K, Lehmann-Che J, Eyquem S, Himber C, et al. A cellular microRNA mediates antiviral defense in human cells. Science 2005;308(5721):557-60.
24. Nybakken K, Vokes SA, Lin TY, McMahon AP, Perrimon N. A genome-wide RNA interference screen in Drosophila melanogaster cells for new components of the Hh signaling pathway. Nat Genet 2005;37(12):1323-32.
25. Zheng L, Liu J, Batalov S, Zhou D, Orth A, Ding S, et al. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. Proc Natl Acad Sci U S A 2004;101(1):135-40.
26. McManus MT, Sharp PA. Gene silencing in mammals by small interfering RNAs. Nat Rev Genet 2002;3(10):737-47.
27. Gerrard L, Zhao D, Clark AJ, Cui W. Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. Stem Cells 2005;23(1):124-33.
28. Lee NS, Rossi JJ. Control of HIV-1 replication by RNA interference. Virus Res 2004;102(1):53-8.
29. Dykxhoorn DM, Palliser D, Lieberman J. The silent treatment: siRNAs as small molecule drugs. Gene Ther 2006;13(6):541-52.
30. Aronin N. Target selectivity in mRNA silencing. 2006;13(6):509-516.
31. Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell 2003;115(2):199-208.
32. Mittal V. Improving the efficiency of RNA interference in mammals. Nat Rev Genet 2004;5(5):355-65.
33. Scacheri PC, Rozenblatt-Rosen O, Caplen NJ, Wolfsberg TG, Umayam L, Lee JC, et al. Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. Proc Natl Acad Sci U S A 2004;101(7):1892-7.
34. Jackson AL, Burchard J, Schelter J, Chau BN, Cleary M, Lim L, et al. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. Rna 2006;12(7):1179-87.
35. Fedorov Y, King A, Anderson E, Karpilow J, Ilsley D, Marshall W, et al. Different delivery methods-different expression profiles. Nat Methods 2005;2(4):241.
36. Lemoine N. Understanding Gene therapy. 2nd ed: BIOS Scientific Publishers; 1999.
37. Davidson JM, Krieg T, Eming SA. Particle-mediated gene therapy of wounds. Wound Repair Regen 2000;8(6):452-9.
38. Schakman O, Thissen JP. Gene therapy with anabolic growth factors to prevent muscle atrophy. Curr Opin Clin Nutr Metab Care 2006;9(3):207-13.

Rizzuto G, Cappelletti M, Maione D, Savino R, Lazzaro D, Costa P, et al. Efficient and regulated erythropoietin production by naked DNA injection and muscle electroporation. Proc Natl Acad Sci U S A 1999;96(11):6417-22.

Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, et al. *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science 1996;272(5259):263-7.

Strang BL, Takeuchi Y, Relander T, Richter J, Bailey R, Sanders DA, et al. Human immunodeficiency virus type 1 vectors with alphavirus envelope glycoproteins produced from stable packaging cells. J Virol 2005;79(3):1765-71.

Ikeda Y, Takeuchi Y, Martin F, Cosset FL, Mitrophanous K, Collins M. Continuous high-titer HIV-1 vector production. Nat Biotechnol 2003;21(5):569-72.

Ni Y, Sun S, Oparaocha I, Humeau L, Davis B, Cohen R, et al. Generation of a packaging cell line for prolonged large-scale production of high-titer HIV-1-based lentiviral vector. J Gene Med 2005;7(6):818-34.

Laakso MM, Sutton RE. Replicative fidelity of lentiviral vectors produced by transient transfection. Virology 2006;348(2):406-17.

Wilson JM. Adeno-associated virus and lentivirus pseudotypes for lung-directed gene therapy. Proc Am Thorac Soc 2004;1(4):309-14.

Bett AJ, Haddara W, Prevec L, Graham FL. An efficient and flexible system for construction of adenovirus vectors with insertions or deletions in early regions 1 and 3. Proc Natl Acad Sci U S A 1994;91(19):8802-6.

Yang Y, Li Q, Ertl HC, Wilson JM. Cellular and humoral immune responses to viral antigens create barriers to lung-directed gene therapy with recombinant adenoviruses. J Virol 1995;69(4):2004-15.

Yang Y, Su Q, Wilson JM. Role of viral antigens in destructive cellular immune responses to adenovirus vector-transduced cells in mouse lungs. J Virol 1996;70(10):7209-12.

Lu H, Sullivan D, Gerber MA, Dash S. Adenovirus induced acute hepatitis in non-human primates after liver-directed gene therapy. Chin Med J (Engl) 2002;115(5):726-31.

Grubb BR, Pickles RJ, Ye H, Yankaskas JR, Vick RN, Engelhardt JF, et al. Inefficient gene transfer by adenovirus vector to cystic fibrosis airway epithelia of mice and humans. Nature 1994;371(6500):802-6.

Qing K, Bachelot T, Mukherjee P, Wang XS, Peng L, Yoder MC, et al. Adeno-associated virus type 2-mediated transfer of ecotropic retrovirus receptor cDNA allows ecotropic retroviral transduction of established and primary human cells. J Virol 1997;71(7):5663-7.

Flotte TR. Adeno-associated virus-based gene therapy for inherited disorders. Pediatr Res 2005;58(6):1143-7.

Gao G, Vandenberghe LH, Wilson JM. New recombinant serotypes of AAV vectors. Curr Gene Ther 2005;5(3):285-97.

Collins CG, Tangney M, Larkin JO, Casey G, Whelan MC, Cashman J, et al. Local gene therapy of solid tumors with GM-CSF and B7-1 eradicates both treated and distal tumors. 2006.

Zhang X, Feng J, Ye X, Yao Y, Zhou P, Chen X. Development of an immunocytokine, IL-2-183B2scFv, for targeted immunotherapy of ovarian cancer. Gynecol Oncol 2006.

Hamada K, Shirakawa T, Gotoh A, Roth JA, Follen M. Adenovirus-mediated transfer of human papillomavirus 16 E6/E7 antisense RNA and induction of apoptosis in cervical cancer. Gynecol Oncol 2006.

Olaf J. C. Hellwinkel JM, Annika Pollmann, Hartmut Kabisch,. Osteosarcoma lines display variable individual reactions on wildtype p53 and Rb tumour-pressor transgenes. The Journal of Gene Medicine 2005;7(4):407-419.

Germano IM, Uzzaman M, Benveniste RJ, Zaurova M, Keller G. Apoptosis in an glioblastoma cells produced using embryonic stem cell-derived astrocytes ressing tumor necrosis factor-related apoptosis-inducing ligand. J Neurosurg 06;105(1):88-95.

Peng DJ, Sun J, Wang YZ, Tian J, Zhang YH, Noteborn MH, et al. Inhibition hepatocarcinoma by systemic delivery of Apoptin gene via the hepatic loglycoprotein receptor. Cancer Gene Ther 2006.

Song W, Sun Q, Dong Z, Spencer DM, Nunez G, Nor JE. Antiangiogenic e therapy: disruption of neovascular networks mediated by inducible caspase-9 livered with a transcriptionally targeted adenoviral vector. Gene Ther 05;12(4):320-9.

Yang WS, Park SO, Yoon AR, Yoo JY, Kim MK, Yun CO, et al. Suicide ncer gene therapy using pore-forming toxin, streptolysin O. Mol Cancer Ther 06;5(6):1610-9.

Aghi M, Kramm CM, Chou TC, Breakefield XO, Chiocca EA. Synergistic ticancer effects of ganciclovir/thymidine kinase and 5-fluorocytosine/cytosine aminase gene therapies. J Natl Cancer Inst 1998;90(5):370-80.

Part III

Principles of Techniques in Protein Study

Topic 15. Protein Expression Technique

Objective:

1. students should know how to express protein
2. students should know advantages and disadvantages of each protein expression system
3. understand the solubility of protein
4. know how to determine the protein concentration
5. understand each protein purification method
6. understand protein sequencing

Protein Expression

Application

Examples:

1. Antigen production
2. Biochemical or cell biology studies
3. structural studies

A. Protein Expression in *E.coli* : Prokaryotic system (most popular, use plasmid)

Advantage :

1. the techniques necessary to express protein are relatively simple
2. the amount of time to generate an overexpressing strain is very short
3. cheap to grow and able to manipulate the *E. coli* genetic and physiology to increase the protein product

Disadvantage:

1. expressed protein from E. coli are not properly modified
2. protein expressed in large amounts in *E. coli* often precipitate into insoluble aggregates called "inclusion bodies" from which they can only be recovered in an active form by solubilization in denaturing agents followed by careful renaturation
3. it is relatively difficult to arrange the secretion of large amounts of expressed proteins from *E. coli*, although it has often been possible to secrete small amounts into the periplasmic space and to recover them by osmotic shock

Procedure in summary

1. clone gene of interest into an appropriate vector
2. transfer recombinant DNA into *E. coli*
3. select the positive clones

w the positive clone in an appropriate medium e.g. LB containing antibiotic
late inclusion bodies containing your protein from bacteria (denaturated form of protein
kept at –20 C in high salt concentration), check the protein using SDS-PAGE to see if
protein is pure. If not, go to purification step using an appropriate method
ature the protein
ncentrate the protein
ep refolded protein at –80 C
nalysis

B. Protein Expression in Baculovirus Expression System: Eukaryotic system (use virus)

dvantage :

1. proteins are almost always express at high levels
2. expressed proteins are usually expressed in the proper cellular comparment (membrane proteins are usually localized to the membrane, nuclear proteins to the nucleus and secreted proteins secreted into the medium)
3. the expressed protein is often properly modified

Disadvantage:

1. the techniques to grow and work with the virus are still not very widely used and may be difficult for the beginner
2. proteins are not always properly modified
3. sophisticated and a lot of works

C. Protein Expression in Mammalian System: Eukaryotic system (use COS cell)

Advantage :

1. expressed proteins are usually properly modified
2. expressed proteins almost always accumulate in the correct cellular compartment

Disadvantage:

1. techniques are more difficult, time consuming and expensive than those used to express protein in *E. coli*
2. much more difficult to perform on a large scale; but they are quite practical for small- and medium-scale work by investigators already familiar with mammalian cell culture techniques

Note: The COS cell and virus procedures are suitable for rapid small- and medium-scale protein production

Strategy of Cloning for Protein Expression

ice of domain

tting the range of sequence to clone and express

cloning your protein, it pays to think carefully about what you will do with the expressed al. Will you solve a structure and/or will you want to do some biochemistry? In the worst f you clone in the wrong bit it may be useless for any kind of experiment, let alone for g the structure. If you need help in analysing the sequence, you are welcome to contact the

u want to practise using some sequence analysis software, try to get on internet.

oing the right bit of protein to express is vital to the success of an expression experiment. In ection, we want to get you thinking about the following topics.

- Is the protein globular, non-globular, membrane or multidomain?
- Have you guarded against errors, truncations, frameshifts in the sequence?
- Have you defined the borders of the domain (or domains)?
- Have you examined a multiple sequence alignment of the homologues?

Class of protein

Many classical metabolic enzymes consist of a single folded domain and the purified proteins are therefore relatively easy to handle. However, the majority (at least in Eukaryotes) of proteins are not like this at all. Leaving aside the specialist field of membrane proteins, the first thing to worry about is whether the protein is globular at all: if it isn't, you will not be able to determine it's structure (except as an induced fit in a complex), although you will still be able to investigate the protein by CD & NMR. Multidomain proteins very often consist of a mixture of globular domains interspersed with non-globular regions. The non-globular regions can be important for the function (*e.g.* containing phosphorylation, SH2-binding, SH3-binding sites and so on). To gain structural insight, it is often necessary to solve one module at a time. Defining the domain boundaries is important since the presence of random coil can confound crystallisation attempts and may lead to unpredictable properties of the protein at NMR concentrations. Sometimes it is very hard to precisely define borders and then it is best to clone several constructs in parallel and then work with the smallest folded construct.

Here is a checklist for the fundamental protein types:

- Globular
- Rods
- Elongated proteins with repetitive sequencse

- Non-globular
- Integral membrane
- Multidomain
- May combine elements of any other class

You will need to do a multiple sequence alignment to be sure that you know about your protein.

3. Errors, truncations, frameshifts, incorrect exon assembly

Unfortunately, many sequences in databases are of low quality. Usually you will get away with "point mutation" errors that affect only one amino acid. Although these will cause trouble if, as you should, you check your sample by Mass Spectroscopy.

http://www-db.embl-heidelberg.de:4321/emblGroups/g_65.html

However, sequences of incorrect length are likely to cause severe problems. Many sequences have not been reliably determined at their N-termini and an arbitrary internal methionine may have been designated as the N-terminus. Occasionally the C-terminus is also incomplete. Either terminus could be affected by **frameshift sequencing errors** which can also cause internal problems, concatenate adjacent reading frames, etc. Another major (and increasing) class of error in sequence databases is **incorrect gene prediction** and exon assembly. The latter can have consequences at least as dramatic as for frameshift errors.

The only way to guard against errors like this is to do comparative sequence analysis: again you need to do a multiple sequence alignment.

4. Defining the borders of domains

If you are lucky, your domain(s) may be in a domain database. You can go to the following domain servers: SMART, PFAM or PROSITE.

They will show you many known domains in your protein. But you cannot be completely sure that they have the domain borders correct. You need to examine the set of proteins that contain a given domain. Often, an N- or C-terminal location can be used to define a domain border. Tandemly repeated domains define the domain size but it is easy to wrongly permute the domain borders.

You can only solve the structure of a single domain if it is autonomously folded, say an SH2 or PH domain. However, there is a common class of small rod-forming repetitive domains that are not autonomous, including LRR, ankyrin, tetratricopeptide, armadillo, HEAT and so forth. These typically consist of two or three 2D-structural elements and can only be expressed and folded in groups, so are not usually suitable for NMR analysis (which delayed the structural investigation of this repeat class, although most have representative structures now.)

are also exceptions to the usual organisation of autonomous globular domains that can trouble. Sometimes large non-globular inserts are found: these are likely to cause both ding and crystallisation problems. Sometimes, domains are fused together - either when ent or if one domain is inserted into a loop of another domain. Sometimes, unforeseeable in extensions are present, as in the WW domain structure solved at EMBL some years ago: minimal domain is unfolded, trial and error with larger constructs and monitoring by CD & is the only way to find out what is going on, unless you want to try a different sequence.

laking a multiple sequence alignment

e a multiple sequence alignment is the best way to protect yourself from many potential lems, if you don't have one already to hand, now is the time to do it. Here is a brief guide to ecting sequences and aligning them. If more help is needed, contact mw or try to get on the ware on internet..

ere are two ways to collect a set of related proteins and in practice both methods are usually d. You can retrieve them by keyword search with a tool such as SRS.

this case, it is difficult to collect a clean set as keywords are inconsistently applied and it is y common to bring in some completely wrong proteins, as well as missing some real ones.

e alternative is to do a database search with BLAST or Smith-Waterman on the Bioccelerator d collect all the sequences this way. Web outputs from good servers should let you collect all sequences that are clearly related (*e.g.* at EMBL, we use SRS links to collect the top hits). ten you will need to go through a process of eliminating annoying sequence fragments that ill not help in any future analysis.

ll the sequences to be aligned need to be saved into a single file, which is easy if you collect em via Netscape. You can then align them all with Clustal X which runs on any computer, ncluding fast G3 Macs, or with the Clustal W website at the EBI. Clustal X is recommended as e colour-coded output helps you to see features in the sequences. You can print out coloured ostscript alignments too. Clustal X also has a **quality check** that helps to find dodgy parts of the lignment. These tend to be either misaligned, doubtfully aligned (*i.e.* the signal is too weak to e sure of any alignment) or frameshifted regions of sequences.

aving got your alignment, you should now check the sequences as discussed above *e.g.* for the rmini. You need to look for "globularity" of the sequences too and this is simply done by inding buried core residues *i.e.* conserved columns of hydrophobic residues, which will be nterspersed with unconserved columns for the surface residues. Often, conservation periodicities ndicate whether a block of sequence is alpha-helix or beta-strand. Non-globular regions are haracterised by hydrophilic residues, Prolines and Glycines and generally poor conservation with high tolerance of INDELs. Sometimes it will be useful to have a secondary structure rediction too: In this case try submitting the alignment to the PredictProtein server.

Conclusion

u cannot guarantee to design a perfectly successful cloning and expression experiment but r can optimise your chances of success. Comparative sequence analysis is the best way to ntrol for success in designing a protein expression experiment, just as Mass Spec, CD and 1D MR are essential controls to analyse the behaviour of the expressed product. Using these tools save a lot of frustration. In severe cases, the latter will reveal that you have to start again, so se approaches may need to be used iteratively until you get something to work.

ference

is reference gives some practical hints on analysing protein domains.

rk, P. and Gibson, T.J. (1996) Applying motif and profile searches. *Methods Enzymol. 266,* 2-184.

Choice of Expression System

ere are three systems that most suited for large-scale production of proteins:

cherichia coli. The expression of proteins in ***E. coli*** is the easiest, quickest and cheapest thod. There are many commercial and non-commercial expression vectors available with ferent N- and C-terminal tags and many different strains which are optimized for special plications (for local users: see *strain database*).

ast. Yeast is an eukaryotic organism and has some advantages and disadvantages over *E. coli.* ne of the major advantages is that yeast cultures can be grown to very high densities, which akes them especially useful for the production of isotope labeled protein for NMR. The two ost used yeast strains are *Saccharomyces cerevisiae* and the methylotrophic yeast *Pichia storis.*

culovirus infected insect cells. Insect cells are a higher eukaryotic system than yeast and are ble to carry out more complex post-translational modifications than the other two systems (see *omparison of expression systems*). They also have the best machinery for the folding of ammalian proteins and, therefore, give you the best chance of obtaining soluble protein when ou want to express a protein of mammalian origin. The disadvantages of insect cells are the gher costs and the longer duration before you get protein (usually 2 weeks).

o determine which system is the best choice, ask yourself the following questions:

- **What type of protein do I want to express?**

hen you would like to express a protein of prokaryotic origin, the obvious choice is to use *E. oli.* The method is quick and cheap and the organism has all the machinery necessary for olding and post-translational modifications.

se the protein is from an eukaryotic source, the method of choice will depend on more rs (see below).

o I get soluble protein when I express in *E. coli*?

for the expression of eukaryotic proteins the first method of choice is normally *E. coli* for bove mentioned reasons. However, many eukaryotic proteins don't fold properly in *E. coli* form insoluble aggregates (inclusion bodies). Sometimes it is possible to resolubilize the ein from the inclusion bodies or improve the solubility by expressing the protein at a lower perature. Also expression of your target protein as a fusion protein with a highly soluble ner such as glutathione-S-transferase (GST), maltose binding protein (MBP), or DsbA can rove its solubility. Often, however, it is better to change to an eukaryotic expression system use it is better equipped to fold proteins from an eukaryotic source. Thus, instead of trying 10 different *E. coli* constructs, it is better to **switch expression system**.

oes my protein need post-translational modifications for structure/activity?

ny proteins need to be modified following translation in order to become active and/or adapt proper structure. The simplest of these modifications is the removal of the N-terminal thionine residue, which can occur in all organisms. More complex modifications, like N- and lycosylation, phosphorylation, are exclusively carried out by eukaryotic cells. Keep in mind t not all eukaryotic cells carry out the same modifications. Check table 1 to find out which pression system carries out the post-translational modification(s) you are looking for.

What is the codon usage in my protein?

Not all of the 61 mRNA codons are used equally. The so-called major codons are those that cur in highly expressed genes, whereas the minor or rare codons tend to be in genes expressed a low level. Which of the 61 codons are the rare ones depends strongly on the organism. The don usage per organism can be found in the **Codon Usage Database** tp://www.embl.de/ExternalInfo/protein_unit/draft_frames/frame_which_express_ext.htm.

or more information on the low usage codons per organisms see table 1 and table 2.

1. The 8 least used codons in *E. coli*, yeast, *Drosophila*, and primates.

E. coli	yeast	*Drosophila*	primates	amino acid
AGG	AGG			arginine
AGA		AGA		arginine
AUA		AUA		isoleucine
CUA				leucine
CGA	CGA	CGA	CGA	arginine
CGG	CGG	CGG	CGG	arginine
CCC				proline
UCG			UCG	serine
	CGC		CGC	arginine
	CCG		CCG	proline
	CUC			leucine
	GCG		GCG	alanine
	ACG		ACG	threonine
		UUA		leucine
		GGG		glycine
		AGU		serine
		UGU		cysteine
			CGU	arginine

Source: Zhang *et al.* (1991) *Gene 105*, 61-72.

2. The usage of the 6 least used *E. coli* codons by selected other organisms.

frequencies of more than 15 per 1000 codons are shown in **bold**.

anism	AGG arginine	AGA arginine	CGA arginine	CUA leucine	AUA isoleucine	CCC proline
herichia coli	1.4	2.1	3.1	3.2	4.1	4.3
mo sapiens	11.0	11.3	6.1	6.5	6.9	**20.3**
sophila lanogaster	4.7	5.7	7.6	7.2	6.9	**18.6**
enorhabditis gans	3.8	**15.6**	11.5	7.9	9.8	4.3
ccharomyces revisiae	9.3	**21.3**	3.0	13.4	**17.8**	6.8
smodium leipanum	4.1	**20.2**	0.5	**15.2**	**33.2**	1.0
ostridium steurianus	2.4	**32.8**	0.8	6.0	**52.5**	8.5
rococcus honkoshii	**30.3**	**20.4**	1.0	**18.0**	**44.9**	10.1
ermus aquaticus	13.7	1.4	1.4	3.2	2.0.	**43.0**
abidopsis thaliana	10.9	**18.4**	6.0	9.8	12.6	5.2

urce: Strategies Newsletter (2000) Vol. 3 (1), p. 31.

Usually, the frequency of the codon usage reflects the abundance of their cognate tRNAs. Therefore, when the codon usage of the protein you would like to express differs significantly from the average codon usage of the expression host, this could cause problems during expression. The following problems are often encountered:

- **Interrupted translation**, which leads to a variety of truncated protein products.
- **Frame shifting**.
- **Misincorporation** of amino acids. For instance, lysine for arginine as a result of the AGA codon. This can be detected by mass spectroscopy since it causes a decrease in the molecular mass of the protein of 28 Da.

Inhibition of protein synthesis and cell growth.

consequence, the observed levels of expression are often low or there will be no expression . Especially in cases were rare codons are present at the 5'-end of the mRNA or where cutive rare codons are found expression levels are low and truncated protein products are d.

crease the expression levels of proteins containing rare codons in *E. coli,* two main methods vailable:

- Site-directed mutagenesis to replace the rare codons by more commonly used codons for the same residue; *e.g.* the rare argenines codons **AGA** and **AGG** by the *E. coli* preferred **CGC** codon.

- Co-expression of the genes which encode rare tRNAs. There are several commercial *E. coli* strains available that encode for a number of the rare codon genes:

BL21 (DE3) CodonPlus-RIL	**AGG/AGA** (arginine), **AUA** (isoleucine) and **CUA** (leucine)	Stratagene
BL21 (DE3) CodonPlus-RP	**AGG/AGA** (arginine) and **CCC** (proline)	Stratagene
Rosetta or Rosetta (DE3)	**AGG/AGA** (arginine), **CGG** (arginine), **AUA** (isoleucine) **CUA** (leucine), **CCC** (proline), and **GGA** (glycine)	Novagen

ften your will obtain a mixture of full-length protein and truncated species. Providing the otein with a C-terminal tag (*e.g.* His_6-tag) will help you to purify only the full-length protein ing affinity chromatography.

hen both above-mentioned methods fail to increase expression levels, it is time to change pression system and try to express your protein in yeast or insect cells. In cases where the otein contains many rare *E. coli* codons it is probably better to immediately start with an karyotic system.

References

Protein Expression. A practical approach (Higgins. S.J. & Hames, B.D., eds), Oxford University Press, 1999.

Kane, J.F. (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli. Current Opinions Biotechnol. 6,* 494-500.

Zhang, S., Zubay, G. & Goldman, E. (1991) Low-usage codons in *Escherichia coli,* yeast, fruit fly and primates. *Gene 105*, 61-72.

y, R., Drott, D., Yaeger, K. & Mierenhof, R. (2001) Overcoming the codon bias of *E. coli* nhanced protein expression. *inNovations 12*, 1-3.

hoice of Vector

you have made up your mind about which expression system you want to use, it's time to se the right expression vector. Also here we would like to limit ourselves to expression ors used in *Escherichia coli*, *Pichia pastoris* and baculovirus.

herichia coli vectors
Commercial vectors (see catalog)
Vector features (see catalog)

ia pastoris vectors
Commercial vectors (see catalog)
Vector features (see catalog)

culovirus vectors
Commercial vectors (see catalog)

basic architecture of an *E. coli* expression vector is shown in the figure below and contains following features:

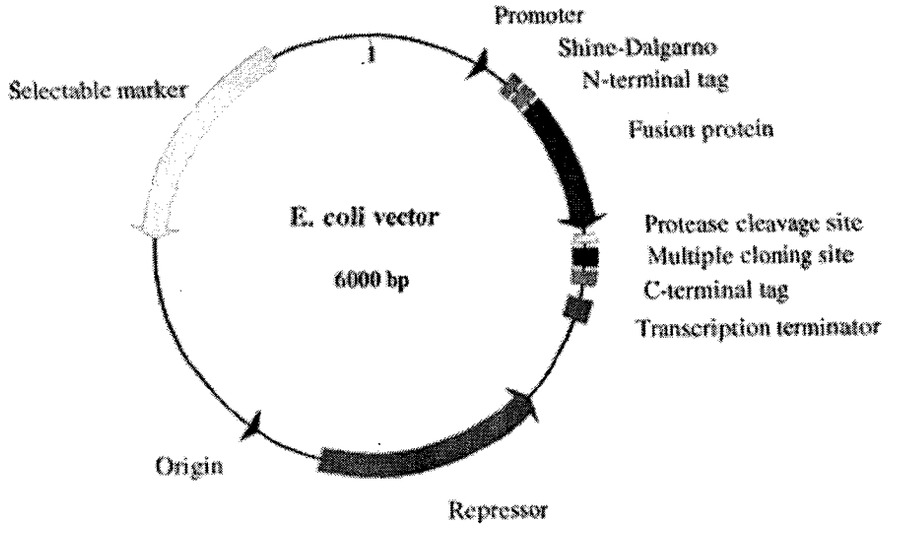


Figure 1. The general structure of *E. coli* vector. Source: Vector NT I program

table marker. In the absence of selective pressure plasmids are lost from the host. ially in the case of very high copy number plasmids and when plasmid-borne genes are to the host or otherwise significantly reduce its growth rate. The simplest way to address roblem is to express from the same plasmid an antibiotic-resistance marker and supplement edium with the appropriate antibiotic to kill plasmid-free cells. The most used antibiotics heir effective concentrations are listed in table 3.

e 3. Most used antibiotics and their effective concentrations

tibiotic		**concentration (mg/ml)**
p	ampicillin	100
b	carbenicillin	100
m	chloramphenicol	34
n	kanamycin	30
r	rifampicin	200
t	tetracyclin	12.5

he use of ampicillin requires special care. The selectable marker, b-lactamase, is secreted into he medium where it hydrolysis all of the ampicillin. This point is already reached when the ulture is barely turbid. From here on, cells that lack the plasmid will not be killed and could vergrow the culture. Some possible solution are:

- grow overnight cultures at 30°C or less.
- spin overnight cultures and resuspend the pellet in fresh medium to remove the produced b-lactamase.
- use the more stable carbenicillin instead of ampicillin.

Regulatory gene (repressor). Many promoters show leakiness in their expression *i.e.* gene products are expressed at low level before the addition of the inducer. This becomes a problem when the gene product is toxic for the host. This can be prevented by the constitutive expression of a repressor protein.

The **lac**-derived promoters are especially leaky. These promoters can be controlled by the insertion of a **lac**-operator sequence downstream the promoter and the expression of the **lac**-

ssor by host strains carrying the **lacI^q** allele (for medium copy number plasmids) or from me or a helper plasmid (for higher copy number plasmids). Alternatively, repression can be ved by the addition of 1% glucose to the culture medium.

in of replication. The origin of replication controls the plasmid copy number.

moter. The promotor initiates transcription and is positioned 10-100 nucleotides upstream of ibosome binding site. The ideal promoter exhibits several desirable features:

- It is strong enough to allow product accumulation up to 50% of the total cellular protein.
- It has a low basal expression level (*i.e.* it is tightly regulated to prevent product toxicity).
- It is easy to induce.

extensive list of possible promoters is given in table 4. The most used promoters are icated in red.

ble 4. Promoters

romoter	source	regulation	induction	level of expression
ac	*E. coli*	lacI, lacI^q	IPTG	low
acUV5	*E. coli*	lacI, lacI^q	IPTG	low
ac (hybrid)	*E. coli*	lacI, lacI^q	IPTG	moderately high
trc (hybrid)	*E. coli*	lacI, lacI^q	IPTG	moderately high
P_{syn} (synthetic)	*E. coli*	lacI, lacI^q	IPTG	
trp	*E. coli*		tryptophan starvation	
araBAD	*E. coli*	araC	l-arabinose	variable
lpp[a]	*E. coli*		IPTG, lactose	
lpp-lac (hybrid)	*E. coli*	lacI	IPTG	
phoA	*E. coli*	phoB (positive) phoR (negative)	phosphate starvation	
recA	*E. coli*	lexA	nalidixic acid	

U	*E. coli*		osmolarity	
-1	*E. coli*		glucose starvation	
A	*E. coli*		tetracyclin	
dA	*E. coli*	cadR	pH	
	E. coli	fnr	anearobic conditions	
	l	l cIts857	thermal (shift to 42°C)	moderately high
pA	*E. coli*		thermal (shift to below 20°C)	
	T7	l cIts857	thermal	
-lac erator	T7	$lacI^q$	IPTG	very high
-lac erator	T3	$lacI^q$	IPTG	
-lac erator	T5	lacI, $lacI^q$	IPTG	
gene 32	T4		T4 infection	
prM-lac erator	*Bacillus*	$lacI^q$	IPTG	
Hb	*Vitreoscilla*		oxygen	
rotein A	*S. aureus*			

the constitutive **lpp promoter** was converted into an inducible one by insertion of the lacUV5 romoter-operator region downstream.

References

Makrides, S. C. (1996) *Microbiol. Rev. 60,* 512-538.
Hannig, G. & Makrides, S. C. (1998) *TIBTECH 16,* 54-60.
Stevens, R. C. (2000) *Structures 8,* R177-R185.

anscription terminator. The transcription terminator reduces unwanted transcription and creases plasmid and mRNA stability.

ine-Delgarno sequence. The Shine-Dalgarno (SD) sequence is required for translation itiation and is complementary to the 3'-end of the 16S ribosomal RNA. The efficiency of nslation initiation at the start codon depends on the actual sequence. The concensus sequence 5'-**TAAGGAGG**-3'. It is positioned 4-14 nucleotides upstream the start codon with the timal spacing being 8 nucleotides. To avoid formation of secondary structures (which reduces pression levels) this region should be rich in A residues.

art codon. Initiation point of translation. In *E. coli* the most used start codon is **ATG**. **GTG** is ed in 8% of the cases. **TTG** and **TAA** are hardly used.

gs and fusion proteins. N- or C-terminal fusions of heterologous proteins to short peptides gs, check with commercial catalog of the vectors) or to other proteins (fusion partners, check th commercial catalog of the vectors) offer several potential advantages:

- **Improved expression.** Fusion of the N-terminus of a heterologous protein to the C-terminus of a highly-expressed fusion partner often results in high level expression of the fusion protein.
- **Improved solubility.** Fusion of the N-terminus of a heterologous protein to the C-terminus of a soluble fusion partner often improves the solubility of the fusion protein.
- **Improved detection.** Fusion of a protein to either terminus of a short peptide (epitope tag) or protein which is recognized by an antibody or a binding protein (Western blot analysis) or by biophysical methods (*e.g.* GFP by fluorescence) allows for detection of a protein during expression and purification.
- **Improved purification.** Simple purification schemes have been developed for proteins fused at either end to tags or proteins which bind specifically to affinity resins (check with commercial catalog of the vectors).

Protease cleavage site. Protease cleavage sites are often added to be able to remove a tag or fusion partner from the fusion protein after expression. Most commonly used proteases are listed in table 5 or check with commercial catalog of the vectors . However, cleavage is rarely complete and often additional purification steps are required.

ble 5. Protease cleavage sites

otease	**recognition sequence**	**source**
ctor Xa	Ile Glu/Asp Gly Arg |	different distributors
terokinase	Asp Asp Asp Asp Lys |	New England Biolabs
rombin	Leu Val Pro Arg | Gly Ser	different distributors
EV otease	Glu Asn Leu Tyr Phe Gln | Gly	Life Technologies
eScission	Leu Glu Val Leu Phe Gln | Gly Pro	Amersham Pharmacia Biotech

ultiple cloning site. A series of unique restriction sites that enables you to clone your gene of terest into the vector.

op codon. Termination of translation. There are 3 possible stop codons but **TAA** is preferred ecause it is less prone to read-through than **TAG** and **TGA**. The efficiency of termination is ncreased by using 2 or 3 stop codons in series.

D. PCR Strategy

Primer design

The gene of interest usually has to be amplified from genomic or vector DNA by PCR (polymerase chain reaction) before it can be cloned into an expression vector. The first step is the design of the necessary primers.

Important features are:

Primer sequence. Especially the 3'-end of the primer molecule is critical for the specificity and sensitivity of PCR. It is recommemded not to have:

- 3 of more **G** or **C** bases at this position. This may stabilize nonspecific annealing of the primer.
- a 3' **thymidine**, since it is more prone to mispriming than the other nucleotides.

Primer pairs should be checked for complementarity at the 3'-end. This often leads to primer-dimer formation.

ases at the 5'-end of the primer are less critical for primer annealing. Therefore, it is possible to d sequence elements, like restriction sites, to the 5'-end of the primer molecule.

mer length. Usually a primer length of **18-30 bases** is optimal for most PCR applications. orter primers could lead to amplification of nonspecific PCR products.

elting temperature (T_m). The specificity of PCR depends strongly on the melting temperature $_m$) of the primers (the temperature at which half of the primer has annealed to the template). ually good results are obtained when the T_m's for both primers are similar (within 2-4 °C) and ove 60°C. The T_m for a primer can be estimated using the following formula:

$$T_m = 2°C * (A + T) + 4°C * (C + G)$$

C content. The GC content of a primer should be between **40** and **60%**.

sign of the 5'-end primer

e 5'-end primer overlaps with the 5'-end of the gene of interest and should contain the lowing elements:

- **Restriction site.** The restriction site should be the same or provide the same sticky end to the first of the restriction enzymes in the multiple cloning site of the vector chosen to clone the gene of interest into. Alternatively, you could pick any restriction enzyme that gives a blunt end upon cleavage (see cloning). Often **Nco I** (CCA**TG**G) or **Nde I** (CA**TATG**) are chosen because the **ATG** within these sites can be used directly to create the ATG start codon and/or the ATG codon for the N-terminal methionine residue (see Utilization of the Nco I cloning site).
- **5'-extension to the restriction site.** Restriction enzymes cleave DNA much less efficient towards the end of a fragment. A 5' extension of the restriction site with 2-10 bases greatly increases the cleavage efficiency of most enzymes. Data on the effect of the extension length and sequence on the cleavage efficiencies of the most used restriction enzymes can be found in the reference appendix of the New England Biolabs catalog.
- **Start codon.** A start codon (usually **ATG**) should be included when the gene of interest is not expressed with an N-terminal tag or fusion partner or when an N-terminal methionine residue is present. It should be checked that the start codon and the gene of interest are in frame with an eventual N-terminal tag and/or fusion partner.
- **Overlap with the gene of interest.** The overlap between the primer and the gene of interest should be long enough to give a T_m of 60°C or more (calculated as shown above).

sign of the 3'-end primer

he 3'-end primer overlaps with the DNA strand complementory to the 3'-end of the gene of terest and should contain the following elements:

- **Restriction site.** The restriction site should be the same or provide the same sticky end to the second of the restriction enzymes in the multiple cloning site of the vector chosen to clone the gene of interest into. Alternatively, you could pick any restriction enzyme that gives a blunt end upon cleavage (see cloning).
- **5'-extension to the restriction site.** Restriction enzymes cleave DNA much less efficient towards the end of a fragment. A 5' extension of the restriction site with 2-10 bases greatly increases the cleavage efficiency of most enzymes. Data on the effect of the extension length and sequence on the cleavage efficiencies of the most used restriction enzymes can be found on pp. 210-211 of the 2000/01 catalog of New England Biolabs.
- **Stop codon(s).** A stop codon (**TAA** is preferred because it is less prone to read-through than **TAG** and **TGA**) should be included when no C-terminal tag is used. To increase the efficiency of termination it is possible to use 2 or 3 stop codons in series.
- **Overlap with the stand complement to the 3'-end of the gene of interest.** The overlap between the primer and the strand complement to the 3'-end of the gene of interest should be long enough to give a T_m of 60°C or more (calculated as shown above). It should be checked that the gene of interest is in frame with an eventual C-terminal tag.

ptimizing PCR

he success or failure of PCR depends on many factors:

rimer concentration. The concentration of each primer should be between 0.1 and 0.5 mM. or most applications **0.2 mM** produces satisfactory results. Too high primer concentrations crease the chance of mispriming, which results in nonspecific PCR products. Limiting primer oncentrations result in extremely inefficient PCR reactions. The primer concentration can be alculated as described in Preparation of oligo solutions.

NA template concentration. The concentration of DNA template depends on the source. ormally used concentration are **100-250 ng** for mammalian genomic DNA and **20 ng** for inearized plasmid DNA (circular plasmid DNA is slightly less efficiently amplified) per 50-ml eaction.

Concentration of dNTPs. The concentration of each dNTPs (dATP, dCTP, dGTP and dTTP) hould be **200 mM**. Too high concentrations of dNTPs inhibit the PCR reaction.

NA polymerase. The PCR reaction conditions and reaction times depend on the type of DNA olymerase used. Different polymerases are commercially available and are summarized in DNA olymerases.

ymerase buffer. All DNA polymerases are supplied with their own optimal polymerase ffer.

ndard Polymerase buffer:

10 mM Tris-HCl pH 8.3 (at room temperature)
50 mM KCl
1.5 mM $MgCl_2$

e standard polymerase buffer works well for a wide range of templates and primers but may t be optimal for any particular combination. Especially the concentration of Mg^{2+} ions is tical and should be optimized. A series of PCR experiments should be carried out with Mg^{2+} ncentrations varying from 1.5 to 4 mM in 0.5-mM steps. High concentrations of chelating ents (such as EDTA) and negatively charged ionic groups (such as phosphates) should be voided. Some suppliers of DNA polymerases have added NH_4^+ ions to their buffers. It has been hown that the presence of NH_4^+ ions results in a high specificity of the primer-template binding ver a broad temperature range.

C content of DNA template. PCR with GC-rich templates(>60%) are especially difficult. This mainly caused by the formation of stable secondary structures that stall or reduce the olymerase reaction. Good results have been obtained by the addition of **glycerol**, **DMSO** (5-0%), **formamide** (5-20%) or **tetramethylammonium chloride** (0.01-10 mM) to the reaction ix. Although little is known of the exact role of these chemicals in PCR. A commercial additive **Q-Solution** from Qiagen.

nnealing temperature. The optimal annealing temperature has to be determined xperimentally. As a starting point, an annealing temperature 5°C below the estimated melting mperature (T_m) can be used.

Cycle hold times. The optimal hold times for the denaturation and annealing steps in each cycle depend strongly on the design of the thermocycler and the wall thickness of the PCR tubes used and should be determined experimentally. The hold time for the extension step depends mainly on the length of the DNA template. As a rule of thumb we use 60 sec per kb. Ramp times are normally kept as short as possible.

Number of cycles. The number of cycles necessary to obtain a suffient amount of PCR product depends strongly on the concentration of the DNA template. In a typical PCR, the maximum amout of product is approx. 10^{12} copies of the template. Starting from one copy, the most efficient PCR would reach this level in 40 cycles. Depending on the nature of the DNA template, we start with many more copies and as a rule of thumb we carry out PCR with **25** cycles for plasmid DNA and **30-35** cycles for genomic DNA.

ibition of PCR by impurities.

urities introduced by the Plasmid Mini/Maxi Prep and/or DNA Purification kits can have ng inhibitory effects on PCR (see table below).

Substance	Inhibitory concentration
SDS	>0.005% (w/v)
phenol	>0.2% (v/v)
ethanol	>1% (v/v)
isopropanol	>1% (v/v)
sodium acetate	>5 mM
sodium chloride	>25 mM
EDTA	>0.5 mM

CR experiments

pical PCR experiment

ypical PCR experiment is carried out under the following conditions:

1X polymerase buffer
0.2 mM each primer
200 mM each dNTP
2.5 U DNA polymerase
DNA template*

* **100-250 ng** for mammalian genomic DNA and **20 ng** for linearized plasmid DNA per 50 ml reaction mix.

nd consists of **three phases**:

1. **Initial denaturation.** 3-5 min at 92-94°C.

2. **Amplification.** This phase consists of 25 to 40 cycles of:

- **Denaturation.** 10-30 sec at 92-94°C.

Annealing. 10-60 sec at the annealing temperature (50-68°C).

Extension. 60 sec * product length (in kb) at 72°C.

inal extension. 5-10 min at 72°C.

erimental conditions for a typical PCR experiment are described in the following protocol.

R with longer DNA templates

PCR with longer DNA templates (3-7 kb) better results are obtained by lowering the nsion tempearture to 68°C. No or lower amounts of PCR product is obtained at 72°C caused higher level of template depurination at this temperature.

NA polymerases

ith regard to PCR we only consider thermostable DNA polymerases.

aq DNA polymerase

aq DNA polymerase is a thermostable enzyme from the thermophilic eubacterium *Thermus uaticus*. It has a half-life of 40 min at 95°C. It is relatively cheap and is, therefore, the enzyme choice for routine and control experiments. Taq DNA polymerase produces PCR products ith a one base 3'-overhang (most often being an A).

aq DNA polymerase is sold by many manufacturers and can also be obtained in the presence of thermolabile inhibitor. During the initial denaturation step the inhibitor is denatured and active olymerase is released ("hot start"). This procedure improves the specificity of the PCR mplification by excluding the nonspecific reactions that occur at lower temperatures.

Proofreading DNA polymerases

Proofreading or high fidelity DNA polymerases are enzymes with a very low rate of nucleotide misincorporation. The first discovered and best known enzyme is **Pfu polymerase** (from *Pyrococcus furiosus*). This enzyme incorporates approx. 767,000 nucleotides before making an error. In comparison Taq DNA polymerase makes an error in approx. every 125,000 nucleotides. It is also one of the most thermostable DNA polymerases known. It can tolerate temperatures exceeding 95°C, enabling it to PCR amplify GC-rich targets. However, Pfu polymerase is more expensive and should, therefore, only be used when high accuracy is required. Pfu polymerase produces blunt-end PCR products.

tures of polymerases

mplify long DNA targets (5-35 kb) it is preferred to use a mixture of Taq and a proofreading A polymerase. Optimized enzyme mixtures can be obtained from different manufacturers.

Cloning Strategy

neral strategy

speed up protein production, we have adopted a strategy of parallel cloning and expression. gene of interest is cloned in parallel into a variety of expression vectors containing different and/or fusion partners, and into vectors for a variety of expression systems. This approach uld not only gain us a lot of time but also result in a larger number of successfully expressed teins.

cloning strategy consists of the following steps:

- the gene of interest is amplified by PCR.
- the PCR product is cloned into a specific cloning or expression vector using one of the cloning methods described below.
- the sequence of a positive clone is checked by sequence analysis.
- the gene of interest is subcloned into a variety of expression vectors (for different expression systems) using a fixed set of restriction enzymes or specific recombination sites.

r expression in *E. coli*, we have selected expression vectors with the following tags and fusion rtners:

Tag	fusion partner
N-terminal His_6-tag	
N-terminal His_6-tag	thioredoxin (**TrxA**)
N-terminal His_6-tag	glutathione-S-transferase (**GST**)
N-terminal His_6-tag	maltose binding protein (**MBP**)
N-terminal His_6-tag	disulfide oxidoreductase (**DsbA**)
N-terminal His_6-tag	**NusA**
C-terminal His_6-tag	

ing methods

Cloning using restriction enzymes

TA cloning and TOPO TA cloning

Recombination cloning systems:

GATEWAY Cloning Technology (Invitrogen)

Creator (Clontech)

Dicistronic cloning

nsformation/Transfection

- Transformation of plasmid DNA to competent *E. coli* cells
- Transfection of viral DNA to insect cell lines

oning using restriction enzymes

Restriction enzymes (restriction endonucleases) are proteins that cut DNA at (or close to) ecific recognition sites (see the catalogs of manufacturers or the Restriction Enzyme Database). Two types of restriction enzymes exist that differ in the way they cut the target DNA:

Blunt end cutters. These enzymes cut both strand of the target DNA at the same spot creating blunt ends.

Sticky end cutters. These enzymes cut both strand of the target DNA at different spots creating 3'- or 5'-overhangs of 1 to 4 nucleotides (so-called sticky ends).

To be able to clone a DNA insert into a cloning or expression vector, both have to be treated with two restriction enzymes that create compatible ends. At least one of the enzymes used should be a sticky end cutter to ensure that the insert is incorporated in the right orientation.
It will save you a lot of time when you could carry out the two digestions simultaneously (double digestion). Not all restriction enzymes work equally well in all commercially available buffers and, therefore, it is worthwhile to check (*e.g.* in the reference appendix of the New England Biolabs catalog) which enzymes are compatible and in which buffer. To ensure efficient digestion the two recognition sites should be more than 10 base pairs apart. If one of the enzymes is a poor cutter or if the sites are separated 10 base pairs or less, the digestions should be performed sequentially. The first digest should be done with the enzyme that is the poorest cutter

the second enzyme added after digestion has been verified by running a sample of the ction mix on an agarose gel.

our cloning work, we have selected two sticky end cutters that create different 5'-overhangs:

- 3'-end: **Nco I**. Its recognition site contains the **ATG** start codon.
- 5'-end: **BamH I**. It is cheap and active in most buffers.

s recommended to use the special **BamH I buffer** (New England Biolabs) for this double gestion.

ethylation of DNA

ector preparation

- Digestion of vector DNA using (preferably) two restriction enzymes.
- **Dephosphorylation of the ends** using calf intestine or shrimp alkaline phosphatase. This reduces the background of non-recombinants due to self-ligation of the vector (especially when a single site was used for cloning).
- **Purification of the digested vector** by agarose electrophoresis to remove residual nicked and supercoiled vector DNA and the small piece of DNA that was cut out by the digestions. This usually reduces strongly the background of non-recombinants due to the very efficient transformation of undigested vector.

sert preparation

- Digestion of insert DNA using (preferably) two restriction enzymes.
- **Purification of the digested insert.** Purification should be carried out by agarose gel electrophoresis when the insert is subcloned into a vector from a vector with the same selective marker or PCR amplified from a vector with the same selective marker. Otherwise, it can be purified using a commercial kit (such as Qiagen's PCR purification kit).

igation

he next step is the ligation of the insert into the linearized vector. This involves the formation of hosphodiester bonds between adjacent 5'-phosphate and 3'-hydroxyl residues, which can be talyzed by two different ligases: *E. coli* DNA ligase and bacteriophage T4 DNA ligase. The tter is the preferred enzyme because it can also join blunt-ended DNA fragments.

he efficiency of the ligation reaction depends on:

the **absolute DNA concentration**. The concentration should be high enough to ensure that intermolecular ligation is favored over self-ligation but not so high as to cause extensive formation of oligomeric molecules.
For pET vectors, good results are obtained at a vector DNA concentration or approx. **1 nM** (*i.e.* 50-100 ng vector DNA per 20-ml ligation mix).

- the **ratio between vector and insert DNA**. The maximum yield of the right recombinants is usually obtained using a molar ratio of insert to vector DNA of approx. 2. If the concentration of insert DNA is substantially lower than that of the vector, the ligation efficiency becomes very low.
  In practise, we set-up ligation reactions with a molar ratio of insert to vector DNA of **1:1**, **2:1**, and **3:1**.

- the **cloning strategy**. Higher yields of the right recombinant are obtained when the vector and insert have been prepared using two restriction enzymes and the digested vector has been gel-purified before the ligation reaction (as shown in the figure below).

ligation of blunt-ended fragments is less effective than that of sticky-ended ones. Blunt-end tion may be enhanced by:

- **high concentrations of blunt-ended DNA fragments**.
- a **high concentration of ligase** (10,000 NEB units/ml).
- a **low concentration of ATP** (0.1 mM).
- the **addition of PEG 4000** [5% (w/v)].
  Reference: Pheiffer, B.H. & Zimmerman, S.B. (1983) Nucleic Acid Res. 11, 7853-7871.

ferences

mbrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning:* A laboratory manual, 2nd d., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 1.63-1.70.

ET System Manual (1999), 8th Ed., Novagen.

cloning

he TA cloning method takes advantage of the terminal transferase activity of some DNA olymerases such as *Taq* polymerase. This enzyme adds a single, 3'-A overhang to each end of he PCR product. This makes it possible to clone this PCR product directly into a linearized loning vector with single, 3'-T overhangs. DNA polymerases with proofreading activity, such as *fu* polymerase, can not be used because they provide blunt-ended PCR products.

A cloning kits are available from different manufacturers.

have created a circular **TA cloning vector, pTA Plus**, based on pPCR-Script Amp ratagene).

gestion of this vector in two sequential reactions with *BamH* I and *Xcm* I gives a linearized ctor with 3'-T overhangs and a low background of non-recombinants.

OPO TA cloning

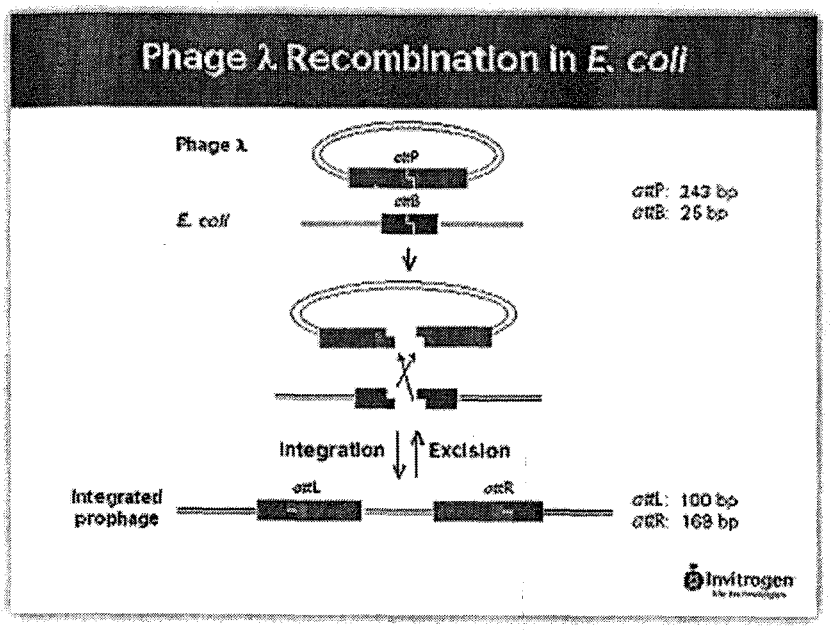
e **TOPO TA cloning** method combines the advantages of TA cloning with the ligation tivity of topoisomerase I. This allows direct ligation of PCR products in just **5 minutes.**

OPO TA cloning kits are available from *Invitrogen.*

ATEWAY Cloning Technology

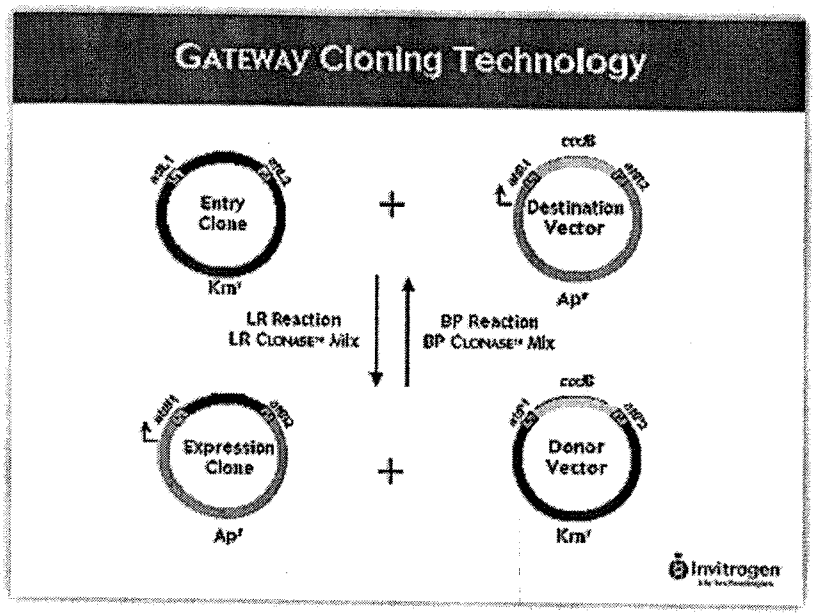
e **GATEWAY Cloning Technology** is based on the site-specific recombination system used phage l to integrate its DNA in the *E. coli* chromosome. Both organisms have specific ombination sites called *att*P in phage l site and *att*B in *E. coli.* The integration process ogeny) is catalyzed by 2 enzymes: the phage l encoded protein Int (Integrase) and the *E. coli* otein IHF (Integration Host Factor). Upon integration, the recombination between *att*B (25 nt) d *att*P (243 nt) sites generate *att*L (100 nt) and *att*R (168 nt) sites that flank the integrated age l DNA (see Figure 1).

e process is reversible and the excision is again catalyzed Int and IHF in combination with the age l protein Xis. The *att*L and *att*R sites surrounding the inserted phage DNA recombine site-ecifically during the excision event to reform the *att*P site in phage l and the *att*B site in the *E. li* chromosome.



urce: Invitrogen product catalog 2001

ure 2. Phage l recombination into the *E. coli* chromosome The **GATEWAY reactions** are *in vo* versions of the integration and excision reactions. To make the reactions directional two htly different and specific site were developed, *att*1 and *att*2 for each recombination site. ese sites react very specifically with each other. For instance in the **BP Reaction** *att*B1 only cts with *att*P1 resulting in *att*L1 and *att*R1, and *att*B2 only with *att*P2 giving *att*L2 and *att*R2. e reverse reaction (**LR Reaction**) shows the same specificity



ource: Invitrogen product catalog 2001

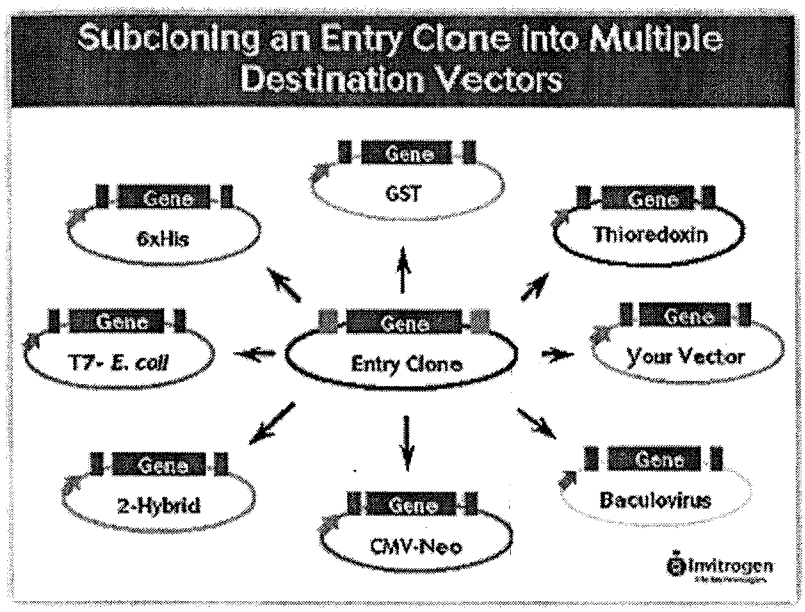
igure 3. The GATEWAY reactions The ultimate goal of the **GATEWAY reactions** is to make an expression clone. This is often a two step process

ep 1 Cloning the gene of interest into an *Entry Vector* using the **BP Reaction.**

ep 2 Subcloning the gene of interest from the Entry Clone (Step 1) into a *Destination Vector* using the **LR Reaction** producing the *Expression Clone.*

t's have a closer look at the **LR Reaction** of Step 2 (see also Figure 2). The gene of interest is oned into an *Entry Vector* and flanked by the *att*L1 and *att*L2 recombination sites. The *Entry ctor* is transcriptionally silent and contains the gene for kanamycin resistance (Km^r). To oduce the *Expression Clone* the gene has to be subcloned into a *Destination Vector* that ntains all the sequence information necessary for expression, the gene for ampicillin resistance Ap^r), and two recombination sites *(att*R1 and *att*R2) that flank a gene for negative selection, cdB (the encoded protein is toxic for the standard *E. coli* strains).

e two plasmids are mixed and the **LR CLONASE Enzyme Mix** is added. The reaction is ectional and specific, so that *att*L1 only reacts with *att*R1 and *att*L2 with *att*R2. The combination yields two constructs: the intended *Expression Clone* and a by-product (labelled Figure 2 as *Donor Vector*). The produced expression clone is under two forms of selection: the tibiotic resistance and the negative selection by the toxic ccdB protein. As a result high levels positive clones (typicaly more than 99%) are obtained after transformation to a standard oning or expression strain like DH5a or BL21 (DE3).



ource: Invitrogen product catalog 2001

igure 4. Subcloning an *Entry Clone* into multiple *Destination Vectors*_One of the main dvantages of the **GATEWAY Cloning Technology** is that once you have made an *Entry Clone* e gene of interest can be easily subcloned into a wide variety of *Destination Vectors* using the R **reaction.**

IPTG Induction and Extraction of Proteins from Bacteria

Induction in bacteria can be performed using one of two basic methods. Fast induction s not work for all proteins and can give you suboptimal yields. Slow induction can enhance solubility of some proteins. The method that's best for you will depend on your particular tein and the application. If you want optimal solubility both should be tested before scaling This protocol is generalized and will vary based on a variety of factors such as the bacterial in, recombinant protein, and parent plasmid.

t induction

1) From a relatively fresh plate (<4 weeks) pick a colony and grow O/N at 30deg C (or deg C) in 1-2ml LB+AMP(or other selection) in a 15ml snap cap tube on a rotator or shaker.

2) Dilute 1:50 (1:100 if 37deg O/N) in 2ml LB+AMP and grow 3-4 hours at 37deg C in ml snap cap tube in a rotator.

4) Prepare 1ml LB+AMP+1mM IPTG in a 15ml conical and prewarm to 37dec C about min before use.

5) After 3-4hrs remove 1ml from tubes at 37deg C and place in labeled 1.5ml tubes. Spin max, 30sec, RT, and remove supe. Freeze pellet at -20 until needed. THIS IS THE NINDUCED CONTROL.

6) Add prewarmed 1ml LB+AMP+1mM IPTG to 15ml snap cap tube and return to 37 g C for 3-4 hours. This will get the final volume back to 2ml and the final concentration of TG to 0.5mM.

7) After 3-4hrs transfer 1ml from induced sample to labeled 1.5ml tubes and spin at max, sec, RT, and remove supe. Freeze pellet at -20 until needed. THIS IS THE INDUCED AMPLE.

8) Sample preparation for SDS-PAGE: Add 100ul of 1X loading buffer (see solutions low) with 1% BME to uninduced and induced samples. Vortex for 10sec to 1min or until ere are no clumps of bacteria. Boil 3-5min, spin at max, 30sec, RT, and load 5-25ul (usually 0ul) depending on gel (amount of protein, size of pellet, Western, etc.).

Slow induction

For slow induction of protein follow fast induction protocol with the following changes:

6) Add **20deg** 1ml LB+AMP+1mM IPTG to 15ml snap cap tube and incubate rotating or shaking **at 20deg C for 12-16 hours**. This will get the final volume back to 2ml and the final concentration of IPTG to 0.5mM.

7) After **12-16hrs** transfer 1ml from induced sample to labeled 1.5ml tubes and spin at max, 30sec, RT, and remove supe. Freeze pellet at -20 until needed. THIS IS THE INDUCED SAMPLE.

NOTES for induction:

*Induction times vary from 2-5hrs.

*IPTG can be varied from 0.1-1.0M.

*If you boil your sample too long they will become viscous from total release of cellular DNA. You can still use them if you can find an area of low viscosity, however, its usually better just to repeat the experiment.

Extraction of Soluble Proteins

1) Wash the bacterial pellet with 2mls of ice cold STE (10mM Tris, pH 8.0; 150mM Nacl; 1mM EDTA) once.

2) Resuspend the bacterial pellet (from a 10ml induced culture) in 800ul of STE containing 100ug/ml of Lysozyme (added immediately prior to resuspension)

3) Incubate on ice for 15 minutes.

4) Add DTT to a final concentration of 5mM (we have stocks at 1M in -20).

5) Add protease inhibitors.

6) Bacteria are then lysed by the addition of N-laurylsarcosine (Sarkosyl) from a 10% (w/v) stock in STE. The final concentration of N-lauryl sarcosine should be 1.5%.

7) Sonicate the cells in the small bath sonicator with for 2 cycles (6min each). The sonicator automatically turns off after 6 minutes.

8) Centrifuge the lysate for 5min at 4deg C.

9) Transfer the supernatant to a new eppendorf tube and add Triton X-100 (from a 10% stock made in STE) to a final concentration of 2%.

10) Take 100ul of Ni-NTA agarose gel and wash twice by centrifugation with ice-cold PBS at 1000rpm for 1min. each

11) Add the beads to the eppendorf tube containing the lysate and the Triton X-100 tube and incubate on a nutator or rocker at RT for 30-60min.

12) Wash the beads 4X with 1ml ice-cold PBS containing 20mM imidiazole at 1000rpm for 1minute each.

13) Add 1X loading buffer with 1% BME, boil 3min, and analyze on SDS-PAGE.

NOTES for extraction:

*This basic protocol will work for FLAG, GST, and 6HIS epitope tags. It has not been tested for MBP, which does not respond will to detergents.

*The inclusion of Imidiazole is specific to 6HIS tags.

*TWEEN20 at 0.1-1% can also be incorporated into the wash buffer to reduce background if required.

Solutions

Loading Buffer (To make a 4X Stock)

50 mM Tris-HCl; pH 6.8 (2.0 ml 1M Tris-HCl; pH 6.8)

2% SDS (0.8 g SDS)

10% Glycerol (4.0 ml 10% Glycerol)

12.5 mM EDTA (1.0 ml 0.5 M EDTA)

0.02 % Bromophenol Blue (8 mg Bromophenol Blue)

Add fresh 2(or beta)-mercaptoethanol (BME) to 1% before using.

Topic 16. Protein Purification Technique

Protein Solubility

- The ionic strength dependence of protein solubility depends on the nature of the protein and on both the concentration and nature of the salt. Differential solubility in ammonium sulfate solutions forms the basis of an effective protein purification step.

The addition of salt can increase protein solubility (**salting in**) or decrease solubility (**salting out**).

Protein Purification

Before one can study the properties or structure of any molecule it generally needs to be isolated in a pure form first. Exceptions to this rule include the use of "knockouts" to define gene function in whole organisms. Purification from a complex mixture, such as a whole cell culture, requires the identification of a property or properties of the molecule in question that differentiates it from the others in the mixture. Techniques which separate molecules based on these properties can then be used as the basis of a purification procedure. Given the complexity of a biological system many different steps may need to be used to achieve a high degree of purity of the desired substance.

- **Source of Protein**
  - In order to purify a protein you need a source.
  - It might be blood or some other biological fluid, but most often it is a cell, usually a specific type - liver, muscle, yeast, bacteria, etc.
  - The cells must be broken open - homogenized - to release the protein in a soluble form.
  - Homogenization conditions must be worked out that release the protein from the cell without damaging the protein.
  - Membrane-bound proteins can also be purified, but different approaches are required.
- **Fractional Precipitation (based on differences in solubility properties)**
  - In concentrated salt solutions, usually ammonium sulfate is used, some proteins are more soluble than others. By varying the concentration of ammonium sulfate, one can achieve some limited purification of proteins. This technique is often used in the first step of a protein purification. In general,
  - Small proteins are more soluble than large proteins.
  - The larger the number of charged side chains, the more soluble the protein.
  - The solubility of protein A is totally independent of the solubility of protein B -- solubility depends on the surface properties of each individual kind of protein.

Separation and purification

eneral strategy of macromolecular purification is the same as is commonly employed for molecules. One caution is that biological macromolecules generally retain their activity relatively small range of conditions (eg it is not a good idea to boil most proteins or place in concentrated acid or base). The commonly used fractionation methods exploit properties particular molecule which are likely to be distinctive.

OPERTY	PROCEDURE
rge:	Ion exchange chromatography Electrophoresis Isoelectric focussing
larity	Adsorption chromatography Paper chromatography Reverse-phase chromatography Hydrophobic interaction chromatography
ze	Dialysis and ultrafiltration Gel electrophoresis Gel filtration chromatography Ultracentrifugation
ecificity	Affinity chromatography
lubility	Repetitive fractionation

Chromatography

romatography is a technique which exploits differential interactions between the molecules to separated in solution, the mobile phase, with molecules immobilised on a solid support, the tionary phase. The mobile phase is normally allowed to pass over the stationary phase in a lumn. The different interactions will cause different molecules to be eluted at different times. ose molecules which are adsorbed only weakly are eluted first, whereas those which interact st strongly with the column matrix are eluted last. Sometimes it is necessary to change the mposition of the buffer during the elution process to remove the most strongly adsorbed olecules.

Column Chromatography

lumn Chromatography

- *Study Questions/Answers from the Handbook for Organic Chemistry Lab*

column chromatography, the stationary phase, a solid adsorbent, is placed in a vertical glass sually) column and the mobile phase, a liquid, is added to the top and flows down through the lumn (by either gravity or external pressure). Column chromatography is generally used as a rification technique: it isolates desired compounds from a mixture.

The mixture to be analyzed by column chromatrography is applied to the top of the lumn. The liquid solvent (the eluent) is passed through the column by gravity or by the plication of air pressure. An equilibrium is established between the solute adsorbed on the sorbent and the eluting solvent flowing down through the column. Because the different mponents in the mixture have different interactions with the stationary and mobile phases, they ill be carried along with the mobile phase to varying degrees and a separation will be achieved. he individual components, or elutants, are collected as the solvent drips from the bottom of the olumn.

Column chromatography is separated into two categories, depending on how the solvent lows down the column. If the solvent is allowed to flow down the column by gravity, or ercolation, it is called **gravity column chromatography**. If the solvent is forced down the olumn by positive air pressure, it is called **flash chromatography**, a "state of the art" method urrently used in organic chemistry research laboratories The term "flash chromatography" was oined by Professor W. Clark Still because it can be done in a "flash."

he Adsorbent

Silica gel (SiO_2) and alumina (Al_2O_3) are two adsorbents commonly used by the organic chemist for column chromatography. These adsorbents are sold in different mesh sizes, as indicated by a number on the bottle label: "silica gel 60" or "silica gel 230-400" are a couple examples. This number refers to the mesh of the sieve used to size the silica, specifically, the number of holes in the mesh or sieve through which the crude silica particle mixture is passed in the manufacturing process. If there are more holes per unit area, those holes are smaller, thus allowing only smaller silica particles go through the sieve. The relationship is: the larger the mesh size, the smaller the adsorbent particles.

Adsorbent particle size affects how the solvent flows through the column. Smaller particles (higher mesh values) are used for flash chromatography, larger particles (lower mesh values) are used for gravity chromatography. For example, 70–230 silica gel is used for gravity columns and 230–400 mesh for flash columns.

Alumina is used more frequently in column chromatography than it is in TLC. Alumina is quite sensitive to the amount of water which is bound to it: the higher its water content, the fewer polar sites it has to bind organic compounds, and thus the less "sticky" it is. This stickiness or activity is designated as I, II, or III, with I being the most active. Alumina is usually purchased in activity I and deactivated with water before use according to specific procedures. Alumina comes in three forms: acidic, neutral, and basic. The neutral form of activity II or III, 150 mesh, is most commonly employed.

Silica gel and alumina are the only column chromatography adsorbents used in the CU organic chemistry teaching labs; please refer to the references for information on other column chromatography adsorbents.

The Solvent

The polarity of the solvent which is passed through the column affects the relative rates at which compounds move through the column. Polar solvents can more effectively compete with the polar molecules of a mixture for the polar sites on the adsorbent surface and will also better solvate the polar constituents. Consequently, a highly polar solvent will move even highly polar molecules rapidly through the column. If a solvent is too polar, movement becomes too rapid, and little or no separation of the components of a mixture will result. If a solvent is not polar enough, no compounds will elute from the column. Proper choice of an eluting solvent is thus crucial to the successful application of column chromatography as a separation technique. TLC is generally used to determine the system for a column chromatography separation.

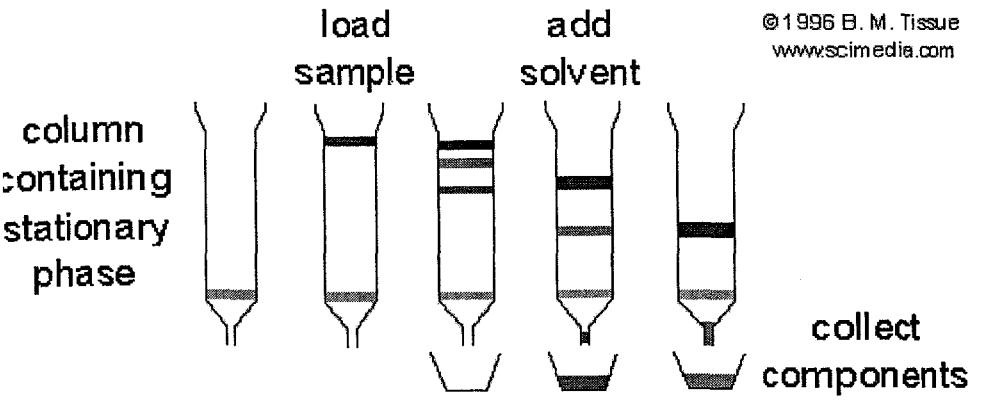
Often a series of increasingly polar solvent systems are used to elute a column. A non-polar solvent is first used to elute a less-polar compound. Once the less-polar compound is off the column, a more-polar solvent is added to the column to elute the more-polar compound.

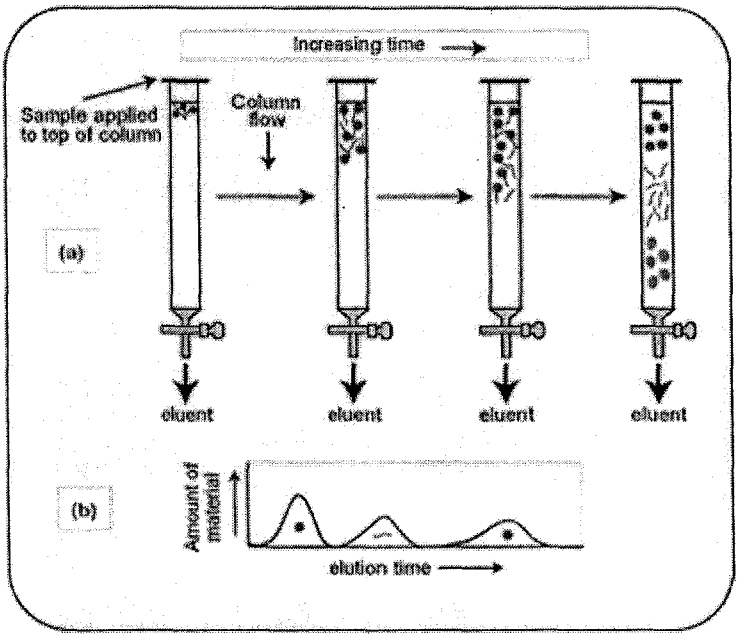
Interactions of the Compound and the Adsorbent

Compounds interact with the silica or alumina largely due to polar interactions. These interactions are discussed in the section on TLC.

Analysis of Column Eluants

If the compounds separated in a column chromatography procedure are colored, the progress of the separation can simply be monitored visually. More commonly, the compounds to be isolated from column chromatography are colorless. In this case, small fractions of the eluent are collected sequentially in labeled tubes and the composition of each fractions is analyzed by thin layer chromatography.





Source: http://matcmadison.edu

The term **Chromatography** applies to a wide variety of separation techniques which are based on the partitioning of a sample between a moving phase and a stationary phase.
Chromatography means:(color migration)
Distribution of analytes between phases

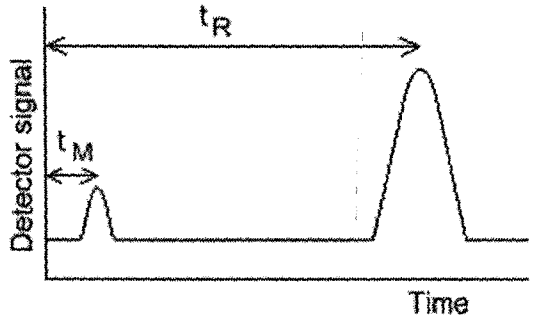
omatography Theory

distribution of analytes between phases can often be described quite simply. An analyte is in ilibrium between the two phases;

$$A_{mobile} \rightleftharpoons A_{stationary}$$

equilibrium constant, K, is termed the *partition coefficient*; defined as the molar centration of analyte in the stationary phase divided by the molar concentration of the analyte he mobile phase.

time between sample injection and an analyte peak reaching a detector at the end of the umn is termed the *retention time* (t_R). Each analyte in a sample will have a different retention The time taken for the mobile phase to pass through the column is called t_M.



term called the *retention factor, k'*, is often used to describe the migration rate of an analyte on column. You may also find it called the *capacity factor*. The retention factor for analyte A is efined as;

$$k'_A = t_R - t_M / t_M$$

$_R$ and t_M are easily obtained from a chromatogram. When an analytes retention factor is less han one, elution is so fast that accurate determination of the retention time is very difficult. High etention factors (greater than 20) mean that elution takes a very long time. Ideally, the retention factor for an analyte is between one and five.

We define a quantity called the *selectivity factor*, a , which describes the separation of two species (A and B) on the column;

$$a = k'_B / k'_A$$

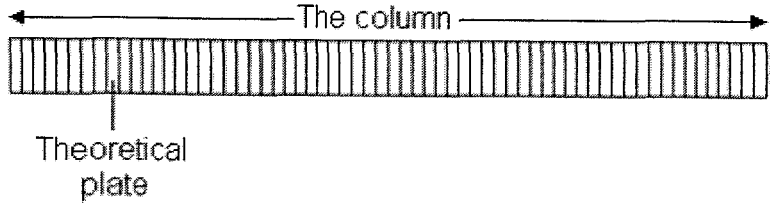
When calculating the selectivity factor, **species A elutes faster than species B. The selectivity factor is always greater than one.**

Band broadening and column efficiency

To obtain optimal separations, sharp, symmetrical chromatographic peaks must be obtained. This means that band broadening must be limited. It is also beneficial to measure the efficiency of the column.

The Theoretical Plate Model of Chromatography

plate model supposes that the chromatographic column is contains a large number of rate layers, called *theoretical plates*. Separate equilibrations of the sample between the ionary and mobile phase occur in these "plates". The analyte moves down the column by sfer of equilibrated mobile phase from one plate to the next.



important to remember that the plates do not really exist; they are a figment of the gination that helps us understand the processes at work in the column.They also serve as a of measuring column efficiency, either by stating the number of theoretical plates in a mn, N (the more plates the better), or by stating the plate height; the *Height Equivalent to a oretical Plate* (the smaller the better).

e length of the column is L, then the HETP is

$$\text{HETP} = L / N$$

number of theoretical plates that a real column possesses can be found by examining a omatographic peak after elution;

$$N = \frac{5.55\ t_R^{\,2}}{w_{1/2}^{\,2}}$$

ere $w_{1/2}$ is the peak width at half-height.

can be seen from this equation, columns behave as if they have different numbers of plates different solutes in a mixture.

e Rate Theory of Chromatography

more realistic description of the processes at work inside a column takes account of the time en for the solute to equilibrate between the stationary and mobile phase (unlike the plate del, which assumes that equilibration is infinitely fast). The resulting band shape of a omatographic peak is therefore affected by the rate of elution. It is also affected by the erent paths available to solute molecules as they travel between particles of stationary phase. we consider the various mechanisms which contribute to band broadening, we arrive at the n Deemter equation for plate height;

$$\text{HETP} = A + B / u + C u$$

ere u is the average velocity of the mobile phase. A, B, and C are factors which contribute to nd broadening.

Eddy diffusion

e mobile phase moves through the column which is packed with stationary phase. Solute

lecules will take different paths through the stationary phase at random. This will cause adening of the solute band, because different paths are of different lengths.

Longitudinal diffusion

concentration of analyte is less at the edges of the band than at the center. Analyte diffuses from the center to the edges. This causes band broadening. If the velocity of the mobile phase igh then the analyte spends less time on the column, which decreases the effects of itudinal diffusion.

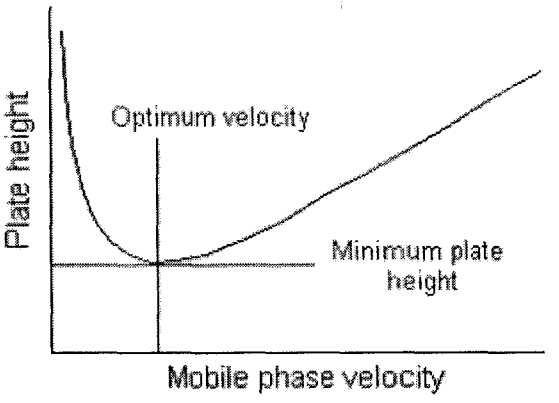
Resistance to mass transfer

analyte takes a certain amount of time to equilibrate between the stationary and mobile se. If the velocity of the mobile phase is high, and the analyte has a strong affinity for the ionary phase, then the analyte in the mobile phase will move ahead of the analyte in the tionary phase. The band of analyte is broadened. The higher the velocity of mobile phase, the rse the broadening becomes.

n Deemter plots

plot of plate height vs. average linear velocity of mobile phase.

A typical Van Deemter plot



uch plots are of considerable use in determining the optimum mobile phase flow rate.

---

esolution

lthough the selectivity factor, a, describes the separation of band centres, it does not take into ccount peak widths. Another measure of how well species have been separated is provided by easurement of the *resolution*. The resolution of two species, A and B, is defined as

$$R = \frac{2[(t_R)_B - (t_R)_A]}{W_A + W_B}$$

aseline resolution is achieved when $R = 1.5$

t is useful to relate the resolution to the number of plates in the column, the selectivity factor nd the retention factors of the two solutes;

$$R = \frac{\sqrt{N}}{4}\left(\frac{\alpha - 1}{\alpha}\right)\left(\frac{1 + k'_B}{k'_B}\right)$$

To obtain high resolution, the three terms must be maximised. An increase in N, the number of theoretical plates, by lengthening the column leads to an increase in retention time and increased band broadening - which may not be desirable. Instead, to increase the number of plates, the height equivalent to a theoretical plate can be reduced by reducing the size of the stationary phase particles.

It is often found that by controlling the capacity factor, k', separations can be greatly improved. This can be achieved by changing the temperature (in Gas Chromatography) or the composition of the mobile phase (in Liquid Chromatography).

The selectivity factor, a, can also be manipulated to improve separations. When a is close to unity, optimising k' and increasing N is not sufficient to give good separation in a reasonable time. In these cases, k' is optimised first, and then a is increased by one of the following procedures:

1. Changing mobile phase composition
2. Changing column temperature
3. Changing composition of stationary phase
4. Using special chemical effects (such as incorporating a species which complexes with one of the solutes into the stationary phase)

Review your learning

You should now be familiar with the terms used in chromatography, how species become separated from one another, and how various conditions can be manipulated to obtain well-resolved chromatograms with a minimum elution time.

1. Gel-Filtration Chromatography

Principle and concept:

Gel filtration chromatography exploits the different sizes of the molecules to effect their separation. The stationary phase contains cross-linked polymeric particles or beads which have pores or holes in them. The sample is normally applied to the top of the column and then eluted with a buffer. Smaller protein molecules find their way into the holes and therefore take longer to elute from the column as they have further to travel. Different polymer gels with different sized pores are used to get optimum separation for a particular size range. If the column is calibrated with standards of known molecular weight a gel filtration column can be used to get rough estimates of molecular weight.

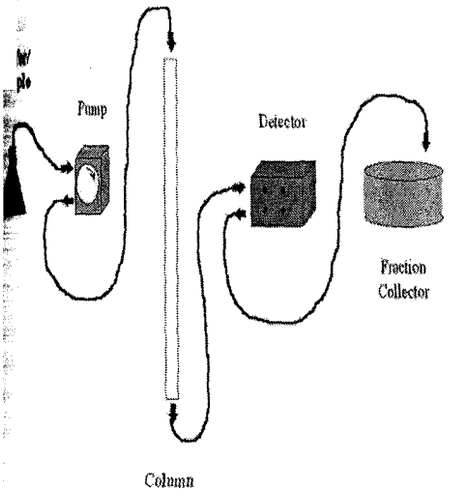
This method separates proteins by size in order of large to small

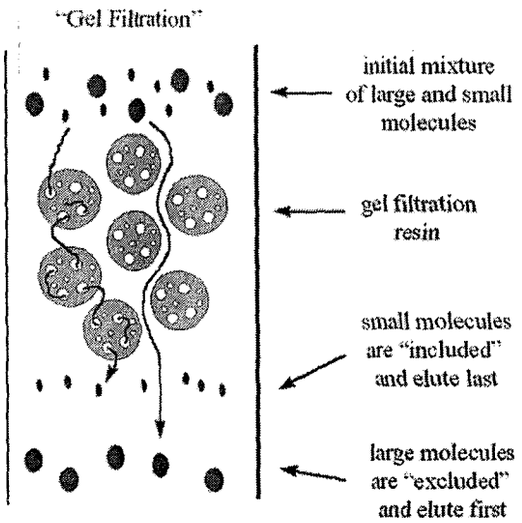
l filtration does not rely on any chemical interaction with the protein, rather it is based on a ysical property of the protein - that being the *effective molecular radius* (which relates to mass most typical globular proteins).

- Gel filtration resin can be thought of as beads which contain pores of a defined size range.
- Large proteins which cannot enter these pores pass around the *outside* of the beads.
- Smaller proteins which can enter the pores of the beads have a longer, tortuous path before they exit the bead.

Thus, a sample of proteins passing through a gel filtration column will *separate based on molecular size*: the big ones will elute first and the smallest ones will elute last (and "middle" sized proteins will elute in the middle).

If your protein is unusually "small" or "large" *in comparison to contaminating proteins* then l filtration may work quite well.





Source: http://matcmadison.edu

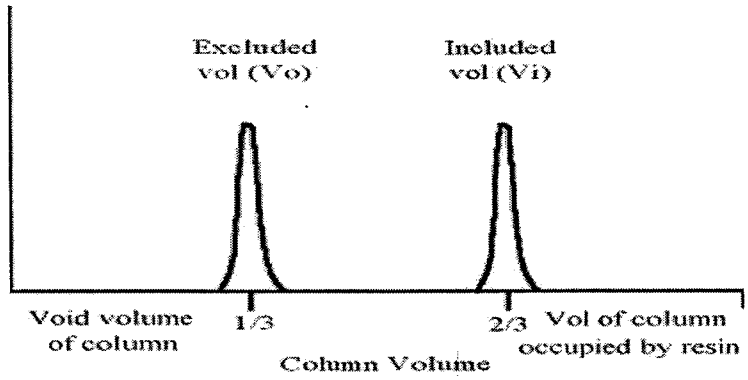
Column Volume, Bed Volume (Vt) = the volume of the gel

Exclusion limit: The MW of the smallest molecule that can not enter the gel.

Excluded volume or void volume (Vo) = All solutes in the sample which are equal to, or larger, than pores of the beads will behave identically: they will all eluted in the excluded volume (Vo) of the column. These solutes will come out of the column first.

Included volume or eluted volume (Vi or Ve) = All small mass solutes in the sample which will be ***completely included*** within the pores of the gel filtration beads. All solutes in the sample which are equal to, or smaller, than this critical size will behave identically: they will

all eluted in the included volume of the column. These solutes will come out of the column later.



Source: http://matcmadison.edu

a general rule of thumb, the excluded volume (Vo) is approximately equal to one third of the umn volume, the included volume is approximately equal to two thirds of the column volume

- In gel filtration the resolution is a function of column length (the longer the better)
- However, one drawback is related to the maximum sample volume which can be loaded. The larger the volume of sample loaded, the more the overlap between separated peaks. Generally speaking, the sample size one can load is limited to about *3-5% of the total column volume.*
- Thus, gel filtration is best saved for *the end stages of a purification* ,when the sample can be readily concentrated to a small volume.
- Gel filtration can also be used to remove salts from the sample, due to its ability to separate "small" from "large" components.
- Finally, gel filtration can be among the most "gentle" purification methods due to the lack of chemical interaction with the resin.

2. Ion-Exchange Chromatography

Principle and Concept:

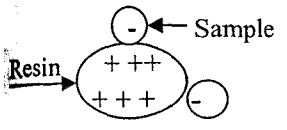
Ion exchange chromatography is the method used to separate proteins by the differences of protein charges in term of both the amount and type of charges (positive or negative). In another word, it separates proteins based on their electrical charge, which depends on their amino acid composition and the pH of the medium. The molecules to be separated bind electrostatically to the stationary matrix. They are eluted from the column using a gradient solution of increasing ionic strength since salts tend to disrupt electrostatic interactions. Alternatively a pH gradient can be used for elution since if pH > pI the average charge is negative and if the pH< pI the average charge is positive.

- Ion exchange resins have fixed charges - either positive or negative.
- Proteins bind to the resin via electrostatic interactions.

The strength of these interactions depends on the net charge on the protein, which is a function of pH and the nature of the weak acid amino acid side chains, and the salt concentration of the buffer - high salt concentrations reduce the interaction. The higher the net charge on the protein at the pH of the environment on the column, the more tightly it sticks to the oppositely charged resin, and the higher the salt concentration required to elute it from the column.

Anion Exchange Chromatography is used to separate negatively charged proteins (anions). The column has positively charged groups on the inert phase which bind negative sites on the protein. The protein may then be exchanged by anions in the solution.

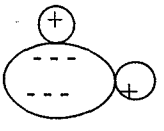
Resins exhibit positive charges and bind to proteins that exhibit negative Charges



Example of resin : DEAE-cellulose contains amino group which exhibit positive charge

Cation Exchange Chromatography is used to separate positively charged proteins (cations). The column contains a polyanionic matrix, usually sulfonate (-SO3-) or carboxyl (-COO-) groups covalently linked to a cellulose or agarose matrix.

Resins exhibit negative charges and bind to proteins that exhibit positive charges



Example of resin : CM-cellulose contains carboxyl group which exhibit negative charge

3. Affinity Chromatography

Principle and concept: Affinity chromatography is a general term which applies to a wide range of chromatographic media. It can be basically thought of as some inert resin to which has been attached some compound which has a *specific affinity* for your protein of interest.

- Thus, a specific antibody attached to an inert resin would be a type of affinity chromatography.
- Other examples might include: a protease inhibitor attached to some matrix, designed to bind a specific protease
- a cofactor bound to some matrix, designed to bind to a particular enzyme
  a metal ion bound to a matrix, designed to chelate a protein with a metal binding site (Metal-Chelate Affinity Chromatography), and so on.

- Affinity chromatography is a more specific interaction in which a ligand specifically recognized by the protein of interest is attached to the column material.
- When a mixture of proteins is passed through the column, only those few that bind strongly to the ligand will stick, while the others will pass through the column.
- By changing the buffer one weakens the interaction between the protein and the ligand, which causes the protein to be eluted from the column.

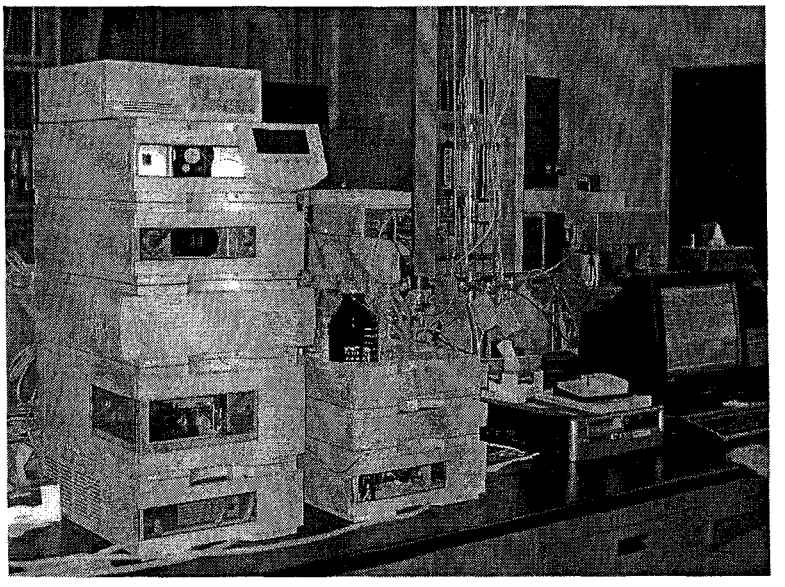
A variation is immunoaffinity chromatography, in which an antibody specific for a protein is immobilized on the column and used to affinity purify the specific protein.

Affinity chromatography is often used to purify cloned & overexpressed proteins. Since the protein has been cloned it is possible to add a specific peptide sequence at the end peptide chain. One possibility is to add a number (usually 6) histidine residues to one end of the protein, which then binds specifically to a column containing immobilised nickel ions. Another possibility is actually to link another complete protein, such a glutathione-S-transferase (GST), to the gene expressing the target protein. GST will bind specifically to beads coated with glutathione. One the GST-"target protein" complex is purified the GST can be removed enzymatically. Affinity columns can also be made using antibodies which bind the target protein. These antibodies can be imobilised on the column.

4. High-Performance Liquid Chromatography

Principle and Concept:

High Performance Liquid Chromatography (HPLC). HPLC uses the same basic chemistry described above, including ion exchange and gel filtration. HPLC uses relatively narrow long columns with non-compressible supports. The mobile phase is then driven over the support using



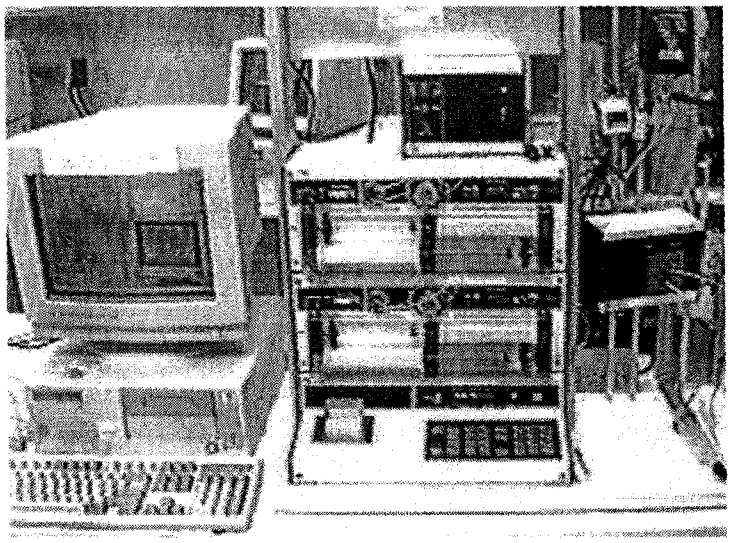
Source: http://steacie.nrc-cnrc.gc.ca

pump at high pressure (5,000-10,000 psi, or 1-20 Mpa). The advantages include:
higher resolution - better separation
speed - less time
high sensitivity - can use small quantities
automation - computer controlled pumps etc.

HPLC is a popular method of analysis because it is easy to learn and use and is not limited by the volatility or stability of the sample compound. The history section illustrates the HPLC's evolution from the 1970's to the 1990's. Modern HPLC has many applications including separation, identification, purification, and quantification of various compounds.

Reverse phase HPLC is often used to separate smaller unfolded proteins or peptides which can be easily refolded following separation. The mobile phase is a non-polar solvent. The stationary phase is a non-polar liquid immobilised on a matrix. The interactions are hydrophobic.

Fast Performance Liquid Chromatography



Source: http://steacie.nrc-cnrc.gc.ca

Liquid chromatography is a term which refers to all chromatographic methods in which the mobile phase is liquid. The stationary phase may be a liquid or a solid. Fast performance liquid chromatography (FPLC) is a type of liquid chromatography where the solvent velocity is controlled by pumps. The pumps controls the constant flow rate of the solvents. The solvents are accessed through tubing from an outside resevoir. The flow rate of the solvent is set through computer input and controlled by pumps.

ere are various columns used in liquid chromatography depending on the type of separation ferred. Each column contains a small diameter packing material. The column is a large (mm tube containing small (u) particles (gel beads) known as stationary phase. The omatographic bed is composed by the gel beads alone when they are inside the column. The ple is introduced into the injector and then carried into the column by the flowing solvent. ce in the column, the sample mixture separates as a result of different components adhering to diffusing intothe gel.As the solvents is forced into the chromatographic bed by the flow rate, sample separates into various zones of sample components. These zones are referred to as nds.

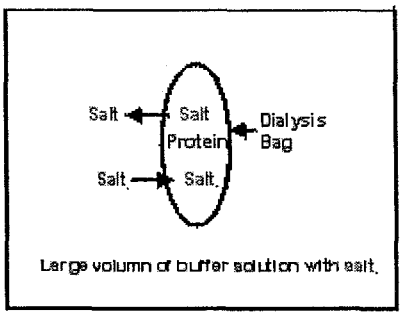
5. Dialysis/Ultrafiltration

nciple and concept:

alysis is a technique that separates molecules according to their size. Smaller molecules are owed to pass through a semi-permeable membrane containing pores too small to allow the olecules being purified to get through. Dialysis may be carried out in hollow tubes or specially signed flat cells. Membranes are usually made from cellulose acetate (cellophane) but rocellulose is sometimes used.

alysis is often used to remove low molecular weight (MW) contaminants from a protein lution or to change the buffer in which the protein is stored. Over a period of 3-4 hours, the side and outside solutions come to equilibrium. As long as there is a vast excess of the external lution (say 1000:1), the final concentration in the dialysis bag will resemble the external edium.

alysis can also be used to separate different-sized proteins as dialysis tubing can be obtained th specific MW cut-offs (the MW cut-off refers to the largest-sized protein, which will be tained by the dialysis membrane). For example, one could remove a contaminating 5000 Da otein from a 50,000 Da protein by dialysing the mixture using dialysis tubing with a MW cut-ff of 10,000. The small protein will be lost to the external medium while the larger protein will e retained in the dialysis bag.



Protein concentration determination (Bradford Method)

nciple and concept:

e Bradford protein assay is one of several simple methods commonly used to determine the al protein concentration of a sample. The method is based on the proportional binding of the e Coomassie to proteins. Within the linear range of the assay (~5-25 mcg/mL), the more tein present, the more Coomassie binds. Furthermore, the assay is colorimetric; as the protein ncentration increases, the color of the test sample becomes darker. Coomassie absorbs at 595 . The protein concentration of a test sample is determined by comparison to that of a series of tein standards known to reproducibly exhibit a linear absorbance profile in this assay. hough different protein standards can be used, we have chosen the most widely used protein our standard - Bovine Serum Albumin (BSA).

breviations:

- mcg = micrograms
- dI = deionized
- mcL = microliters
- BSA = bovine serum albumin
- O.D. = optical density

ocedure

- Prepare a 4-fold dilution of a 2 mg/mL BSA sample by adding 50 mcL of 2 mg/mL BSA to 150 mcL of dI water to make 200 mcL of 0.5 mg/mL BSA.
- Generate test samples for the reference cell, blank, BSA standards and the protein sample to be tested according to Table 1 in disposable cuvettes.
- Note that the "reference cell" and "blank" are identical. A reference cell test sample is only required when using a double-beam UV-visible spectrophotometer for absorbance measurements.
- Note that a dilution of the protein sample may be required for the resulting absorbance to fall within the linear range of the assay.
- Allow each sample to incubate at room temperature for 10-30 minutes. (Record the actual incubation time in your notebook.)
- Measure the absorbance of each sample at 595 nm using a UV-visible spectrophotometer. Be sure to allow the instrument to warm up for at least 15 minutes prior to use.
- Plot the absorbance of each BSA standard as a function of its theoretical concentration. The plot should be linear.
- Determine the best fit of the data to a straight line in the form of the equation "y = mx + b" where y = absorbance at 595 nm and x = protein concentration.
- Use this equation to calculate the concentration of the protein sample based on the measured absorbance. Note: If the absorbance of the test sample is outside of the absorbance range for the standards, then the assay must be repeated with a more appropriate dilution, if any. The linear range for the assay (and for most spectrophotometers is 0.2 - 0.8 O.D. units.

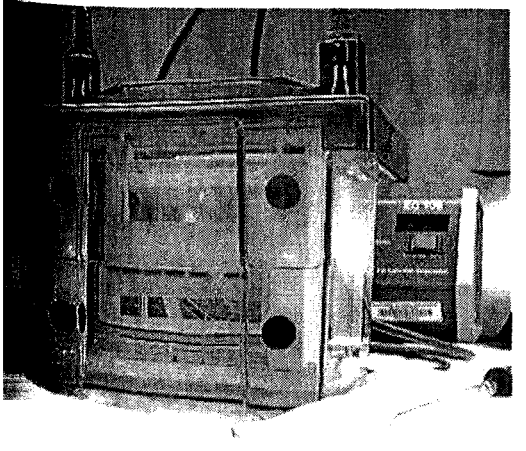
ble 1. Preparation of test samples for the Bradford protein assay.

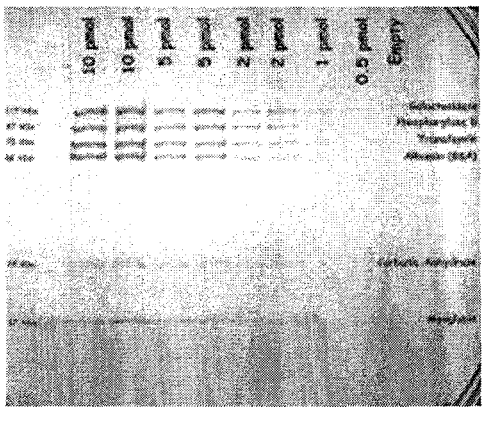
Test Sample	Sample Volume, mcL	Vol. Water, mcL	Vol. Bradford Reagent mcL
Reference Cell	0	800	200
Blank	0	800	200
BSA Standard - 5 mcg/mL	10	790	200
BSA Standard - 10 mcg/mL	20	780	200
BSA Standard - 15 mcg/mL	30	770	200
BSA Standard - 20 mcg/mL	40	760	200
BSA Standard - 25 mcg/mL	50	750	200
Protein Sample	50	750	200

Topic 17. Protein Electrophoresis Techniques

Electrophoresis - provides a very high resolution technique for separating macromolecules. As the name implies the separation relies on the different rates of movement of molecules in an electric fields. Thus it ultimately relies on the charge on the molecules, much the same as ion exchange chromatography. Since a limited amount of material can be applied to a gel, electrophoresis is normally used for analytical rather than preparative purposes.
Most electrophoresis experiments in Biochemistry are run in gels made from polyacrylamide or from agarose. The sample to be separated is placed at one end of the gel and an electric voltage is applied. The separation will depend on the net charge on the molecules as well as their effective sizes since the gel matrix acts like a gel filtration column. The bands are made visible by staining with dyes. Proteins are typically stained with Coomassie brilliant blue or if more sensitivity is required they are silver stained.

SDS-PAGE gels. If separation is to be based on molecular weights then the charge differences need to be minimised so that the separation is based solely on the gel filtration effect. This is generally done by first boiling the proteins and treating them with b-mercaptoethanol to denature the molecules. They are then treated with SDS (sodium dodecyl sulfate) a detergent which binds to the protein, keeps it unfolded and gives the molecules an approximately uniform charge per molecular weight. These extended peptide then move through the gel, the smallest moving fastest. This is the opposite to gel filtration since in this case the matrix is a cross-linked polymer and the shorter peptides move through the pores more quickly.
Native PAGE gels are run when it is desirable to see the extent of oligomerisation of the native protein. The pH needs to be chosen in such a way that the protein of interest carries the correct charge to migrate through the gel. The smaller the oligomer, the faster it moves through the gel, since oligomers of the same protein will have approximately the same charge per mass ratio and they will therefore be sorted by size.

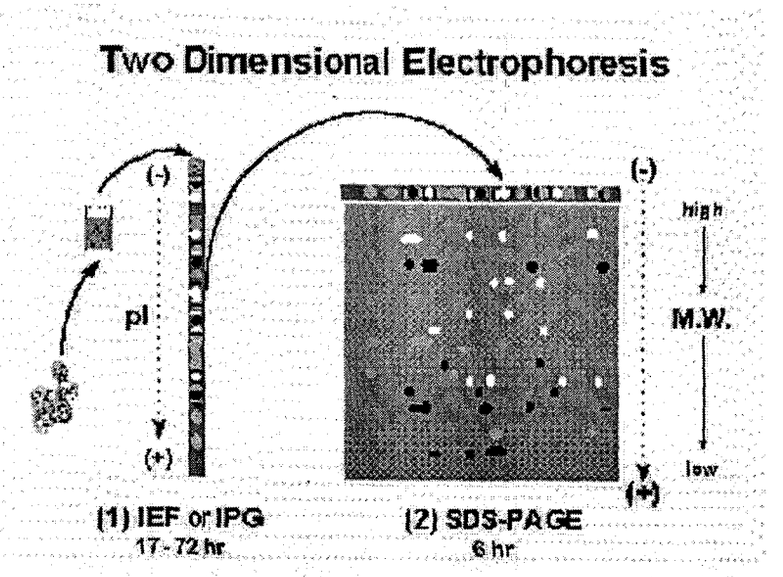




Source: Promega product catalog

lectric focussing (IEF)- is a variation on the general gel electrophoresis method. A very low ecular weight acrylamide is used which forms large pores. There is a pH at which any romolecule carries no net charge. This pH is termed the isoelectric point or pI. Imagine that ave an electrophoresis experiment in which there is a stable and continuous pH gradient ing from the cathode to the anode. In this case each molecule will migrate to a position on el where the pH is equal to its pI. At this point the molecule will feel no net force and it will moving.

-dimensional gels (2D) where the separation in one dimension is based on gel trophoresis and the other on isoelectric focussing afford the opportunity to separate on two erties, size and pI at the same time. That is separation will be based on a combination of ge, size and shape in one direction and pI in the other. The proteins are usually first rated by their pI by isoelectric focussing and then by size using SDS PAGE. One can rate all the proteins in bacterium using this method on one large gel.



Source: http://www.lecb.ncifcrf.gov/phosphoDB/2d-description.html

Theory of SDS-PAGE

.Polyacrylamide Gel Electrophoresis (SDS-PAGE) is probably the worlds most widely used hemical method. In the early 60's scientists first appreciated the utility of polyacrylamide as a convenient and versatile alternative to starch gels Ornstein 1964, Davis 1964, thus eloping polyacrylamide gel electrophoresis or PAGE. The inclusion of ionic detergent ium Dodecyl Sulphate (SDS) to the gel and the sample was an important addition to this k. Shapiro et al. were one of the first to make use of this approach (Shapiro et al. 1967). mmli showed that proteins could be reliably fractionated by SDS-PAGE, which he described figure legend in a Nature paper (Laemmli, 1970). This paper used the method to study ifically the proteins of bacteriophage T4, and made use of Tris based buffers, as used by t researchers today. At about the same time Weber and Osborn showed that the method had eral applicability and also showed, in a very systematic study, that the relative mobility of eins in SDS-PAGE gels correlated quite well with their molecular weights (Weber and orn 1969). However these gels were based on phosphate buffers, which subsequently have n less widely used. After these papers appeared SDS-PAGE gels became wildly popular, first ubes, then as slabs and then as minigels. Ultimately 2-dimensional gel techniques, using lectric focusing of proteins in one direction followed by regular SDS-PAGE in the other ction were developed. The first influential paper in this area was published by O'Farrell Farrell 1975). People also began to appreciate that it was possible to transfer proteins out of S-PAGE gels onto agarose, or much more conveniently, onto nitrocellulose membranes and re stain them with antibodies. The most influential early paper making use of nitrocellulose mbranes was that of Towbin et al. (Towbin et al 1979). Later studies used other kinds of mbranes, notably the nylon like material PVDF, which allowed proteins transferred from S-PAGE gels to be probed with antibodies but also subjected to direct peptide sequencing atsudaira, 1975).

e separation of macromolecules in an electric field is called *electrophoresis*. A very common thod for separating proteins by electrophoresis uses a discontinuous polyacrylamide gel as a pport medium and sodium dodecyl sulfate (SDS) to denature the proteins. The method is led sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The most mmonly used system is also called the Laemmli method after U.K. Laemmli, who was the first publish a paper employing SDS-PAGE in a scientific study.

DS (also called lauryl sulfate) is an anionic detergent, meaning that when dissolved its olecules have a net negative charge within a wide pH range. A polypeptide chain binds ounts of SDS in proportion to its relative molecuar mass. The negative charges on SDS troy most of the complex structure of proteins, and are strongly attracted toward an anode positively-charged electrode) in an electric field.

olyacrylamide gels restrain larger molecules from migrating as fast as smaller molecules. ecause the charge-to-mass ratio is nearly the same among SDS-denatured polypeptides, the al separation of proteins is dependent almost entirely on the differences in relative molecular ass of polypeptides. In a gel of uniform density the relative migration distance of a protein (Rf, e f as a subscript) is negatively proportional to the log of its mass. If proteins of known mass e run simultaneously with the unknowns, the relationship between Rf and mass can be plotted, d the masses of unknown proteins estimated.

tein separation by SDS-PAGE can be used to estimate relative molecular mass, to determine relative abundance of major proteins in a sample, and to determine the distribution of teins among fractions. The purity of protein samples can be assessed and the progress of a ctionation or purification procedure can be followed. Different staining methods can be used detect rare proteins and to learn something about their biochemical properties. Specialized hniques such as Western blotting, two-dimensional electrophoresis, and peptide mapping can used to detect extremely scarce gene products, to find similarities among them, and to detect l separate isoenzymes of proteins.

lecular mass versus molecular weight

lecular mass (symbol m) is expressed in Daltons (Da). One Dalton is defined as 1/12 the mass carbon 12. Most macromolecules are large enough to use the kiloDalton (kDa) to describe lecular mass. Molecular weight is not the same as molecular mass. It is also known as relative lecular mass (symbol Mr, where r is a subscript). Molecular weight is defined as the ratio of mass of a macromolecule to 1/12 the mass of a carbon 12 atom. It is a dimensionless ntity.

hen the literature gives a mass in Da or kDa it refers to molecular mass. It is incorrect to press molecular weight (relative molecular mass) in Daltons. Nevertheless you will find the n molecular weight used with Daltons or kiloDaltons in some literature, often using the breviation MW for molecular weight.

lyacrylamide gels for SDS-PAGE

ny systems for protein electrophoresis have been developed, and apparatus used for SDS-GE varies widely. The methodology used on these pages employs the Laemmli method. ference to the Laemmli method in a materials and methods section eliminates the need to scribe the buffers, casting of gels, apparatus, etc. Unless the paper employs some modification the method, the only details of SDS-PAGE that should be reported in a methods section are rcent total acrylamide (%T) in a gel, relative percentage and type of crosslinker (%C), and haps a reference to the gel dimensions. We use a "mini-gel" system, with 3 1/4" x 4" gel ssettes.

S-PAGE can be conducted on pre-cast gels, saving the trouble and hazard of working with ylamide. The following description applies to shop-made casting and running apparatus that much cheaper than commercially available equipment. In addition to cost effectiveness, an vantage of making one's own gels the first time is a deeper understanding of the process.

gardless of the system, preparation requires casting two different layers of acrylamide tween glass plates. The lower layer (separating, or resolving, gel) is responsible for actually arating polypeptides by size. The upper layer (stacking gel) includes the sample wells. It is signed to sweep up proteins in a sample between two moving boundaries so that they are mpressed (stacked) into micrometer thin layers when they reach the separating gel.

S-PAGE Procedure

st Protoaol

x for about 20mls running gel solution, enough for two minigels, total is 20.01 mls. te "ul" means microliter, many browsers have trouble displaying Greek characters rrectly.

	12%	10%	9%	8%	7.5%	6%
rylamide/Bisacryclamide Mix	10.8 mls	9 mls	8.1 mls	7.21 mls	6.76 mls	5.4 m
M Tris/HCl pH=8.8	7.5 mls	7.5 mls	7.5 mls	7.5 mls	7.5 mls	7.5 m
stilled water	1.4 mls	3.2 mls	4.1 mls	4.99 mls	5.44 mls	6.8 m
% SDS	200 uls	200 uls	200 uls	200 uls	200 uls	200 u
% Ammonium Persulfate	100 uls	100 uls	100 uls	100 uls	100 uls	100 u
MED	10 uls	10 uls	10 uls	10 uls	10 uls	10 ul

acking Gel Solution, good for 2 minigels, ~4.8% acrylamide:

rylamide/Bisacrylamide Mix	2 mls
ater	5.6 mls
M Tris/HCl pH=6.8	1.25 mls
% SDS	100 uls
% Ammonium Persulfate	50 uls
MED	5 ulsTotal 10.05 mls

crylamide/Bisacrylamide = 22.2 g acrylamide, 0.6 g bis-acrylamide (37:1 cross-linker tio) to 100 ml water, filtered.

eservoir/Running buffer = 57.6 g Glycine, 12 g Tris base, 4 g SDS, water to 4 litres

tain solution = 2.5 g Coomassie Brilliant Blue R-250, 450 mls methanol, 100 mls glacial cetic acid, water to 1 liter.

estain solution = 300 mls methanol, 400 mls acetic acid, water to 4 liters.

ample buffer 5X = make up 100 mls and store away 5-10 mls aliquots.

M Tris/HCl pH = 6.8	31.25 mls
DS powder	10 g
lycerol	25 mls
romophenol Blue (2% in ethanol)	750 uls
-mercaptoethanol	5 uls
ater	to 100 mls

ond Protocol

terials and Methods

ipment:

-Rad Mini-Trans-Blot Cell
ctrophoretic Blotting Apparatus
aker (orbital type)

emicals:

urce Cat. No.
Blott ABI 400994 TM
rylamide Bio-Rad 161-0100
ido Black Bio-Rad 161-0402
monium persulfate Bio-Rad 161-0700
mophenol blue Bio-Rad 161-0404
omassie©Brillant Blue R-250 Bio-Rad 161-0400
yclohexylamino-1-propane-
fonic acid (CAPS) Aldrich 16,376-7
ycerol BRL 5514UA
ycine Bio-Rad 161-0717
ercaptoethanol Sigma M-6250
thanol ABI 400470
N'-methylene bisacrylamide Bio-Rad 161-0200
N,N,N-Tetramethylene-
mine(TEMED) Bio-Rad 161-0800
nceau Sigma P7767
rva Blue G Serva 35050
dium dodecyl sulfate (SDS) Bio-Rad 161-0301
dium thioglycolate Sigma T-0632
cine Aldrich 16,378-3
s (hydroxymethyl)aminomethane Bio-Rad 161-0716
iton X-100 (NP-40) Bio-Rad 161-0407
ea Bio-Rad 161-0730

DS-PAGE Systems

number of systems used with gel electrophoresis have been described . These are the Tris-
lycine [4 5]
d Tris-Tricine buffer systems for mini-gles. Tris-glycine is used for separating proteins larger
an [6]
kDa; the Tris-Tricine buffer system is used to resolve peptides and low-molecular weight
oteins
well as larger proteins. Either buffer system can be used prior to electroblotting.

ris-Glycine stock solutions and buffers

Lower tris (4 X) Dissolve 36.34 g tris base and 0.8 g SDS in 150 mL D.I. H2O.
itrate with 6N HCl to pH 8.8. Add D.I. H O to final volume of 200 [2]
L.
Upper tris (4 X) Dissolve 12.11 g tris base and 0.8 g SDS in 150 mL D.I. H2O.
itrate with 6N HCl to pH 6.8. Add D.I. H O to final volume of 200 [2]

mL.
30% Acrylamide Dissolve 30g acrylamide and 0.8 g bisacrylamide in D.I. H O to final 2
30.8% T, 2.6% C) volume of 100 mL.
Electrophoresis Buffer Dissolve 3.03 g tris base, 14.4 g glycine and 1 g SDS in 1 L D.I. H O. 2
Ammonium persulfate Dissolve 100 mg ammonium persulfate in 1 mL D.I. H O. Store the 2
10%) solution at 5 C. Discard after one week. o
Sample Preparation (2X) Mix 1 mL glycerol, o.5 mL 2-mercaptoethanol, 0.3 g SDS, 0.25 mL
olution of 0.05% (w/v) bromophenol blue and 1.25 mL Upper Tris (4X)
uffer in D.I. H O to a final volume of 5 mL. 2
Filter solutions with 0.2 μ Nalgene© filter.

Tris-Glycine Gel Solutions

The following proportions are for preparing 2 gels (0.5 x 60 x 90 mm) for mini-gel systems.

Stock Solutions	**10% Resolving Gel (8 mL)**	**15% Resolving Gel (8 mL)**	**4% Stacking Gel (4 mL)**
30% Acrylamide	2.66 mL	4.0 mL	0.52 mL
Upper tris (4X)	-	-	-
Lower tris (4X)	2.0 mL	2.0 mL	-
D.I. H2O	3.34 mL	2.0 mL	2.47 mL
De-gas for 5 minutes			
TEMED	4 μL	4 μL	4 μL
Ammonium persulfate (10%)	40 μL	40 μL	40 μL

Electrophoresis Conditions for Tris-Glycine

Sample preparation: Mix one volume (1-10 μL) of sample solution with one volume of ample preparation solution (2X) and heat in boiling water bath for 2 minutes. The sample now ready to load into the sample well for electrophoresis.

When using the mini-gel apparatus, the recommended electrophoresis condition for optimal esolution with minimal thermal band distortion is 7 mA constant current (50-90 volts) for pproximately 1 hour. For maximum separation, allow the bromophenol blue dye to run to ithin 1 cm of the end of the gel.

Tris-Tricine stock solutions and buffers

Anode Buffer (10 X) Dissolve 121.9 g tris base in 400 mL D.I. H O. 2
2.0 M Tris HCl) Titrate to pH 8.9 with HCl. Add D.I. H O to final volume 500 2
mL.
Cathode Buffer (10 X) Dissolve 60.55 g, 89.58 g tricine and 5 g SDS in 400 mL D.I.
1.0 M Tris, 1.0 M Tricine, H O. Titrate to pH 8.25 with HCl. Add D.I. H O to final 1.0% 2 2
DS) volume 500 mL.
Acrylamide Solution Dissolve 48 g acrylamide and 1.5 g bis-acrylamide in 100 mL
49.5% T, 3% C) D.I. H O. 2
Gel Buffer (3 X) Dissolve 181.5 g tris base, 1.5 g SDS in 400 mL D.I. H O. 2
Titrate to pH 8.45 with HCl. Add D.I. H O to final volume 2
f 500 mL.
0% glycerol (v/v) Dissolve 250 mL glycerol in 250 mL D.I. H O. 2
Sample Preparation Mix 1.0 mL of 100% glycerol, 0.5 mL 2-mercaptoethanol, 2.5
olution (2X) mL of 20% (w/v) SDS, 0.62 mL Tris (pH 6.8, 1M), 0.01 mL
f 10% Serva BlueG (w/v) in D.I. H O to a final volume of 2

mL.
Filter solutions with 0.2 µ Nalgene© filter.

Tris-Tricine Gel Solutions

The following proportions are preparing 2 gels (0.5 x 60 x 90 mm).

Stock Solutions 10% Resolving 4% Stacking

Gel (15 mL) Gel (12.5 mL)

Acrylamide solution 3.0 mL 1.0 mL

50% glycerol 4.0 mL -

Gel Buffer (3X) 5.0 mL 3.1 mL

H_2O 3.0 mL 8.4 mL

De-gas for 5 minutes

TEMED 7.5 µL 7.5 µL

Ammonium persulfate (10%) 75 µL 100 µL

Electrophoresis Conditions for Tris-Tricine

Sample preparation: Mix 1 volume of sample with 1 volume of sample preparation solution (2 X) and heat in boiling water bath for 2 minutes.

Electrophoresis Conditions: 20 mA constant current for 2-3 hours. For maximum separation, allow the serva blue dye to run to within 1 cm of the end of the gel.

Electroblotting

CAPS Buffers[1]

Stock CAPS (10 X) Dissolve 22.13 g CAPS in 900 mL of D.I. water. Titrate with 2N NaOH (20 ml) to pH 11, and add D.I. water to a final volume of 1L.

Store at 4EC.

Electroblotting buffer Prepare 2L by mixing 200 mL of the 10X CAPS buffer with 200 ml 10 mM CAPS in of methanol and 1600 mL of D.I. water.

10% MeOH)

Tris-Glycine Buffer7,8

Electroblotting buffer Dissolve 3.0 g Tris and 14.4 g glycine in a solution of 900 mL H_2O [25 mM Tris base, and 100 mL methanol.

192 mM glycine, 10%

(v/v) methanol (pH 8.3)]

Procedure:

1. Remove the gel from the electrophoresis cell and soak it in 100 mL of electroblotting buffer for 5 minutes.
2. Meanwhile, wet ProBlott™ with 100% methanol for a few seconds, then transfer it to the blotting buffer.

Note: Be sure the ProBlott™ membrane is completely moistened with methanol before transferring it to the blotting buffer. It if dries out, wet it with methanol again.

3. Dip the sponges and filter papers in a separate container of blotting buffer before starting to assemble the transblotting sandwich.
4. Assemble the transblot sandwich in the following order starting from anode side: sponge, filter paper, 2 sheets of ProBlott™, gel, filter paper, sponge.

Make sure there are no bubbles between the gel and the membrane.

5. Pour 1L of electroblotting buffer into the transblot cell and insert the transblot sandwich. Electroblotting conditions are dependent of the buffer, as follows:

APS:50V (170 mA-100mA) at room temperature for 30 minutes. Tris-Gly:40V (300 mA) room temperature for 1-4 hours.

Note: Electroblotting times vary and are dependent on protein sample. Longer electroblotting times may be necessary with proteins that are 70 kDa or larger.

. After transfer is complete, remove ProBlott from transblotting sandwich and rinse with D.I. ™ water before staining.

Protein Detection

rotein sampleson ProBlott can be detected with conventional staining techniques, such as ™ oomassie© Brilliant Blue, Ponceau S or Amido Black.

oomassie© Brilliant Blue Staining Solution

.1% Coomassie Blue R-250 in 1% acetic acid/40% MeOH

issolve 1.0 g Coomassie blue in 400mL of methanol. Stir for one hour. Add 10 mL acetic acid nd

90 mL D.I. water. Stir for 30 min, then filter with a Nalgene© filter (0.45µ pore size).

estaining solution

0% MeOH in D.I. water

oomasie Blue Staining Procedure:

. Remove the ProBlott membranes (both primary and secondary) from the transblot ™ andwich, and rinse them with D.I. water.

. Saturate the ProBlott membranes with 100% MeOH for a few seconds. ™

. Stain the ProBlott membranes (one at a time) with the staining solution using constant ™ rbital shaking. Protein bands should appear within one minute.

CAUTION: If the ProBlott membrane is over-stained, it may require a ™ longer destaining time to prevent a high background.

. Destain the ProBlott membranes by soaking them in destaining solution. For quicker ™ destaining and improved contrast, change the destaining solution several times, allowing the membranes to soak for 1-2 minutes between changes.

. Rinse the membrane thoroughly with D.I. water, and excercise the bands of interest with a clean razor blade.

Amido Black Stainig solution

.1% Amido Black in 1% acetic acid/40% MeOH

Dissolve 1.0 g Amido Black in 400 mL of methanol. Stir this solution for at least one hour. Add 10

mL acetic acid and 590 mL of D.I. water. Stir for another 30 minutes, then filter the solution with Nalgene filter (45µ pore size).

The staining procedure for Amido Black is the same as that described for Coomasie Blue, with the substitution of D.I. water as the destaining solution.

Ponceau S Staining solution

0.2% Ponceau S in 1% acetic acid

Dissolve 0.4 g of Ponceau S in 198 mL of D.I. water and stir for 30 minutes. Add 2 mL of acetic acid to the mixture.

1. Remove the ProBlott membranes from the transblotting sandwich and rinse them with D.I. ™ water.

2. Stain the ProBlott membranes in the Ponceau S staining solution with constant orbital ™ shaking. Protein bands should appear within one minute.

3. Destain with a D.I. water rinse. 4. Excise the bands of interest with a clean razor blade.

Denaturing Discontinuous Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. Assemble each gel sandwich by stacking, in order, the notched aluminum plate, two 0.75-mm spacers, and a glass plate. It is important that the spacers are aligned properly, with the ends flush with the top and bottom edges of the plates, and the flanged edge of the T-shaped spacer positioned against the outside edges of the glass plate. Place a piece of wax paper between each sandwich to facilitate separation after polymerization (optional).

2. Fit the gel sandwiches (usually four at a time) tightly in the multiple gel caster. Fill any remaining space in the mold by including additional glass plates.

3. Place the front face plate on the caster, clamp it in place against the silicone gasket, and verify the alignment of the glass plates and spacers.

4. Prepare the separating gel solution as indicated by the table below. Do not add TEMED or ammonium persulfate until ready to pour.

5. Fill a 50-ml syringe with the solution and slowly inject it into the caster until the gels are 6 cm high, allowing 1.5 cm for the stacking gel.

6. Overlay each gel with 100 ul H2O-saturated isobutyl alcohol or 0.1% SDS. Allow gels to polymerize for approximately one hour.

7. Remove any remaining gel overlay solution by blotting the top of each gel with a piece of Whatman 3mm paper.

8. Prepare the stacking gel solution as described below. Fill a 10-ml syringe with the solution, and inject into each gel sandwich until the top is reached.

9. Carefully place an appropriate comb into each gel, taking care not to trap any bubbles. Allow gel to polymerize for about 1 hour.

10. Remove the front faceplate of the caster. Carefully remove the gels, and separate using a razor blade.

    - *The gels can be stored with the combs in place tightly wrapped in plastic wrap inside a sealable bag at 4°C for 2 to 3 weeks. Keep the gels moist. Do not store gels in the multiple caster.*

11. When ready to run the gel, remove the comb and rinse the sample wells with running buffer (1x SDS/electrophoresis buffer, see below). Place a line indicating the bottom of the each well on the front glass plate, or use a well template.

12. Fill the upper and lower buffer chambers with running buffer. The upper chamber should be filled to 1 to 2 cm above the notched plate.

13. Using the marks on the glass plate or the well template as a guide, pipette the prepared samples into the wells.

14. Electrophorese the samples at 10 mA constant current per 0.75-mm gel until the dye front reaches the top of the separating gel (45 minutes). Increase current to 20 mA per gel, and continue run until the bottom of the gel is reached (1 hour).

Recipes used in SDS-PAGE

SEPARATING GEL *a*

	Final acrylamide concentration in the separating gel (%) *b*									
Stock Solutions	5	6	7	7.5	8	9	10	12.5	13	15
30% acrylamide/ 0.8% bisacrylamide	5.00	6.00	7.00	7.50	8.00	9.00	10.00	12.50	13.00	15.00
4x Tris-Cl/SDS, pH 8.8	7.50	7.50	7.50	7.50	7.50	7.50	7.50	7.50	7.50	7.50
dH2O	17.50	16.50	15.50	15.00	14.50	13.50	12.50	10.00	9.50	7.50
10% ammonium persulfate *c*	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
TEMED	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

a. These recipes produce 30 ml of separating gel, adequate for casting four minigels of dimensions 0.75 mm x 7.3 cm x 8.3 cm.

b. Units of all numbers in the table are milliliters.

c. Best to prepare fresh.

Preparation of the separating gel

In a 50-ml conical tube, mix 30% acrylamide/0.8% bisacrylamide solution, 4x Tris-Cl/SDS, pH 8.8, and dH2O. Degas the solution under vacuum 10 to 15 minutes. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

STACKING GEL (4.0% T 2.6% C)

In a 50-ml conical tube, mix 1.30 ml of 30% acrylamide/0.8% bisacrylamide, 2.50 ml of 4x Tris-Cl/SDS, pH 6.8, and 6.10 ml dH2O. Degas under vacuum 10 to 15 minutes. Add 50 µl of 10% ammonium persulfate and 10 µl TEMED. Swirl gently to mix. Use immediately. Produces 10 ml of stacking gel, sufficient for four minigels. The stacking gel is prepared in this way regardless of the acrylamide concentration used in the separating gel.

REAGENTS USED IN GELS

30% acrylamide/0.8% bisacrylamide (30.8% T 2.6% C)

Mix 60.0 g acrylamide and 1.6 g *N,N'*-methylene-bisacrylamide in a total volume of 200 ml dH2O. Filter solution through a 0.45 µm filter and store at 4°C protected from light. Discard after 30 days.

- **CAUTION:** *Acrylamide and bisacrylamide are potent neurotoxins and are absorbed through the skin. Wear a mask while weighing the powder. Gloves and a lab coat should be worn when handling the solution. Do not mouth pipette.*

4x Tris-Cl/SDS, pH 6.8 (0.5 M Tris-Cl/0.4% SDS)

Dissolve 12.1 g Tris base in 80 ml dH2O. Adjust pH to 6.8 with 1 N HCl. Add dH2O to 200 ml total volume. Filter the solution through a 0.45 µm filter, add 0.8 g electrophoresis-grade SDS, and store at 4°C.

Tris-Cl/SDS, pH 8.8 (1.5 M Tris-Cl/0.4% SDS)

Dissolve 182 g Tris base in 600 ml dH2O. Adjust pH to 8.8 with 1 N HCl. Add dH2O to 1000 ml total volume. Filter the solution through a 0.45 µm filter, add 4 g electrophoresis-grade SDS, and store at 4°C.

SDS/electrophoresis buffer (0.125 M Tris, 0.960 M glycine, 0.5% SDS)

Dissolve 15.1 g Tris base, 72.0 g glycine, and 5.0 g electrophoresis-grade SDS in a total volume of 1000 ml dH2O. Do not adjust the pH of the solution (should be pH 8.3 when diluted). Store at 4°C. Dilute to 1x for working solution.

Rapid silver staining protocol for SDS PAGE gels

Protocol 1

1. Fixation (30% Methanol and 5% Acetic acid): Add 50ml of fixing solution for 10 min
2. Water washes: Several changes of water for at least half an hour
3. Hypo (10mg/100ml Sodium thiosulfate): Add Hypo for one min and quickly wash the gel with water
4. Silver nitrate (0.1 gm with 100 µl Formaldehyde): Add 50 ml of the silver nitrate for 8-10 min
5. Developer (1.5% with 50 µl Formaldehyde): Quickly wash the gel with developer to remove the excess stain and gently rinse the gel in Developer until bands appear
6. Stop solution (2M citric acid): Add 5 ml of stop solution

Protocol 2

1. Rinse gel on glass plate briefly with distilled water.

2. Wash gel in fresh 50% ETOH-12% HAc, with shaking, 3 times, 1 hour each. (Fix IEF gel ON in 10% TCA).

3. Wash gel four times with 10% ETOH-5% HAc for one hour each. One wash may be done overnight.

4. Rinse gel with several changes of water over five minutes.

5. Incubate gel in fresh 200 ml of 0.6 g K2Cr04 (0.0034 M) and 40 µl HNO3 (0.0032 M) for 5 minutes.

6. Rinse gel with several changes of water until color fades, approximately 10 min.

7. Incubate in 150 ml of 0.012 M AgNO3 (0.3 g/150 ml) Illuminated for 30 min, rocking several times to incorporate stain.

8. Rinse gel with three changes of water over 1 minute.

9. Rapidly rinse three times with 100 ml 0.28 M Na2CO3 and 0.5 ml 37% formaldehyde (8.9 g Na2CO3 and 1.5 ml formaldehyde/300 ml). Remove the darkened rinse solution quickly each time.

10. Agitate in 100 ml solution in 9 until desired intensity is reached and then stop development with 100 ml of 3% HAc.

11. Wash twice with water and store in water or in 10%MeOH-3% glycerol. Stain will not fade and can be photographed at your convenience.

Color Silver Stain

1. Fix gel in 50% ETOH and 10% HAc for 2 hours or more.

2. Wash gel in 50% ETOH and 10% HAc for 2 hours, then in 25% ETOH-10% HAc for 1 hour two times, and then in 10% ETOH-5% HAc for 1 hour twice.

3. Equilibrate gel in 0.012M AgNO3 solution (0.4 g/200 ml) for 2 hours.

4. Rinse with water for 10-20 seconds.

5. Reduce in the following solution for 10 min.: 200 ml 0.75N NaOH (6 g/200 ml) and dissolve 17.5 mg NaBH4 and 1.5 ml 37% Formaldehyde.

6. Wash gel in 0.75% Na2CO3 for 1 hour and store normally. Color will not fade and can be photographed at some time in the future.

Detailed Procedure for RNase Activity Staining on SDS-PAGE

1. Remove glass plates that have been stored soaking in 3% hydrogen peroxide and dry thoroughly with Kimwipes.
2. Assemble plates and seal with melted 1.5% agarose.
3. Prepare separating and stacking gels in the order listed in the formula:
   -- After pouring separating gel, slowly drop Tris-saturated butanol (use 1M Tris pH 9.0 for this) to form layer on the top
   of the gel.
   -- When the gel is polymerized, wash off the butanol-saturated water with sterilized water and dry water and residual
   polyacrylamide thoroughly with filter paper.
   -- After pouring stacking gel, enter comb on its top on angle to avoid air bubbles.
   -- When set, remove comb gently and rinse wells thoroughly with running buffer.
4. Attach plates to running tank, add running buffer to top and bottom tanks, and remove air bubbles from beneath the gel.
5. Load samples, connect power supply and run gel until the front ion line reaches the bottom.
6. Remove plates from the tank, separate the gel from the plates, and remove the stacking portion.
7. Wash the gel in 25% iso-propanol in 0.01 M Tris pH 7.0 for 2 x 10 min (to remove SDS from the gel).
8. Wash the gel in 2 μM ZnCl2/0.01 M Tris pH 7.0 for 2 x 10 min (to remove iso-propanol from the gel).
9. Incubate gel in 0.1 M Tris pH 7.0 at 51oC for 50 min.
10. Wash the gel in 0.01 M Tris pH 7.0, for 10 min.
11. Stain the gel in 0.2% toluidine blue O/0.01 M Tris pH 7.0 for 10 min.
12. Destain the gel in 0.01 M Tris pH 7.0 for 10 min and 2 x 20 min.
13. Rinse the gel in 10% glycerol/0.01 M Tris pH 7.0 for 10 min.
14. Sandwich and dry the gel with cellophane.
   -- cut two pieces of cellophane to the size of the frame
   -- wet the cellophane thoroughly with water
   -- smooth one of the two cellophane sheets over the solid backing plate
   -- center the gel on the cellophane
   -- smooth the second sheet over gel
   -- clamp the frame over this sheet using several binder clamps
   -- set aside to dry overnight
   -- remove dried gel from frame, cut off excess cellophane, and store the gel flat.

To record the gels, we photograph them using a 35 mm camera mounted above a white light box, and using Kodak "technical pan" film. Place the gel on a glass plate on the light box, covering up the rest of the light box surface with black paper to block the light around the gel. Use of a yellow plastic filter underneath the gel improves definition. We take a range of photographs at a shutter speed of 1/8 second and f-stops of 5.6 to 11. Adjust if necessary for your equipment.

	Separating Gel (11%)	Stacking Gel
30% acrylamide	4.7 mL	750 µL
2% bis	1.9 mL	300 µL
1 M Tris-HCl	5.0 mL (pH 9.0)	320 µL (pH 6.8)
RNA solution	25-30 mg (volume=X µL)	-
0.1 M Tris pH 9.15	(730 - X) µL	-
H2O	-	3.6 mL
10% APS*	100 µL	40 µL
TEMED*	10 µL	4 µL

* add just before pouring
(Adjust separating gel recipe as necessary according to the concentration of your RNA sample. See protocol for preparing RNA below.)

Running Buffer		2x Sample Loading Buffer	
Glycine	21.64 g	50 mM Tris, pH 6.8	0.50 mL 1 M Tris, 6.8
Tris base	5.00 g	2% SDS	1.00 mL 20% SDS
SDS	1.50 g	10% Glycerol	1.25 mL 80% glycerol
H2O	1.50 L	0.025% BPB	0.25 mL 1% BPB
		H2O	bring to 10.00 mL

Preparing RNA to cast in RNase activity gels

1. Dissolve 100 mg/ml of torula yeast RNA (Sigma R-6625) in 1 M Tris pH 9.0 (100 ml is a convenient volume to make).
2. Phenol extract RNA solution, spinning 10 minutes at 7000 rpm.
3. Chloroform extract RNA twice, spinning 10 minutes at 7000 rpm.
4. Precipitate with 1/10 volume of 3 M NaOAC and two volumes of EtOH. Spin 10 minutes at 12,500 rpm.
5. Wash pellet once with 70% EtOH.
6. Dry pellet in lyophilizer or speed-vac, with tube covered with Parafilm that has holes poked in it.
7. Dissolve pellet in 0.1 M Tris pH 9.0 (about half the starting volume).
8. Check A260 of 1:1000, 1:5000 and 1:10,000 dilutions to quantitate.

Semi-Dry Electrophoretic Blotting of Protein

Transfer Buffer

Bjerrum and Schafer-Nielsen transfer buffer containing SDS:
48 mM Trizma base, 39 mM glycine, 20% methanol, 0.00375% SDS

Dissolve 5.82 g of Trizma base and 2.93 g of glycine in about 700 ml of H2O.
Add 1.875 ml of 20% SDS and 200 ml of methanol.
Adjust the volume to 1 liter with H2O.

Electrophoretic transfer with Bio-Rad Trans-Blot SD.

1. Following electrophoresis, equilibrate the gel in transfer buffer.

2. Cut nitrocellulose membrane to the dimensions of the gel. Wet the membrane in the transfer buffer for 15-30 min.

3. Cut two pieces of filter paper (Whatman 3MM) to the dimensions of the gel. Saturate the filter paper with the transfer buffer.

4. Place a pre-soaked sheet of the filter paper onto the platinum anode of the Trans-Blot SD.

5. Place the pre-wetted blotting media on top of the filter paper.

6. Place the equilibrated gel on top of the transfer membrane, aligning the gel on the center of the membrane.

7. Place another sheet of pre-soaked filter paper on top of the gel.

8. Carefully place the cathode onto the stack. Press to engage the latches with the guide posts without disturbing the filter paper. Place the safty cover on the unit.

9. Transfer mini-gels for 15-30 minutes at 10-15 V. Large gels can be transfered for 30-60 minutes at 15-25 V. Do not exceed 25 V. A current limit (3 mA/cm^2 for large gels; 5.5 mA/cm^2 for mini-gels) is recommended to prevent over-heating.

Coomasie Blue Staining of Protein Gels

aterials

- Protein gel from Exercise 4.2
- 0.25% (w/v) Coomasie Brilliant Blue R 250 in methanol-water-glacial acetic acid (5-5-1), filtered immediately before use.
- 7% (v/v) acetic acid
- Commercial destaining unit (Optional)

ocedure

1. Place a gel prepared as in Exercise 4.2 in at least 10 volumes of Coomasie Blue staining solution for 2-4 hours. Rock gently to distribute the dye evenly over the gel.
2. At the conclusion of the staining, wash the gels with several changes of water.
3. Place the gels into a solution of 7% acetic acid for at least 1 hour.
4. If the background is still deeply stained at the end of the hour, move the gels to fresh 7% acetic acid as often as necessary.
   If a commercial destainer is available, this will decrease the time required for stain removal. Follow the manufacturer's directions for use of the destainer. 14
5. Place the gels into containers filled with 7% acetic acid as a final fixative.
6. Photograph the gels or analyze the gels spectrophotometrically.

otes

omasie Brilliant Blue R 250 is the most commonly used staining procedure for the detection proteins. It is the method of choice if SDS is used in the electrophoresis of proteins, and is nsitive for a range of 0.5 to 20 micrograms of protein. Within this range, it also follows the er-Lambert law and thus can be quantitative as well as qualitative. The major drawback is the ngth of time for the procedure and a requirement for destaining. Overstaining results in a nificant retention of stain within the gel, and thus a high background stain, which might literate the bands. The length of time for staining must be carefully monitored, and can range m 20 minutes to several hours. If maximum sensitivity is desired, one should try 2 hours for a % gel and 4 hours for a 10% gel. Destaining must be monitored visually and adjusted cordingly.

CuCl2 Staining of SDS-PAGE Gels

is protocol is 2-3 times more sensitive than coomasie blue staining, it is much quicker, and the s can be stored at 4° for many months without protein degredation. Reference: Lee, C., Levin, Branton, D., (1987). Copper Staining: a Five-Minute Protein Stain for SDS-PAGE Gels. *al. Biochem.* **166**: 308-312.

lutions

upric Chloride Stain

M cupric chloride 5.1 g cupric chloride

to 100 ml with Q

rocedure

Run a standard SDS-PAGE gel.

Following electrophoresis, transfer directly to the Cupric Chloride Stain solution and shake for minutes.

Rinse in Q and transfer to saran wrap on a black background to visualize the bands.

The gels can be photographed under white light and stored for many months at 4°. The bands ill not difuse unless they are cut from the gel.

Topic 18. Protein Sequencing Technique

Chemistry

The Edman chemistry uses different reagents from earlier implementations e.g. gas phase sequencer (Hewick *et al.* 1981, Hunkapiller *et al.* 1983) but the principles remain the same. The base N-methylpiperidine makes the immobilised sample basic, after which the amino group of the terminal amino acid residue reacts with phenylisothiocyanate to produce an anilinothiazolinone amino acid. Liquid or gaseous trifluoroacetic acid cleaves the derivatized amino acid from the end of the peptide chain. Another reaction converts the anilinothiazolinone amino acid to a phenylthiohydantoin (PTH) amino acid which is then identified by reverse phase chromatography on a C18 column using an on-line HPLC. We use the current Applied Biosystems operating procedures and programs for both the sequencer and the on-line PTH-amino acid analyzer.

This process is repeated for each new terminal amino acid that results from a cycle of derivatization. sequencing, length of The process can not be continued indefinitely for two reasons: firstly the cycle of derivatization and hydrolysis is, on average, only 94% efficient and secondly there is acid cleavage of peptide bonds. After 20 cycles (amino acids), there is 0.9420 =29% of the initial quantity of amino acid produced by a derivatization and hydrolysis cycle. One factor decreasing efficiency is loss of sample from the support. For samples on PVDF, 1 to 10% of the sample may be lost off the support in each cycle Matsuidaira (1989b); this work used Immobilon-P and it is possible that later formulations of PVDF will lose less protein. It is also common to see major losses of sample at the end of a peptide immobilized on a Polybrene treated glass fiber filter, and hydrophobic peptides do not stick well to Polybrene.

Another factor reducing efficiency of the reaction occurs when a serine residue appears in the sequence. The hydroxyl group of serine may be esterified by acid, and when the amino group of the serine residue becomes N-terminal during the sequence, the esterified hydroxyl group may participate in an OThe acid used to cleave the derivatized amino acids from the end of the chain also cleaves peptide bonds within the chain. There is specific cleavage after aspartate residues in anhydrous acid (Brandt *et al.*, 1980) as well as non-specific cleavage. Also Ne the yield of the new amino acid gets smaller relative to the background, it becomes impossible to determine which amino acid has been newly released from the original protein. Thus even if there is sufficient sample, 500 pmole or more, we cannot reliably interpret sequence data after about 50 amino acids or less, depending on the protein. Isoaspartate formation also reduces sequence yield. Formation of isoaspartate appears relatively slow, even under the alkaline conditions favouring its formation (Geiger and Clarke, 1987; Violand *et al.*, 1990). However if it occurs, the sequence will stop at that point (Teshima *et al.*, 1991).

Blocked proteins

The reaction described above requires a free -amino group -amino group on the terminal amino acid; if one is not present, we cannot derive any sequence data from the sample without further treatment. The most common moieties blocking the terminal amino group are either acetyl or other acyl groups, or pyroglutamate residues formed by cyclization of glutamine or glutamate

residues. Other blocking groups have also been found (Wold 1981) and it is believed that other less well defined reactions occur in vitro that also block the terminal amino group.

Sample preparation

The three most common ways of supplying samples are: in a volatile solvent, in a buffer solution, in a gel.

Samples in volatile solvents are dried in 15 al aliquots on a glass fiber disk which has been coated with Polybrene which retains proteins on the disk by non-covalent interactions.

When a sample is in buffer, it is spun on to a disk of PVDF in a modified centrifugal filter (ProSpin cartridge) or the solution may be drawn through the PVDF by absorption forces (ProSorb). In this way, salts and detergents pass through the PVDF membrane while the protein remains on the PVDF which is loaded into the sequencer Sheer, 1990). Binding of samples is reported to be more efficient in non-acidic solution (Baumann, 1990).This technique should not be used for proteins which are to be manipulated further because of the difficulty of removing proteins from PVDF (see later).Amicon sell ion exchange membranes (Microcon-SCX) for the same purpose, using a centrifuge to pass the solution through the membrane; the peptides or proteins can be eluted from these membranes, unlike with PVDF where removal of adsorbed molecules is difficult.

Samples purified by gel electrophoresis are blotted to a sheet of PVDF, which is stained and the band of protein cut out and loaded into the sequencer. This technique is described later.

Another method of loading samples is to attach them covalently to a support. A commercially available support is the Sequelon family of modified PVDF from Millipore. In Applied Biosystems instruments, some peptides sequence poorly on these supports, while some give better data. We obtained better results when sequencing a synthetic peptide applied to a Polybrene treated glass fibre filter then the same peptide attached to a Sequelon membrane. However when identifying phosphorylation sites, it is necessary to covalently bind peptides to a PVDF based support to prevent sample losses from the solvents used.

Identification of amino acids

To identify the PTH amino acids produced by the sequencer, they are injected on a C18 column, which is equilibrated in acetate buffer at about pH 3.9 and eluted with a gradient of acetonitrile. The elution positions of arginine, histidine, and pyridylethyl-cysteine are sensitive to pH and buffer concentration. A mixture of PTH amino acid standards are injected at the beginning of each sequencer run for identification of amino acids and quantitation of the amino acids released from the protein being sequenced.

To see cysteine in a sequence, it must be alkylated, preferably before the peptide is loaded on the sequencer, to prevent destruction by acid. Applied Biosystems recommends alkylation with 4-vinylpyridine according to the procedure listed in Appendix A. Another procedure is that of Andrews and Dixon (1987). Alkylation with 4-vinylpyridine produces a pyridylethyl-cysteine residue which can be identified in analyses of both PTH-amino acids (from the sequencer) and PTC-amino acids, which are produced in our procedure for amino acid analysis. Another preferred reagent is N-isopropyliodoacetamide (Molecular Probes, Oregon) (Krutzsch and

an, 1993) which gives carboxamidoacetamide-cysteine, whose PTH derivative is easily rated. Reaction of cysteine with iodoacetate produces carboxymethyl-cysteine (Gurd 1972) ch may be confused with glutamine in an analysis of PTH-amino acids; note that iodoacetate also alkylate methionine (Jones *et al.*, 1994). Cysteine can be alkylated after electrophoresis VDF (Ploug *et al.*, 1992).

ds of serine , and threonine are low because these residues can be esterified by TFA, and the rs can dehydrate, giving dehydroalanine from serine. In our sequenator, DTT reacts with the ydroalanine giving a peak that elutes between alanine and tyrosine. The corresponding ducts from threonine are seen with large amounts of threonine. The acyl group that esterifies hydroxyl group may undergo a ON shift, acylating the terminal amino group, thereby cking it and preventing further sequencing. This shift may explain the drop in yield seen after ne and threonine when sequencing isolated peptides (Aebersold, 1989). Prolonged exposure protein to formic acid appears to esterify serine and threonine (Tarr and Crabb, 1983) but noethanol appears to reverse the reaction.

stidine and arginine may give lower yields, reportedly because of poor extraction from the id supports for the sample.

ptophan is also prone to oxidation by sunlight and other agents (Pirie, 1971). Methionine may oxidized in proteins but this does not appear to be a common problem (Brot and Weissbach, 83).

odified amino acids

osphorylated amino acids cannot be identified by standard sequencing procedures. osphoserine and phosphothreonine eliminate the phosphate group giving a low yield of a rivative; the released phosphate is poorly extracted from the sample support (Soderling and alsh, 1982). Phosphotyrosine is insoluble under standard conditions and not easily identified y chromatography. However if the phosphoproteins are labelled with P32, the position of the belled residue can be determined after digestion with a proteinase of known specificity Shannon and Fox, 1995). Chemical conversion of phosphoserine is another method of dentifying this residue (Meyer *et al*, 1986).

Heavily glycosylated proteins are reported not to give good data because the hydrophilic carbohydrate chains interfere with the non polar reagents and solvents used in the sequencing reaction.

In some proteins, palmitic acid (C16) is attached to cysteine, serine or threonine through a (thio)ester link (Sefton and Buss, 1987); myrisitic acid (C14) is attached through an amide link to amino terminal glycine residues. One method of identifying acylated cysteine is to deacylate with hydroxylamine and alkylate the cysteine (Bizzozero *et al.*, 1990).

ways of protein sequencing

N-terminal protein sequencing
Internal protein sequencing
C-terminal proein sequencing

dman Degradation Method: (most popular)
nciple and Concept:
he Edman reaction is a series of chemical reactions which remove one amino acid at a time om amino terminus of a protein, releasing a derivatized amino acid(phenylthiohydantoin PTH) hich may be chromatographically identified (reversed phase). A simplified schematic diagram f the process is shown in Figure 1. Release of amino acids from the amino terminus in a quential manner is possible because the Edman procedure consists of three chemical reactions hich proceed under different pH conditions.

The reagents and extracting solvents are passed over the immobilized film of protein. The rst step is the coupling step (Figure 1) which occurs a high pH values and results in formation f phenylthiocarbamoylated (PTC) amino groups on the protein (both the alpha amino group of he protein and the epsilon amino groups of any lysyl residues are modified). This leaves the ntact peptide shorten by one amino acid. The Edman procedure can be repated on the second mino acid. The coupling step will be incomplete if the pH is insufficiently high. The second eaction is the cleavage step which occurs at low pH, resulting in release of an nilionothiazolinone (ATZ) form of the amino acid and regeneration of a free amino terminus on he protein. The cleavage reaction does not occur at high pH values. The ATZ-amino acids are oluble in organic solvents such as ethyl acetate and so can be extracted from the insoluble rotein for analysis. The ATZ-amino acid is converted to the phenylthiohydantoin (PTH) erivative in a separate reaction generally exposure to strong acid. The PTH-amino acids are more stable and a bit more amenable to chromatographic resolution.
The model 470 is a gas phase sequenator in the sense that the coupling buffer (12.5% riethylamine) for the Edman reagent and the cleavage reagent (100%) trifluoroacetic acid are elivered in gaseous phase with argon as the carrier gas. This reduces the loss of protein from the ample cartridge. All other solvents and reagents are delivered as liquids.

1. N-Terminal Sequence Analysis

Procedure : Protein Sequence Analysis and Sample Preparation

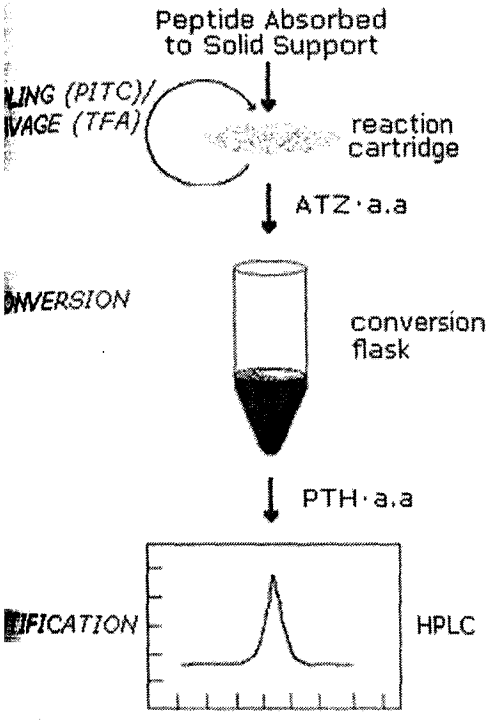


Figure 1: Overview Sequence Analysis is performed on new, state of the art instrumentation (Procise 494 HT, Perkin-Elmer) using Edman chemistry. Edman degradation can be used to sequence proteins rapidly, and with little starting material. The process is easily automated, and sequencers can operate sometimes with as little as a few picomoles of sample. The machine performs the coupling, cleavage, conversion and identification steps for each amino acid residue, and researchers often choose to identify 10-15 residues from the protein. The protein may then be identified by searching a database to match the sequence and properties of the protein.
See description of the process followed.

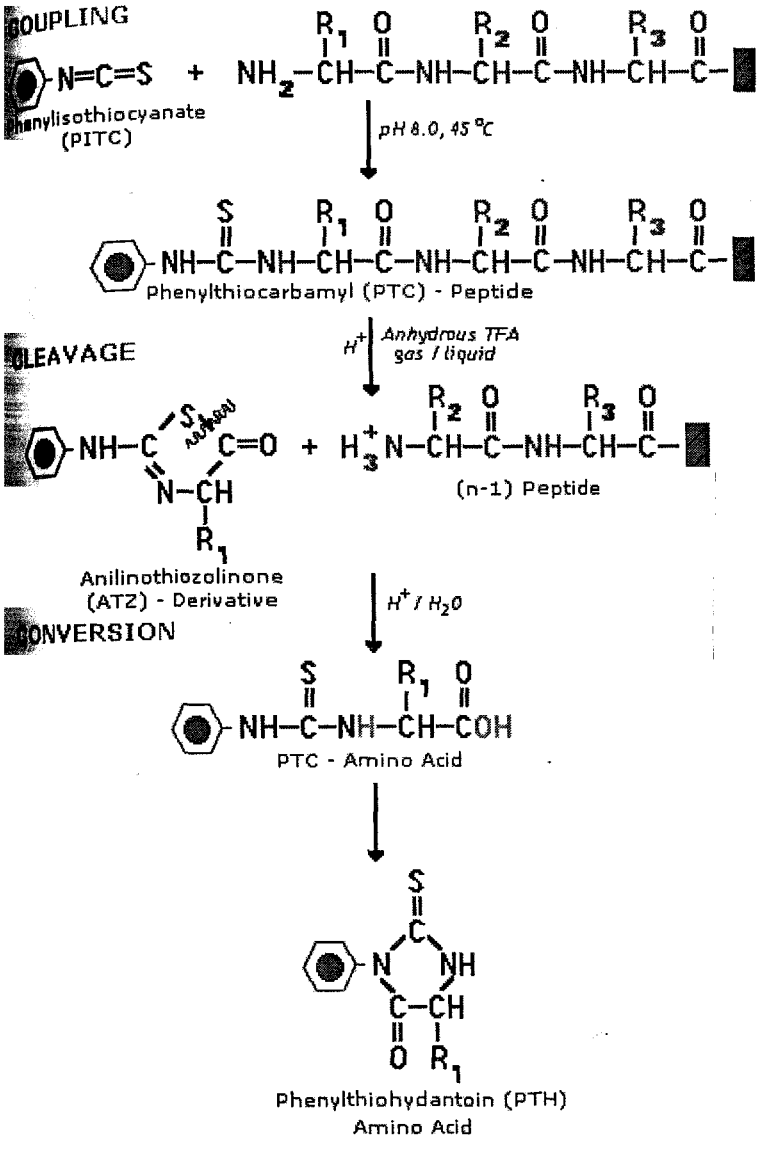
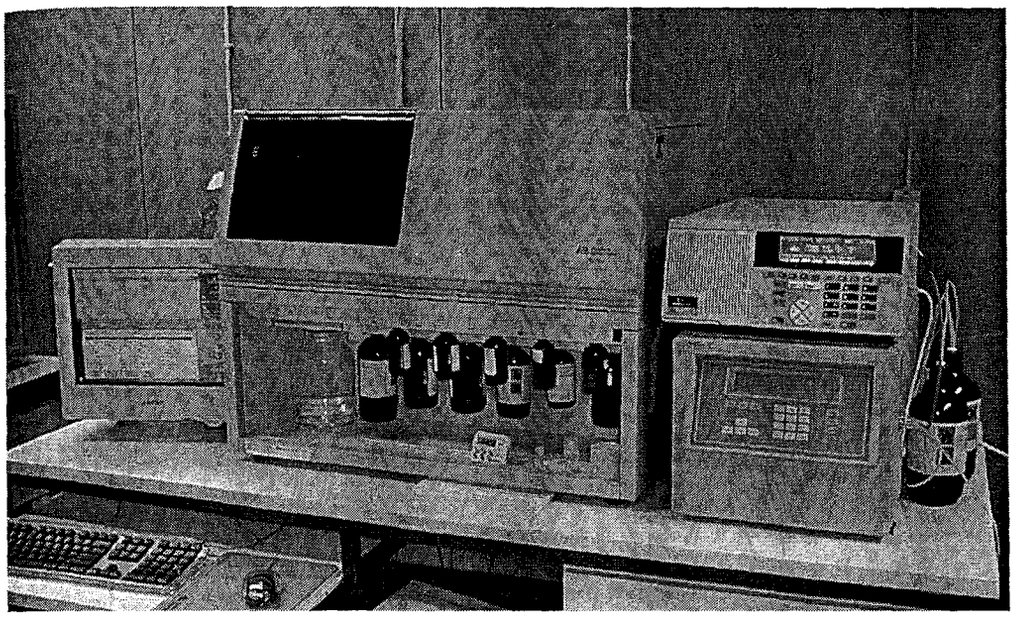


Figure 2: Chemical Description
In Edman sequencing, a protein is attached to a solid support such as a chemically modified glass disk or a porous polyvinylidene fluoride membrane in the reaction cartridge. It is then coupled to phenylisothiocyanate (PITC) at pH 8 and 45°C. The free N-terminal amino group reacts with the carbon of the isothiocyanate group to give the phenylthiocarbamyl (PTC) derivative of the peptide. The next step is cleavage of the PTC derivative using anhydrous trifluoroacetic acid to give the anilinothiozolinone (ATZ) derivative of the N-terminal amino acid, and the peptide with one fewer amino acid, which is free to undergo further couplings and cleavages. The ATZ residue is then filtered into the conversion flask, where it is converted to the phenylthiohydantoin (PTH) amino acid. This is a two step process. First, the ATZ derivative is hydrolyzed under aqueous, acidic conditions to give the PTC amino acid. The acid then cyclizes to give the stable PTH derivative. These derivatives are then injected into an HPLC column where its retention time is compared with that of known PTH amino acid standards.

Source: http://www.biotech.uiuc.edu/centers/Proteomics/Proteinscience



http://www.itqb.unl.pt/Services/Analytical_Services/Protein_Sequencing

Automate protein sequencing

2. Internal Protein Sequencing

Internal protein sequencing is a technique to provide the capability of generating internal quence information from proteins that have blocked N-termini. The protein can be digested ith enzymes either in solution, on PVDF, or in an SDS-PAGE gel. The resulting peptides can are extracted and then separated with the Microblotter's capillary HPLC. The resulting peptide ctions can then be collected utilizing the Microblotter's dynamic solenoid blotting the ctions on a strip of PVDF. After collecting the peptides on the PVDF strip, the peptides can be quenced using the PROCISE 494 or other automate protein sequencer.

estern Blotting

aterials and Methods

quipment:

o-Rad Mini-Trans-Blot Cell
ectrophoretic Blotting Apparatus
aker (orbital type)

hemicals:

urce Cat. No.
oBlott ABI 400994 TM
crylamide Bio-Rad 161-0100
mido Black Bio-Rad 161-0402
mmonium persulfate Bio-Rad 161-0700
romophenol blue Bio-Rad 161-0404
oomassie©Brillant Blue R-250 Bio-Rad 161-0400
cyclohexylamino-1-propane-
lfonic acid (CAPS) Aldrich 16,376-7
lycerol BRL 5514UA
lycine Bio-Rad 161-0717
-mercaptoethanol Sigma M-6250
ethanol ABI 400470
,N'-methylene bisacrylamide Bio-Rad 161-0200
,N,N,N-Tetramethylene-
amine(TEMED) Bio-Rad 161-0800
onceau Sigma P7767
rva Blue G Serva 35050
odium dodecyl sulfate (SDS) Bio-Rad 161-0301
odium thioglycolate Sigma T-0632
ricine Aldrich 16,378-3
ris (hydroxymethyl)aminomethane Bio-Rad 161-0716
riton X-100 (NP-40) Bio-Rad 161-0407
rea Bio-Rad 161-0730

DS-PAGE Systems

number of systems used with gel electrophoresis have been described . These are the Tris-lycine [4 5]

d Tris-Tricine buffer systems for mini-gles. Tris-glycine is used for separating proteins larger
an 6
0kDa; the Tris-Tricine buffer system is used to resolve peptides and low-molecular weight
oteins
well as larger proteins. Either buffer system can be used prior to electroblotting.

ris-Glycine stock solutions and buffers

ower tris (4 X) Dissolve 36.34 g tris base and 0.8 g SDS in 150 mL D.I. H2O.
itrate with 6N HCl to pH 8.8. Add D.I. H O to final volume of 200 2
L.
pper tris (4 X) Dissolve 12.11 g tris base and 0.8 g SDS in 150 mL D.I. H2O.
itrate with 6N HCl to pH 6.8. Add D.I. H O to final volume of 200 2
L.
0% Acrylamide Dissolve 30g acrylamide and 0.8 g bisacrylamide in D.I. H O to final 2
0.8% T, 2.6% C) volume of 100 mL.
ectrophoresis Buffer Dissolve 3.03 g tris base, 14.4 g glycine and 1 g SDS in 1 L D.I. H O. 2
mmonium persulfate Dissolve 100 mg ammonium persulfate in 1 mL D.I. H O. Store the 2
0%) solution at 5 C. Discard after one week. o
mple Preparation (2X) Mix 1 mL glycerol, o.5 mL 2-mercaptoethanol, 0.3 g SDS, 0.25 mL
lution of 0.05% (w/v) bromophenol blue and 1.25 mL Upper Tris (4X)
ffer in D.I. H O to a final volume of 5 mL. 2
ilter solutions with 0.2 µ Nalgene© filter.

ris-Glycine Gel Solutions

e following proportions are for preparing 2 gels (0.5 x 60 x 90 mm) for mini-gel systems.

tock Solutions	**10% Resolving el (8 mL)**	**15% Resolving Gel (8 mL)**	**4% Stacking Gel (4 mL)**
0% Acrylamide	2.66 mL	4.0 mL	0.52 mL
pper tris (4X)	-	-	-
ower tris (4X)	2.0 mL	2.0 mL	-
.I. H2O	3.34 mL	2.0 mL	2.47 mL
e-gas for 5 minutes			
EMED	4 µL	4 µL	4 µL
mmonium ersulfate (10%)	40 µL	40 µL	40 µL

lectrophoresis Conditions for Tris-Glycine

Sample preparation: Mix one volume (1-10 µL) of sample solution with one volume of mple preparation solution (2X) and heat in boiling water bath for 2 minutes. The sample now ready to load into the sample well for electrophoresis.

When using the mini-gel apparatus, the recommended electrophoresis condition for optimal solution with minimal thermal band distortion is 7 mA constant current (50-90 volts) for pproximately 1 hour. For maximum separation, allow the bromophenol blue dye to run to ithin 1 cm of the end of the gel.

ris-Tricine stock solutions and buffers

Anode Buffer (10 X) Dissolve 121.9 g tris base in 400 mL D.I. H O. 2
2.0 M Tris HCl) Titrate to pH 8.9 with HCl. Add D.I. H O to final volume 500 2
mL.
Cathode Buffer (10 X) Dissolve 60.55 g, 89.58 g tricine and 5 g SDS in 400 mL D.I.

1.0 M Tris, 1.0 M Tricine, H O. Titrate to pH 8.25 with HCl. Add D.I. H O to final 1.0% 2 2
SDS) volume 500 mL.
Acrylamide Solution Dissolve 48 g acrylamide and 1.5 g bis-acrylamide in 100 mL
(49.5% T, 3% C) D.I. H O. 2
Gel Buffer (3 X) Dissolve 181.5 g tris base, 1.5 g SDS in 400 mL D.I. H O. 2
Titrate to pH 8.45 with HCl. Add D.I. H O to final volume 2
of 500 mL.
50% glycerol (v/v) Dissolve 250 mL glycerol in 250 mL D.I. H O. 2
Sample Preparation Mix 1.0 mL of 100% glycerol, 0.5 mL 2-mercaptoethanol, 2.5
Solution (2X) mL of 20% (w/v) SDS, 0.62 mL Tris (pH 6.8, 1M), 0.01 mL
of 10% Serva BlueG (w/v) in D.I. H O to a final volume of 2
5 mL.
Filter solutions with 0.2 µ Nalgene© filter.

Tris-Tricine Gel Solutions

The following proportions are preparing 2 gels (0.5 x 60 x 90 mm).

Stock Solutions	10% Resolving Gel (15 mL)	4% Stacking Gel (12.5 mL)
Acrylamide solution	3.0 mL	1.0 mL
50% glycerol	4.0 mL	-
Gel Buffer (3X)	5.0 mL	3.1 mL
H O	3.0 mL	8.4 mL 2
De-gas for 5 minutes		
TEMED	7.5 µL	7.5 µL
Ammonium persulfate (10%)	75 µL	100 µL

Electrophoresis Conditions for Tris-Tricine

1. Sample preparation: Mix 1 volume of sample with 1 volume of sample preparation solution (2 X) and heat in boiling water bath for 2 minutes.
2. Electrophoresis Conditions: 20 mA constant current for 2-3 hours. For maximum separation, allow the serva blue dye to run to within 1 cm of the end of the gel.

Electroblotting

CAPS Buffers1

Stock CAPS (10 X) Dissolve 22.13 g CAPS in 900 mL of D.I. water. Titrate with 2N NaOH (20 ml) to pH 11, and add D.I. water to a final volume of 1L.
Store at 4EC.
Electroblotting buffer Prepare 2L by mixing 200 mL of the 10X CAPS buffer with 200 ml (10 mM CAPS in of methanol and 1600 mL of D.I. water.
10% MeOH)

Tris-Glycine Buffer7,8

Electroblotting buffer Dissolve 3.0 g Tris and 14.4 g glycine in a solution of 900 mL H O [25 2 mM Tris base, and 100 mL methanol.
192 mM glycine, 10%
(v/v) methanol (pH 8.3)]

Procedure:

1. Remove the gel from the electrophoresis cell and soak it in 100 mL of electroblotting buffer for 5 minutes.
2. Meanwhile, wet ProBlott with 100% methanol for a few seconds, then transfer it to the TM

lotting buffer.
ote: Be sure the ProBlott membrane is completely moistened with methanol TM
efore transferring it to the blotting buffer. It if dries out, wet it with methanol
gain.
. Dip the sponges and filter papers in a separate container of blotting buffer before starting to
ssemble the transblotting sandwich.
. Assemble the transblot sandwich in the following order starting from anode side:
ponge, filter paper, 2 sheets of ProBlott , gel, filter paper, sponge. TM
Make sure there are no bubbles between the gel and the membrane.
. Pour 1L of electroblotting buffer into the transblot cell and insert the transblot sandwich.
lectroblotting conditions are dependent of the buffer, as follows:
APS:50V (170 mA-100mA) at room temperature for 30 minutes. Tris-Gly:40V (300 mA)
t room temperature for 1-4 hours.
Note: Electroblotting times vary and are dependent on protein sample. Longer
lectroblotting times may be necessary with proteins that are 70 kDa or larger.
. After transfer is complete, remove ProBlott from transblotting sandwich and rinse with D.I. TM
water before staining.
Protein Detection
Protein sampleson ProBlott can be detected with conventional staining techniques, such as TM
Coomassie© Brilliant Blue, Ponceau S or Amido Black.
Coomassie© Brilliant Blue Staining Solution
0.1% Coomassie Blue R-250 in 1% acetic acid/40% MeOH
Dissolve 1.0 g Coomassie blue in 400mL of methanol. Stir for one hour. Add 10 mL acetic acid
and
590 mL D.I. water. Stir for 30 min, then filter with a Nalgene© filter (0.45μ pore size).
Destaining solution
50% MeOH in D.I. water
Coomasie Blue Staining Procedure:
1. Remove the ProBlott membranes (both primary and secondary) from the transblot TM
sandwich, and rinse them with D.I. water..
2. Saturate the ProBlott membranes with 100% MeOH for a few seconds. TM
3. Stain the ProBlott membranes (one at a time) with the staining solution using constant TM
orbital shaking. Protein bands should appear within one minute.
CAUTION: If the ProBlott membrane is over-stained, it may require a TM
longer destaining time to prevent a high background.
4. Destain the ProBlott membranes by soaking them in destaining solution. For quicker TM
destaining and improved contrast, change the destaining solution several times, allowing the
membranes to soak for 1-2 minutes between changes.
5. Rinse the membrane thoroughly with D.I. water, and excercise the bands of interest with a
clean razor blade.
Amido Black Stainig solution
0.1% Amido Black in 1% acetic acid/40% MeOH
Dissolve 1.0 g Amido Black in 400 mL of methanol. Stir this solution for at least one hour. Add
10
mL acetic acid and 590 mL of D.I. water. Stir for another 30 minutes, then filter the solution with
Nalgene filter (45μ pore size).

The staining procedure for Amido Black is the same as that described for Coomasie Blue, with the substitution of D.I. water as the destaining solution.

Ponceau S Staining solution

0.2% Ponceau S in 1% acetic acid

Dissolve 0.4 g of Ponceau S in 198 mL of D.I. water and stir for 30 minutes. Add 2 mL of acetic acid to the mixture.

. Remove the ProBlott membranes from the transblotting sandwich and rinse them with D.I. TM water.

. Stain the ProBlott membranes in the Ponceau S staining solution with constant orbital TM shaking. Protein bands should appear within one minute.

. Destain with a D.I. water rinse.

. Excise the bands of interest with a clean razor blade.

In gel Digestion Procedure for obtaining internal peptide sequences:

In Gel Digestion:

. Cut the gel section containing the protein of interest into 1 x 2 mm pieces and place in an Eppendorf tube. Repeat for the blank section of gel.

. Add 250µl 200 mM NH_4HCO_3 / 50% CH_3CN to the gel pieces. (REMOVES CB)

. Wash for 30 minutes at room temperature on a rocker table.

. Remove wash and SAVE.

. Add 100µl NH_4HCO_3 / 50% CH_3CN to the gel pieces.

. Calculate the volume of 45 mM DTT needed to give a final [DTT] of 1 mM in the sample.

. Add the above volumes of DTT and then incubate at 37°C for 20 minutes.

. After incubation, add an equal volume of 100 mM iodoacetic acid or 200 mM MNS (methyl 4-nitrobenzene sulfonate) as DTT.

. For carboxymethyl cysteine, incubate in the dark for 20 minutes. For s-methyl cysteine (MNS treatment), incubate at 37°C for 40 minutes.

0. After incubation, remove supernatant and SAVE.

1. Repeat steps 2-4.

2. Speedvac gel pieces to complete dryness.

3. Immediately before use, make up a 0.033 mg/ml enzyme stock solution in 200 mM NH_4HCO_3 by mixing 5.0 µl of 0.1 mg/ml enzyme stock with 10.0 µl 200 mM NH_4HCO_3.

4. Add a volume of 0.033 mg/ml trypsin or lysyl endopeptidase that equals the initial gel volume (ie. 15µl per 15 mm^3 of gel.

5. Incubate 37°C for 15 minutes. (LOOK TO MAKE SURE GEL IS TOTALLY IMMERSED).

6. If the gel pieces are not totally immersed, add additional enzyme solution until they are just covered.

7. Incubate at 37°C for 24 hours.

8. Extract peptides by adding at least 100µl 0.1% TFA, 60% CH_3CN and shaking on a rocker table at room temperature for 60 minutes.

9. Remove supernatant containing peptides and repeat steps 18 and 19.

0. Speedvac dry the combined washes.

1. Redissolve the dried peptides in 20µl 0.05% TFA / 25% CH_3CN.

2. Sample is ready for reverse phase HPLC separation of peptides on the 173A Microblotter.

-PVDF membrane digestion procedure for obtaining internal peptide sequence ormation:

PVDF Digestion:

ll tubes and pipette tips rinsed with 0.1% TFA/50% ACCN prior to use!

Excise bands and cut into 1 mm x 2mm pieces.
Pre-wet with MEOH and place into microcentrifuge tube.
Destain for 1 min with 0.1% TFA/50% ACCN
Rinse 5X with H2O. Sonicate 5 min per wash.
Add 50µl digestion buffer to PVDF
igestion buffer = (100 mM TRIS, 1% Reduced Triton X-100, 10% ACCN, pH 8.0)
Add 5µl 45 mM DTT and incubate 55° C for 30 min.
Add 5µl 100 mM Iodoacetic acid and incubate 30 min @ room temp in dark.
Remove digestion buffer and wash 1X with 50µl digesiton buffer.
Add 50µl digestion buffer to PVDF.
0. Add 2µ Trypsin (Modified Trypsin-Promega) at 0.1 mg/ml.
1. Incubate 37° C for 4 hrs.
2. Add another 2µl Trypsin (0.1mg/ml).
3. Digest overnight at 37° C.
4. Sonicate sample for 5 min, centrifuge and remove supernatant (save).
8. Extract again with 25µl of 0.1% TFA/50% ACCN by sonicating for 5 min.
9. Centrifuge and save supernatant.
0. Extract again with 25µl of 0.1% TFA by sonicating for 5 min.
1. Centrifuge and save supernatant.
2. Pool all saved supernatants and store at -20° C. The sample is now ready for separation on he 173A Microblotter.

3. C-Terminal Protein Sequencing

Carboxy-terminal sequencing

There are three strategies to obtain C-terminal sequence data.

A few laboratories offer C-terminal sequencing by chemical methods; hundreds of picomoles of sample are needed and 5 amino acids of sequence is a standard result. Proline may not be identified.

A traditional procedure is to cleave amino acids from the C-terminus using a broad specificity carboxypeptidase, usually carboxypeptidase Y or a mixture of carboxypeptidases A and B. By noting the time at which levels of different amino acids reach a maximum, one can deduce the sequence of 4 or 5 residues (Hayashi, 1977). Sequencing of cytochrome P-450 from the C-terminus has been unreliable, possibly because of the presence of proline (Black and Coon, 1986). An example of this procedure is in Asano *et al.* (1986).

More recently, Patterson *et al.* (1995) released amino acids using a differing concentrations of carboxypeptidase and measuring the mass of the remaining peptide by mass spectrometry; this method is limited by the ability of the mass spectrometer to distinguish the loss of different amino acids from the peptide.

A third technique for obtaining C-terminal sequence data is isolation of the C-terminal peptide and determine its sequence by standard N-terminal sequencing; if there is sufficient peptide, one may be able to determine the complete sequence of the C-terminal peptide, especially if the amino acid composition can be determined to see if all amino acids present according to amino acid analysis are also found on sequencing. To isolate the C-terminal peptide, Pierce and Clontech sell a column of immobilized anhydrotrypsin. After digesting the protein with trypsin, the peptides are applied to the column. The internal peptides, with a C-terminal lysine or arginine bind to the anhydrotrypsin but the C-terminal peptide washes through the column. This technique has been applied to other purifications (Hirabayashi and Kasai, 1992).

Techniques to isolate the C-terminal peptide using esterification of carboxyl groups, digestion of the protein and isolation of the C-terminal peptide, which lacks a free carboxy group, have been described by Furka *et al.* (1983) and Dopheide and Ward (1981). Ion exchange has also been used (Gorman and Shiell, 1993).

A technique that may involve smaller volumes of reagents is to digest a protein with endoproteinase lys-C and couple the peptides to a Sequelon-DITC membrane. All peptides except the C-terminal peptide are covalently bound to the membrane through the _-amino group and all peptides are bound through the N-terminal amino group. Treatment with acid cleaves the N-terminal amino acid from the peptide in a reaction like that found in Edman sequencing. The C-terminal amino acid is released from the membrane while all other amino acids remain bound through their C-terminal lysine (Dixon *et al.*, 1993).

Sample Requirements

Samples must contain only one protein or peptide to avoid ambiguous data. If there is more than one polypeptide, an amino acid will be released from each polypeptide in every sequencing cycle. If the amounts of polypeptide are very different, the amounts of amino acid will be also, enabling each amino acid to be assinged to the correct sequence. If there are two protein present in similar amounts, the sequence data will be ambiguous. Because of differences in amino acid yields, and drops in yields after serine, threonine and proline, large differences in amounts of the proteins are needed to confidently determine the sequences.

Although the instrument can sequence 10 pmole or less protein, more sample is preferred. Larger amounts of sample increase the chance of identifying amino acids whose yield is poor, and obtaining a long sequence run. More importantly, with small quantities, samples are easily lost during handling, so that the amount seen for analysis is much less than the amount estimated. With a small quantity of sample, if there is no data, the question remains as to whether the sample is blocked and therefore unsequenceable, or whether there was insufficient sample.

Samples in volatile solvents suitable for direct application to the sample support should be in less than 200 µl, the volume of samples containing salts and other non-volatile material should be less than 1 ml; a PVDF strip up to 8 mm long can be loaded but further strips must be shorter to fit into the cartridge. A high concentration sample is preferable to a larger area of low concentration sample.

Substances which can interfere with sequencing reactions are listed in appendix D.

It is reported that dialysis may easily introduce contaminants, so dialysis should be done with thoroughly cleaned tubing and should be done in the presence of salt or acid to prevent contaminants adsorbing to the protein. We have seen samples that have been dialyzed and contain contaminants. Reportedly "molecular biology grade reagents ;contain insoluble material and UV-absorbing material (Matsuduira, 1989).

One of our preferred methods of final sample preparation is reverse phase chromatography using a gradient of water and acetonitrile with 0.1% trifluoroacetic acid. This procedure often gives a pure sample, and leaves the sample in a volatile solvent. Although samples may appear pure on electrophoresis or ion exchange chromatography, they often contain more than one peptide chain when analyzed on the sequencer. The data from such samples usually can not be interpreted reliably. Despite the resolution that reverse phase HPLC has for peptides, chromatography under one set of conditions may yield more than one peptide in a single peak so that rerunning under different conditions is desirable.

Not all of a sample applied to the sequencer will sequence. This is a standard observation. Hewlett-Packard report that treating a sample with 6M guanidine-HCl immediately before application to their instrument increased the initial yield from 52% (sample in 2% TFA) to 68% (Application Note 93-3). Lottspeich (1990) argues that incomplete reaction of applied protein is caused by interactions of the protein with the support, making portion of the protein unavailable to reactants.

Peptides can also be applied to PVDF; without further treatment, the loss of peptide from the PVDF is significant, but 100 µg of Polybrene applied to PVDF reduces sample wash out, at the

ense of reduced recovery of charged amino acids (Werner *et al.*, 1996). The addition of lybrene reputedly makes sequencing samples on PVDF as efficient as sequencing samples on ss fibre disks.

sses of Protein

is claimed that small amounts of protein are easily lost during drying (Grego *et al.*, 1985; lson *et al.*, 1986; Esch, 1984). One group reports that the amount of protein they see after tensive preparation is often less than 10% of the estimated value (Stone *et al.*, 1989b). When lating a tryptic peptide, Lu and Lai (1986) found that drying a tryptic peptide and dissolving it 50% formic acid gave about 10% of the sequenceable protein seen when the peptide was llected in a polypropylene tube and applied directly to the sequencer; formic acid may have locked the N-terminal thus reducing yields on the sequencer. Tempst *et al.* (1990) suggest that concentrations below 0.1 µg/µl proteins may adsorb to polypropylene tubes. Gary Hathaway npublished results) as a rule of thumb assumes 10 ng of protein/cm2 is adsorbed to olypropylene; 1 ml in a 1 ml pipet tip is 2.9 cm2, in a 1.5 ml tube is 6 cm2. Nevertheless, olypropylene tubes are the preferred tubes to use because glass may adsorb proteins. olystyrene, which is used in ELISA plates because of its protein adsorbing properties, should be voided. One company reports that their low protein binding syringe filters adsorb 2 µg of rotein/cm2, which shows how readily small quantities of protein may be lost.

n controlled pore glass, initial adsorption of protein is increased by the pI of a protein and the bsequent rate of loss is an inverse function of the molecular weight of the protein (Messing, 969). Initial adsorption is 11% for cytochrome c with rates of loss of 5% for this protein, which as the worst of those studied. Acidic urea solutions removed protein from the glass.

roteins differ in their adsorption to silica; the process can be saturated, appears insensitive to onic strength and is affected by conformation of the protein (Morrissey and Stromberg, 1974). epeated washings appear to remove some of the adsorbed protein (Bull, 1957). Adsorbance of p to 4 mg/mm2 of glass have been reported (Morrissey and Stromberg, 1974; Bull, 1957), which corresponds to 6 µg in a 10 x 50 mm tube.

Up to 50 ml of solution can be concentrated on small bore columns, at flow rates up to 1 ml/min Simpson *et al.* (1989a); recoveries are claimed to be 90%. This group suggests using 0.01% Tween 20 to reduce losses of peptides during handling.

Don Hunt's lab in Chemistry routinely handle very small amounts of peptides by collecting in phthalate or siliconized polypropylene tubes and freezing samples as soon as they are collected by placing them on dry ice. However acetonitrile dissolves material from siliconized tubes which may interfere with subsequent operations. They concentrate samples by drying in a Speed-Vac but do not go to dryness. Solutions used to try to redissolve peptides include 30% acetonitrile and 70% formic acid. Another group found that time and temperature do not have large influences on adsorption of a peptide to polypropylene tubes, but addition of TFA to 33% improves recovery of peptide to 90% from 50% (Erdjument-Bromage *et al.*, 1993).

An illustration of protein losses from pipetting into successive tube is provided by Stevenson *et al.* (1994) who showed that the non-ionic detergent Nonidet P-40 reduced the loss.

One report lists recovery of 10-30% from Centricon-10 units (Jones *et al.*, 1994) but >80% recovery when desalting with gel filtration columns. Amicon have since shown that treatment of Centricons with a number of solutions, including 5% Tween-20, 5% polyethylene glycol, 5% Triton X-100 and 5% SDS reduced losses (Amicon document 2301). Amicon have developed Centriplus concentrators for which they report 90% recovery from 15 ml of 1 µg/ml solution of BSA and 50% recovery from a 250 ng/ml solution.

Long, high speed centrifugation increases recovery of proteins during trichloroacetic acid precipitation; even at 0.1 µg/ml, some proteins give 80% recovery (Hwang and Chu, 1996).

We have observed complete loss of sample when digesting 20 pmole of BSA in a gel slice, but addition of Tween 20 or PVP-360 allowed recovery of peptides (Shannon, 1995, unpublished observations). However Tween 20 interferes with mass spectrometry, while PVP-360, which is compatible with mass spectrometry, elutes in the latter part of a reverse phase separation of peptides. Lottspeich at a meeting (1996) suggested PPG-4000 to reduce sample loss.

Supplying a sample for sequencing

The most convenient container for samples is usually a 1.7 ml microcentrifuge tube. Polypropylene is the preferred material for storing protein samples, and the 1.7 ml size allows ready retrieval of samples, whether liquid or pieces of PVDF. The tube must be labelled with unique information; "1" is not a unique identification. Tubes with frosted surfaces are preferred because ink easily rubs off smooth tubes, especially after removal from a freezer.

You also need to decide what information is useful. The ideal sequence run gives continuous data from one polypeptide, but not all runs are ideal. Many intact proteins give no data because the N-terminal is blocked; if there is sufficient protein, there will be a background of amino acid. Normally we stop the run as soon as it is clear that no data is being obtained. In some samples, there are two or more polypeptides, even though other analytical methods suggest that a sample is pure; sometimes we can decide which amino acid belongs to which sequence but often, we have at each cycle, two or more amino acids and we cannot decide which sequence each belongs to; in such cases, you must decide if this data is of any use. Sometimes there will be amino acids which are not seen, either because amounts are so low, or because of the presence of unalkylated cysteine or because of some modified amino acid which is not seen on our analytical system.

Results

The raw data from a sequencing run is a series of chromatograms of PTH-amino acids. Amino acids are identified by their elution times. The chromatogram also has peaks from the injection (these occur in the first 4 minutes of the chromatogram), and the derivatives PMTC, DPTU and DPU formed by reactions involving PITC. These derivatives are useful for aligning successive chromatograms. Ideally there is one, large, easily identified amino acid. In impure samples and late in a run from even a pure sample, there are many amino acid peaks. To identify the amino acid produced in a sequencing cycle, successive chromatograms are compared to identify which amino acid(s) increase in a cycle.

The chromatographic data collected by our data system is archived on floppy disks. We provide a sheet listing the amino acids released by each cycle of the sequencer, and our interpretation of the sequence. We also compare the sequence with the library of sequences from the National

omedical Foundation if a comparison is desired. We compare the sequences with the fasta ogram of Pearson and Lipman (1988). At the University of Virginia we are fortunate to have e of the authors of this widely used sequence comparison program on campus and a well aintained library of sequences.

locked Proteins

a protein yields no sequence data, it may be blocked, although sometimes samples do not ntain adequate amounts of protein. It has been reported that the majority of soluble proteins ve blocked N-termini (Brown and Roberts, 1976; Driessen, 1985). One group could sequence ve of fourteen proteins from a total cell extract (Hanash *et al.*, 1991). In both blocked and blocked proteins, alanine, serine, methionine and pyroglutamate are overrepresented at the N-rminal. If a large amount of sample is present, it will generate some background amino acids m acid cleavage of internal peptide bonds. To confirm the presence of significant amounts of otein in a sample that did not give data, we can digest the sample with cyanogen bromide after has been applied to the sequencer. This process takes a day, and does not give any interpretable quence data, because there will be several sequences. It will however tell if protein was applied the sequencer (Simpson and Nice, 1984).

you do have a N-terminal blocked protein and you need the N-terminal sequence, you can try blocking although success is not assured. Normally it is more productive to digest the protein d isolate and sequence petpides. One problem is that it is not easy to determine what is locking the N-terminus. If the blocking group is a pyroglutamate residue, also known as 5-xopyrrolidine-2-carboxylic or pyrrolidone carboxylic acid, it may be removed by treatment with yroglutamate aminopeptidase as described by Podell and Abraham (1978). This reaction has lso been used by Strydom *et al.* (1985) and Zalut *et al.* (1980). This reaction can be performed n samples that have been electroblotted on to PVDF (Moyer *et al.*, 1990). If this reaction is onducted, a control is advisable to see if the reaction works. We have used [Phe]2-TRH, whose lution from a C18 column is slightly altered after deblocking with pyroglutamate minopeptidase and an amino acid analysis of the peptide isolated from HPLC will show the loss f the pyroglutamate residue. Because of variability of the commercial preparation of the nzyme, Fischer and Park (1992) recommend determining the optimal activity for each batch of nzyme. These authors have successfully deblocked 100 pmol or less of protein. At high oncentrations of enzyme, they observed cleavage at proline. Similarly Dilone *et al.* (1994) bserved removal of Gly-Pro from a peptide by pyroglutamate aminopeptidase.

Cyclization of terminal glutamine residues occurs slowly, but is accelerated by phosphate and ther buffers, extremes of pH and temperature (Khandke *et al.*, 1989); in 0.1M acetic acid at -20°C, one peptide showed 15% conversion in 7 months. During long purifications, glutamine esidues may cyclize to pyroglutamate (Blombäck, 1967). Enzymes which catalyze the yclization reaction have been described (Fischer and Spiess, 1987; Busby *et al.*, 1987). If reatment with pyroglutamate aminopeptidase has no effect, the blocking group may be an acyl roup.

There are two strategies to obtain sequence data from proteins with N-acetyl groups. One is to digest the protein and sequence peptides; if the N-terminal peptide is isolated, it can be deblocked and sequenced by Edman sequencing or mass spectrometry. The second is to deblock the whole protein. For a review, seen Chin and Wold (1986). Digestion of proteins is described

later, so here are described the two methods for deblocking polypeptides, namely enzymatic and chemical.

The commercially available enzymes for removing N-acetyl amino acids use peptides as substrates, but not large proteins. Pierce Chemical Company sells an acylamino acid releasing enzyme. This enzyme is said to work only on peptides less than 30 residues and releases acyl amino acids or acyl dipeptides ; the suggested protocol uses 100 nmole of protein. Some references to the procedure, including identification of the removed amino acid are Nakamura *et al*, 1974; Jones and Manning, 1985; Tsunasawa and Sakiyama 1984.

The chemical methods used acid. The most recent uses trifluoroaccetic acid and methanol, which reportedly makes this method successful on all amino acids, rather than just N-acetyl serine (although almost all data shown is for N-acetyl serine), and the random acid cleavage of the polypeptide chain is claimed to be small (Bergman *et al.*, 1996). In my hands, this method did not deblock some acetylated peptides, apart from one with N-acetyl serine, and did not deblock cytochrome c, which has a terminal acetyl group. Other methods, which probably are only successful for N-acetyl serine, are heating the peptide with 1 M HCl at 110°C for 10 to 20 minutes which may remove an acetyl group (Fordyce *et al.*, 1979; Chin and Wold, 1985) or add anhydrous TFA to a polypropylene tube and incubate at 40°C for 1 hour (LeGendre *et al.*, 1993) or other acidic conditions (Ozols, 1989). Hulmes *et al.* claim deacetylation by exposure to trifluoroacetic acid vapour at room temperature for about a week; they report cleavage at serine residues also with efficiencies of about 25%. Similarly, Wellner *et al.* (1990) found that limited acid treatment did not deblock cytochrome c, but did deblock peptides with N-acetyl serine or N-acetyl threonine at the N-terminal.

A survey of acetylated proteins found that 43% of N-terminal residues are serine or threonine suggesting that limited acid treatment has significant chances of success (Persson *et al.*, 1985). Examples of the determination of an acetyl blocking group are seen in Haniu *et al.*, (1984); Asano *et al.*, (1986).

Proteins with an N-formyl group at the N-terminus may be partially deblocked by heating for 2 h at 55°C in 25% trifluoroacetic acid; success is variable (Shively *et al.*, 1982). Another procedure is to incubate the sample in a 1.5 ml tube with 30 µl of 0.6M HCl for 24 hours at 25°C as starting conditions (LeGendre *et al.*, 1993). When we tried this procedure on a bacterially expressed protein on which N-formyl methionine was expected, there was only slight acid cleavage, rather than deblocking.

A recent review of obtaining sequence data from acetylated proteins used relatively large amounts of protein and found that the specificity of the enzyme will result in little or no data from some proteins (Krishna, 1992). Uses of this enzyme in real sequencing include Schinina *et al.* (1996). An interesting way to use this enzyme is described by Tsunasawa *et al.*, (1990). The authors digested a protein, blocked the newly released -amino groups with PITC and oxidized to phenylcarbamyl groups, then removed the acetylated amino acid with the enzyme. The mixture of peptides was then applied to the sequencer and only the N-terminal peptide was sequenced. Krishna *et al.* (1991) obtained limited sequence data also using chemically blocking the N-termini of peptides. To obtain data, hundreds of pmole of protein are required; identification of the N-terminal amino acid is possible with large amounts of protein. A similar technique using ion exchange chromatography is described by Allen (1989). A procedure for isolating N-terminal

peptides using adsorption of dinitrophenol derivatized amino groups has been reported (Kaplan and Oda, 1987). Another technique was described by Akiyama *et al.* (1994).

A systematic application of the above techniques to proteins blotted on PVDF was decribed by Hirano *et al.* (1993). The procedure appears relatively insensitive, and may introduce peptide bond cleavages from the attempted acid based deblocking steps which may confuse latter data, especially for large proteins.

Some intracellular proteins have myristic acid on their N-terminal glycine (Sefton and Buss, 1987). The residues which determine whether a terminal glycine is myristoylated are within ten residues of the N-terminal (Towler *et al.*, 1988) and include a lysine at position 7 (Kaplan *et al.*, 1988). A N-Acetyltransferase from yeast displays substrate specificity which is affected by parts of the protein distant from the N-terminal (Lee *et al.*, 1990); the enzyme from rat liver is mainly affected by the first three amino acids (Yamada and Bradshaw, 1991). N-acetylation is believed to be a co-translational event (Yamada and Bradshaw, 1991).

Blocked N-terminal peptides can be selectively isolated from digests by cation exchange chromatography (Gorman and Shiell, 1993).

We suggest that if a protein is blocked, it be digested and the peptides be sequenced, unless obtaining the N-terminus is vital.

Sequencing Proteins From Gels

Many proteins are isolated by gel electrophoresis. To obtain sequence data, the two normal methods are to either transfer to PVDF and sequence, or to digest the protein and sequence peptides. Transferring the protein to PVDF and then attempting sequencing is the more simple approach but success is limited because of the frequency of blocked proteins. Digestion of protein from gels is described separately.

When preparing samples for electrophoresis, boiling may hydrolyze peptide bonds, especially Asp-Pro bonds (Geiger and Clarke, 1987); indeed boiling may be used for peptide mapping (Rittenhouse and Marcus, 1984) although boiling for 3 minutes has been considered to give negligible hydrolysis (Kowit and Maloney, 1982), or according to Rittenhouse and Marcus (1984), about 1%. Tubulin is a protein that is very susceptible to non-enzymatic hydrolysis of peptide bonds, namely Asp-Pro, Gly-Ser; in addition, heating may cause ß-elimination by cysteine, producing dehydroalanine, which may react with lysine, leading to cross links between peptides (Correia *et al.*, 1993). The Asp-Pro hydrolysis appears acid catalyzed, because of the weak buffering of Tris at pH 6.8, and the drop in pH with temperature; phosphate pH 7 gives adequate pH control without raising the pH to where disulfide interchange can become a problem (Cannon-Carlson and Tang, 1997). Kubo (1995) noted that even at pH 6, inorganic cations do not cause hydrolysis, while Davagnino *et al.* (1995) preserved immunoglobulins by using a higher pH Tris buffer. Another danger of excessive heating is loss or oxidation of all reducing agent, with possible reformation of disulfide bonds.

Sample preparation of hydrophobic peptides at low temperatures for electrophoresis has been described recently and is claimed to increase the solubilisation of peptides and reduce streaking (Hennessey and Scarborough, 1989)

Harrington (1990) claim that using diacrylylpiperzaine as the cross-linker in gels increases the yield of sequenceable protein by two. They mention that bisacrylamide breaks down to formaldehyde which may react with amino groups and cause blockage of the N-terminal.

N-terminal blockage is a potential problem but there are no definitive reports on its prevalence; Reim and Speicher (1992) find 10-15% blocking during transfer, which is within the range of normal variability in yield. The following steps may be taken to reduce the chance of N-terminal blockage: use of reducing agent in the solution of protein, pre-electrophoresis of gels to remove potential reactants, degassing of gel solutions to reduce the amount of catalyst needed and hence the concentration of free radicals, use of reducing agent in the electrophoresis buffers, recrystallization of SDS, long polymerization of gels (up to 3 days) to reduce concentrations of potential reactants, using sucrose instead of urea in a sample buffer, heating samples to 37°C for 10 minutes; (Walsh *et al.*, 1988; Moos *et al.*, 1988; Speicher, 1989). It has been claimed that the most important reagent to reduce is ammonium persulphate (A. Louie, Applied Biosystems seminar, 1991). One possible reactant in gels which could lead to blocked N-termini is free acrylamide (Geisthardt and Kruppa, 1987). Data of Bosisio *et al* (1980) suggests that 50 mM acrylamide may be present in a typical gel. Moos *et al.* (1988) recommend conducting electrophoresis at near neutral pH, rather than the usual pH 8.5 and suggest that electrophoresis at lower than normal pH is the most important factor in preventing protein blockage. However it has since been reported that a low pH for electrophoresis improves sequencing yields of some proteins but decreases the sequencing yields of other proteins (Stone *et al.*, 1989a);blocking during electrophoresis The low pH also causes some proteins to smear on gels. Lysine and terminal amino groups can be modified in gels by Schiff base formation, a process reversed at high pH (Kosiarz *et al.*, 1978) which may be why it does not seem to be a problem in sequencing where the sample is exposed to base. Ward *et al.* (1990) suggest that electroblotting causes considerable loss of tryptophan.

Hemoglobin (5μg) in sample wells increases the amount of protein transferable out of a gel on to membranes, possibly by blocking reaction of acrylamide with proteins, thereby incorporating some protein into the gel (Gillespie and Gillespie, 1997). Overnight incubation dcreased protein yield rather than increasing it, and pre-electrophoresis had no benefit. This report agreed with an earlier finding that keratin, a common contaminat of gels, is partially solubilized by reducing agents, causing streaks and giving the appearance that 2-mercaptoethanol is a source of contaminats (Ochs, 1983).

Pre-running of gels may decrease resolution, but casting the stacking gel with resolving gel buffer and performing pre-running with stacking gel buffer in the upper tank maintains resolution (Dunbar and Wilson, 1994)

Another possible reaction that occurs in acrylamide gels is alkylation of cysteine by free acrylamide. This phenomenon has been observed by several laboratories, some of which have used the cysteinyl-S-propionamide adduct to identify cysteine (Ward *et al.*, 1990, Ploug *et al.*, 1989). However only some cysteine residues may react (Chiari *et al.*, 1992). Controlled alkylation prior to electrophoresis is preferred.

A double, inverse gradient gel system may increase resolution of proteins of similar sizes (Zardoya *et al.*, 1994). To separate large proteins by electrophoresis, FMC recommend their agarose gels, which can separate proteins up to 600,000 (FMC, 1991).

To separate small peptides as peptides, small as 1,000 daltons, Schägger and von Jagow (1987) suggest electrophoresis using Tricine buffers. Novex (Encinitas, CA), Schleicher and Schuell, Integrated Separation Systems (Natick, MA) and the Nest Group sell precast gels that are claimed to separate peptides as small as 1,500. The Nest Group claim that their gels have longer shelf life and better resolution, in part due to the use of high quality SDS than gels made in the normal laboratory. Integrated Separation Systems also sell a stain that is claimed to efficiently fix and stain small peptides, but the stain is not stated to be compatible with further analysis of the peptides. Hoefer list an electrophoresis system for peptides down to 1,500 in their 1992 catalogue.

Elution of proteins from gels

Below are methods for removing proteins from gels, but these methods have limited uses. Normally to obtain sequence data from an intact protein, transfer to PVDF is the method of choice. If the protein is blocked, the normal method is to digest in the gel and then elute, separate and analyze peptides.

There are devices for direct elution of protein from gels. ISCO, BioRad, Hoefer, Schleicher and Schuell, and Amicon make such apparatus, and another has been described by Hunkapiller *et al.* (1983). The Amicon unit claims recoveries of 80% to 95%, with unfixed proteins; they suggest staining with cold KCl, sodium acetate or Coomassie without acid. Promega claim that proteins stained during electrophoresis with their ChromaPhor Protein Visualization System yield 75% recovery. One report that stains with 4M sodium acetate finds 70% recoveries when 100 µg of protein is applied to a gel (Ohhashi *et al.*, 1992).

One procedure for passive elution of proteins from gels is that of Ward *et al.* (1991), in which the gel slice is soaked in 3 x 1 ml of 20 mM Tris-HCl pH 7.4, 0.02% Tween 20 for 20 hours at 25oC. Coomassie ,effect on elution of proteins stained BSA was poorly eluted from a gel (1%) whereas 90% of unstained BSA was passively eluted after 70 hours (Ward *et al.*, 1991). To remove Coomassie, staining was only for 15 minutes when possible, then the gel pieces were washed with 3% SDS, 50% 2-propanol to remove Coomassie, then washed with water for 24 hours. The SDS is to dissociate the Coomassie and the propanol keeps the protein fixed (Ward, personal communication).

Promega sell a Protein Recovery System to be used in conjunction with their Protein Visualization System which stains proteins during electrophoresis (Promega Technical Bulletin 100). Promega claims recovery of 50% or better for proteins with molecular weights of 31, 55, 97 Kd (Larson and Shultz, 1993). The Recovery System uses homogenization of the gel followed by extraction with 100 µl of 50 mM ammonium bicarbonate, 0.1% SDS. The supernatant from centrifuging the crushed gel contains the protein. After an essential concentration step, the SDS and dye are removed by acetone precipitation. The dye inhibits some proteinases. The system appears to be based on earlier reports of passive elution of proteins from gels (Hager and Burgess, 1980; Bray and Brownlee, 1973).

Another method of from gels is to soak the gel slice, containing protein visualized with 1 M KCl, in 1 ml of 70% formic acid for 10 minutes and then centrifuge the mixture on to a piece of PVDF. The ProSpin cartridges sold by Applied Biosystems appear ideal for this method.

Recoveries of protein from electrophoresis step at 10-20%. The method has been used with proteins from 13,000 to 67,000 (Hermansen *et al.* 1992).

Another method for eluting proteins from gels is to add 200 µl of 50% trifluoroacetic or formic acid, crush the gel with a Kel-F pestle and sonicate; recoveries decrease with the protein size (Asquith *et al.*, 1992)

Feick and Shiozawa (1990) report that elution of proteins from gels with formic acid:acetonitrile:2-propanol:water 50:25:15:10 v/v/v/ followed by gel filtration gives recoveries of over 60% in most cases.

Unstained proteins can be detected in gels by phase contrast (Johnson *et al.*, 1990) but the commercial instrument has been withdrawn from the market.

Zinc or copper staining is claimed to be almost as sensitive as silver staining and reversible, so that the stain does not interfere with blotting or sequencing (Fernandez-Patron *et al.*, 1995a; Ortiz *et al.*, 1992). When treated with the right solutions, proteins can be digested and blotted with minimal loss compared to unstained proteins, even after storage for years. Other stains are the Rapid Reversible Stain of Diversified Biotech which is claimed not to affect biological properties of proteins and to be removable for elution of the protein from gels. Nile Red stains proteins in SDS gels, and with modifications, in IEF gels, without interfering with subsequent operations (Bermudez *et al.*, 1994). Fluorescence detects 10 ng of protein (Alba *et al.*, 1996). ESA (1997) sell Rev-Pro reversible stain for PVDF and nitrocellulose membranes; it is claimed to be more sensitive than Coomassie Blue, and compatible with sequencing.

Proteins can be removed from gel slices by homogenization and spinning in an Amicon Microcon concentrator fitted with a Micropure insert for filtration (Amicon document 1901; Sheer, 1994). In further studies, Amicon (Krowczynska *et al.*, 1995) showed that unstained gels give the best recovery, or using KCl stained gels while Coomassie or zinc imidazole staining reduces recovery. The elution buffer is less important, with best results from 50 mM Tris-Cl, 0.1 mM EDTA, 150 mM NaCl or PBS.

An interesting method is reverse staining of proteins with zinc, followed by electroelution on to a small reverse phase column, from which the protein is eluted with a gradient of acetonitrile (Fernandez-Patron *et al.*, 1995b). This method may separate chromatographically proteins which are not resolved by electrophoresis. However it requires construction of special apparatus and the best sensitivity reported was 50 pmole of protein.

Removal of gel matrix

Another method of recovering proteins that are purified by electrophoresis is to use FMC's agarose based ProSieve electrophoresis system, in which the agarose gel is melted to allow extraction of the protein. Recoveries of 90% are reported (Morgan *et al.*, 1991). Bands are not as sharp as acrylamide gels. Epicentre Technologies (1202 Ann Street, Madison, WI 53713, 800-284-8474) sell GELase™ to digest agarose but we suggest that the absence of proteolytic enzymes be determined before using this product.

BioRad sells the reversible cross linker N,N'-bis-acrylylcystamine for acrylamide. Although this product has been available for several years, it has not been widely used. One reason is that

Achieving a completely soluble gel may require elevated temperatures during polymerization, purification of the reversible cross linker (Hansen, 1981) and the use of formamide to avoid non-reducible cross links (Hansen *et al.*, 1980). This cross linker is less reactive than bis-acrylamide and may give gels that do not perform well (Gelfi and Righetti, 1981). Ghaffari *et al.* (1988) obtained satisfactory resolution, although less than from bis-acrylamide and 50-70% recovery of proteins at a 1 μg load.

Sample concentration from multiple gels

To analyze proteins separated on 2-D gels, Bauw *et al.* (1990) cut out bands from multiple gels and loaded up to 20 gel segments on top of another gel which concentrated the protein. The second gel has a custom made well which concentrates the sample, which can be gel pieces or dilute protein solution, horizontally to give a band 5 mm2 (Rasmussen *et al.*, 1991).

Another technique needs only a different spacer to form a funnel web well. The data shows excellent recovery from multiple gel slices (Lombard-Platet and Jalinot, 1993). A variation runs the protein from multiple pieces of gel into an agarose gel, from which the protein is extracted by melting (Rider *et al.*, 1995).

Another gel concentration technique uses a concentration gel inside a Pasteur pipet, leaving the protein in 1-2 μl of gel (Gevaert *et al.*, 1996).

Proteins from multiple gels can be concentrated on one piece of PVDF by homogenizing the gel and spinning the extract on to PVDF (Warlow *et al.*, 1995). The protein can then be sequenced directly or digested on PVDF.

Blotting to PVDF

The standard procedure now for sequence analysis of proteins separated by electrophoresis is to blot on to PVDF,(polyvinylidene difluoride) membranes (Immobilon-P from Millipore) as described by (Matsudaira, 1987; LeGendre and Matsudaira, 1988). Nitrocellulose cannot be used in the sequenator because of its susceptibility to the harsh conditions employed.

We suggest amount of sample loading 100 pmoles on a gel to increase the chances of having sufficient sequenceable protein present after losses during electrophoresis and transfer. One simple rule for proteins with molecular weights up to 30,000 is that if a stained blot can be photocopied, it can be sequenced. After blotting, it is wise to stain the gel to see if all protein transferred from the gel. Also a second piece of PVDF will indicate how much protein passes through the primary membrane. The amount of protein passing through PVDF is less with some brands (see Selection of PVDF) and may be reduced by presoaking the gel in transfer buffer (Mozdzanowski and Speicher, 1990).

It is reported that the amount of protein on PVDF can be quantitated, down to 0.5 μg, by eluting the bound Coomassie Blue (Kain and Henry, 1990). One laboratory has quantified immunostained and Coomassie stained protein on PVDF by wetting the membrane with methanol to make it transparent and then scanning (Parrado *et al.*, 1993). quantitation of protein In our laboratory, Dr. E. Baramova loaded about 1.3 nmole of a protein on a gel and was able to sequence over 25 amino acids.

Amino acid analysis of proteins electroblotted to PVDF shows typical recoveries of 28%-47% (Tous *et al.*, 1989, Nakagawa and Fukuda, 1989). Stone and Williams (1993) report an average blotting efficiency of 34% for proteins between 14,000 and 116,000. Henzel *et al.* (1994b) report that thin gels increase transfer efficiencies, especially for larger proteins; 0.5 mm thick gels give transfer efficiencies of 60-90%.

Transverse gradient gel electrophoresis can determine the optimal concentration of for electroblotting, by showing resolution and transfer efficiency on one gel (Smejkal and Gallagher, 1994). As a guide, these authors show complete transfer of proteins up to 45,000 with gels up to 12% acrylamide, up to 60,000 with a 7% gel and up to 120,000 with a 6% gel.

Although not tested with PVDF, ultrasound has been used to transfer proteins to nitrocellulose (Kost *et al.*, 1994).

Novex in their catalogue suggest Tricine gels because they are more porous than the corresponding Tris containing gels and therefore may give better transfer.

A new approach to efficient transfer of proteins is the use of different buffers on each side of the PVDF transfer membrane (Laurière, 1995). On the gel side, high pH and SDS promotes elution of the protein from the gel, which on the other side of the PVDF, low pH and methanol promote binding to the PVDF. Under these conditions, large and small proteins transfer efficiently to the membrane.

Selection of PVDF

The PVDF that is best known is Immobilon-P from Millipore. PVDF is also available from Applied BioSystems (ProBlott), BioRad and Schleicher & Schuell (Westran); these companies claim that their products are superior to Immobilon in retaining proteins. Westran is PVDF on a polyester support. In support of the claims of Applied Biosystems, Promega in their Probe-Design kit for generating and separating peptides, state that peptides are more easily eluted from Immobilon and better retained on ProBlott. Millipore has released a second version of Immobilon, the Immobilon-PSQ membrane which they claim binds more proteins and gives much higher yields for . In 1992 Porton Instruments released Hyperbond™ membrane which is claimed to retain small peptides better than existing high density PVDF and to allow more efficient sequencing reactions with more data obtained as a result.

In our laboratory, we have compared three brands of PVDF. Ken Klotz, Department of Cell Biology, transferred 50 pmole of ß-lactoglobulin to the PVDF. ProBlott from Applied Biosystems gave twice as much signal as Immobilon and slightly higher repetitive yields; Westran also retained more sequenceable protein than Immobilon but less than ProBlott. The experiment described here used only one protein and one condition for transfer, (10 mM CAPS, pH 11, 10% methanol, 500 mA, 40 min from 1.5 mm thick 10% gel) so that the advantage of ProBlott may not always be as great under other conditions; also it should be remembered that many people have successfully obtained data with Immobilon. It is claimed that three brands of PVDF selection of PVDF perform similarly when excess SDS is removed but ProBlott and Bio-Rad PVDF are superior when excess SDS is present (Speicher *et al.*, 1990). Later this lab showed that ProBlott and Bio-Rad's Transblot bind small proteins better than Immobilon-P, and that Immobilon-P is more sensitive to SDS (Mozdzanowski and Speicher, 1992a). PVDF which

binds proteins more tightly gives sequencing performance as good as looser binding PVDF, and better yields of tryptophan (Reim and Speicher, 1992).

Jungblut *et al.* (1990) reported only small differences in efficiency of transfer of a range of proteins on to ProBlott™, Immobilon™, a polypropylene membrane (Selex-20) and a siliconized glass fibre (Glassybond). When the increased initial yields obtained with the siliconized glass fibre are considered (Eckerskorn and Lottspeich, 1990), the authors suggest the siliconized glass fibre is the best medium for electroblotting and sequencing of proteins. However this membrane is not in widespread use.

Another study of transfer membranes showed that ProBlott usually binds more protein that the other tested membranes. When the efficiency of transfer and sequencing are considered, ProBlott appears to usually be the best choice of membrane (Baker *et al.*, 1991). Immobilon PSQ was not included in this study. A Millipore representative suggests that because both Immobilon PSQ and ProBlott have 0.1 μm pores, rather that 0.45 μm of the original Immobilon, they will bind protein in a similar manner. Micron Separations (MSI) sell PVDF-Plus, with 0.45 μm pores, which they say is comparable to Immobilon-P. LeGendre *et al.* (1993) report recoveries of 30-80% of protein on Immobilon-P and 100% on Immobilon-Psq and ProBlott.

Buffers for electroblotting

Some studies claim that there is not a systematic difference between CAPS and Tris-glycine buffers (Speicher, 1989; Stone *et al.*, 1989a), although some proteins exhibited less efficient transfer out of gels, or less efficient binding on PVDF membranes in one of the buffers. More recently Mozdzanowski and Speicher (1990) claim that Tris-glycine and Tris-borate buffers (50 mM Tris, 50 mM boric acid, not titrated, pH ~8.3, 10% methanol, 225 mA, 2-4h) give 40-50% recovery of blotted proteins on the first PVDF membrane compared to about 30% recovery obtained with CAPS buffer and sodium borate. Tris buffers must be removed to prevent interference with sequencing. Baker *et al.* (1991) also suggest that Tris-borate pH 8.5 is often superior to CAPS. Speicher (1989) used 12.5 mM Tris, 96 mM glycine pH 8.3 for 2 to 4 hours as his routine conditions which he claimed gave quantitative transfer of most proteins, the exceptions being those in the top 15% of the gel (a range of acrylamide concentrations were used; the concentration used which gave the reduced recovery was not identified) where recovery was still 75% or more.

Bauw *et al.* (1990) report that Tris-borate (50 mM Tris, 50 mM borate, 8 h, 35 V) gives better recovery of proteins than Tris-glycine although transfer is slower.

Bhavsar *et al.* (1994) find that addition of BSA as a carrier protein improves efficiency of transfer. Addition of BSA would interfere with sequencing applications, but deoxycholate may improve transfer efficiency, or decrease it.

Transfer time for electroblotting

Speicher (1989) suggests that transfer times need to be empirically optimized but CAPS needs shorter times than Tris-glycine buffers. Later this author (Mozdzanowski and Speicher, 1990) suggests that 2 hours of blotting is sufficient and that longer transfer times do not remove proteins that remain in the gel, nor do they move proteins from the first PVDF membrane to the second, a finding confirmed by Jungblut *et al.* (1990). Mozdzanowski *et al.* (1992) showed that

overtransfer does not occur i.e. once a protein has been bound to PVDF, it does not elute during the transfer.

A study by Baker *et al.* (1991) shows that a large protein does not transfer as fast as a small protein. With some membranes, there was loss of protein in less than an hour of transferring but no data after one hour was shown.

Effect of methanol and SDS on electroblotting

Gültekin and Heerman, 1988 report that in the transfer buffer prevents binding of proteins to PVDF; methanol is reported to reduce the elution of some proteins from gels (Eckerskorn *et al.*, 1988). Mozdzanowski and Speicher (1990) suggest 10% methanol, because less methanol decreases binding to PVDF; he also reports that SDS usually decreases binding to PVDF. A few membrane proteins, transfer from gels were transferred more efficiently without methanol; apparently methanol was not needed to dissociate SDS from the protein. Mozdzanowski and Speicher, 1990) claim that most losses losses of protein associated with electrophoresis and blotting occur during transfer and are independent of the buffer used but are caused by high concentrations of (Speicher *et al.*, 1990). To avoid the losses, they recommend the presoaking of the gel in transfer buffer. For some proteins below 20,000, presoaking the gel increases the amount of blotted protein by 70%, even though less protein is transferred out of the gel (Mozdzanowski *et al.*, 1992). The amount of protein bound to PVDF is about 50% of that in the gel. Eckerskorn *et al.* (1988) also report that methanol decreases the movement of large proteins out of gels. Speicher explains that methanol is needed to dissociate SDS from the protein; too little methanol will cause incomplete SDS removal and hence poor binding to PVDF, whereas too much methanol will remove SDS from the protein in the gel. Because the charge on the SDS converts the electric field into movement of the protein, removal of SDS in the gel by high concentrations of methanol will result in inefficient transfer of the protein from the gel. Guidelines from Applied Biosystems are to use 20% methanol for transferring protein below 20,000, 10% methanol for proteins between 20,000 and 80,000 and no methanol for larger proteins (Applied Biosystems seminar, 1991). For large proteins, it may be necessary to add to the transfer buffer to help move the proteins out of the gel.

in the transfer buffer reduces the amount of protein bound by PVDF several fold, but addition of 0.1% NaCl partially restores binding efficiency (Tovey and Baldo, 1989).

Transfer of large proteins

Large proteins are the most difficult to transfer and sequence because of low yields on transfer and then relatively high backgrounds of amino acids during sequencing. Some factors which may increase the success of electroblotting are listed here. Besides reducing the methanol concentration, or even omitting methanol (Pluskal *et al.*, 1986), another technique that may improve extraction of proteins from the gel is to presoak the gel in transfer buffer for 30 minutes (Mozdzanowski and Speicher, 1990). Increasing current density may increase efficiency of transfer of large proteins at the expense of smaller proteins (Jungblut *et al.*, 1990). A blotting procedure varying the current is claimed to give good recoveries of both large and small proteins (Otter *et al.*, 1987).

ing from a gel with a low concentration of may also help and is more successful than eing methanol concentration in the transfer buffer (Mozdzanowski *et al.*, 1992). Adding % to the transfer buffer may help.

he Max Planck Institute in Munich, one recommended blotting procedure is to use 10 mM um borate (no adjustment of the pH) overnight in the cold room with cold circulating water 00 mA.

nsfer of small peptides

lipore report that proteins less than 20,000 transfer easily and are insensitive to methanol centration. Because they can diffuse, equilibration of the gel in transfer buffer should be less n ten minutes. Porton Instruments claim that their Hyperbond ™ membrane is more efficient n existing PVDF membranes at binding small proteins. To enhance the binding of small tides to PVDF, Millipore suggest using 20% methanol in the transfer buffer and reducing the ctric field by 50%; a procedure is described by Otter *et al.* (1987). They also recommend their w product, Immobilon-PSQ.

ing semi-dry blotting apparatus

is type of apparatus is less efficient in transferring proteins than a tank transfer apparatus W. Fox, unpublished observations; reported in poster of Mozdzanowski and Speicher, 1989)

study using a Millipore semidry blotting apparatus, found that the amount of membrane roteins extracted from a gel was inversely related to the molecular weight of the protein (above 2,000). Addition of 0.1% SDS to the cathode buffer increased the transfer of larger proteins and 5% or 40% methanol in the anode buffer improved the retention of proteins less than 12,000 on he PVDF. However the high methanol concentrations decreased the amount of larger proteins ound to the PVDF. These modifications are not applicable to transfer in a tank type apparatus Lissilour and Godinot, 1990). Parameters affecting transfer efficiency in a semi-dry apparatus ave also been studied by Jungblut *et al.* (1990).

Other procedures for electroblotting

Millipore has a book of electroblotting protocols for different applications, *Protein Blotting Protocols for the Immobilon -P Transfer Membrane.*

One of the protocols that may be useful for some situations is for electroblotting of large proteins from non-denaturing gels; this procedure may be useful because of the difficulty of transferring large proteins. No methanol is used in the transfer buffer; data on the amount of protein transferred is not shown. A procedure for transfer of basic proteins from non-denaturing gels to nitrocellulose has also been published (Van-Seuningen and Davril, 1990)

Non-denaturing gels can be coupled with SDS-PAGE; two denaturing gels may be used to concentrate the sample (Trudel and Asselin, 1994).

Xu and Shively (1988) recommend transfer times of 100 minutes, at a lower voltage than specified by Matsuduira (1987) and report decreases in recovered protein if the electroblotting is continued. These authors also claim that coating the Immobilon-P membranes with Polybrene

increases the yield of transferred protein, as reported by Matsuduira (1987) and Walsh *et al.* (1988), but no longer recommend the use of Polybrene (conference report).

Amount of PVDF used in electroblotting

PVDF is very hydrophobic and poorly wetted by the butyl chloride used in the sequencer to extract ATZ-amino acids from the PVDF which binds the protein being sequenced. It has been found that putting a large, or very small amount of PVDF in the sequencer is more likely to cause variable flow of butyl chloride over the PVDF membrane, causing variation in the amounts of amino acid seen (Speicher, 1989); this reported variation is consistent with our observations. The suggested ideal size for PVDF membrane is 4 x 9 mm, with two pieces this size; larger amounts can be loaded but are prone to erratic flows of reagents and hence erratic levels of amino acids. Another piece of data on the amount of PVDF is that of Stone *et al.* (1989a) who found that loading protein on a large amount of PVDF decreased the repetitive yield during sequencing.

Staining of electroblotted proteins

Destaining is also reported to affect the yield of sequenceable protein (Speicher, 1989); destaining with 50% methanol/10% acetic acid was reported to reduce yields of sequenceable protein by 30% or more compared to destaining with 50% methanol alone. Phang *et al.* (1996) report that omission of acetic acid from staining and destaining solutions slows the processes but does not affect the final results. The staining procedure suggested is to stain with 0.5% Coomassie blue in 40% methanol and destain in 50% methanol (Speicher, 1989) (time unspecified but possibly five minutes and not more than half an hour). This author has since reported that Coomassie may reduce the yield of sequenceable protein, although he was not working with electroblotted proteins, but rather proteins adsorbed on PVDF; Ponceau S was suggested as a stain which causes less loss of protein (Purcell & Speicher, 1989). Choli *et al.* (1989) report that 40% of lactoglobulin spotted on to PVDF is sequenced, but only 24% of that which is stained with Coomassie; Ponceau S was not tested. Drying the blot after staining enhances visibility of proteins (Sanchez *et al.*, 1992).

Another destaining procedure is to soak the gel in 5% acetic acid:1-butanol 9:1 biphasic mixture; this system is claimed to be more rapid than conventional destaining (Molnar *et al.*, 1990).

Ponceau S is less sensitive than Coomassie. Staining is stronger at acidic pH and is reversible at basic pH (Aebersold, 1989). Schleicher & Schuell claim that destaining in 10% 2-propanol/10% acetic acid gives permanent staining (instructions for Westran).

Promega have three protein staining systems. They claim that their ChromaPhor Protein Visualization System, which stains proteins during electrophoresis, is compatible with blotting and sequencing but they do not show data on recoveries (Breitlow, 1992). This company also reports loss of small and hydrophilic proteins from PVDF during staining and destaining with Coomassie blue. The suggested conditions are to stain for 30 s in 40% methanol, 5% acetic acid, 0.02% Coomassie blue and destain for 1 min in 40% methanol, 5% acetic acid. They do not address the issue of loss of sequenceable protein caused by Coomassie blue or acetic acid (Shultz, 1992).

In addition to staining proteins after transfer, Thompson and Larson (1992) successfully transferred proteins stained with Coomassie Blue or ChromaPhor stain. By soaking the stained gel in 1% SDS, 50 mM Tris-Cl pH 7.5, these workers presumably coated the protein with SDS so that an electric field would move the protein out of the gel, even though it had been fixed. After soaking in SDS, the gel was soaked in transfer buffer for 15 minutes.

Copper phthalocyanine 3,4',4",4"'-tetrasulphonic acid has been proposed as more sensitive than Coommassie blue, easily removed with 0.5 M NaHCO3/20% ethanol (Bickar and Reid, 1992). Another chelating method uses a Ferrozine/ferrous complex or ferrocyanide/ferric complex for protein on PVDF or nitrocellulose (Patton *et al.*, 1994). The iron based stains are as sensitive as gold stains, but are reversible and after reversal, cause no interference with sequencing or digestion.

An alternative to staining is to dry PVDF for 10 minutes after blotting, rewet in 20% methanol and look at it in front of white light or place in 20% methanol over a light box. Protein bands will appear more translucent than background (LeGendre and Matsudaira, 1989). By wetting the PVDF in 50% methanol and scanning at an unspecified, protein bands can be quantitated to 1 μg (Parrado *et al.*, 1993).

Merewether *et al.* (1995) used Coomassie Brilliant Blue G stain instead of R, because there was only one, late eluting peak. Amido Black gave negative deflections in the baseline, interfering with low level peptide collection.

Blotting to other supports

A procedure to transfer proteins from polyacrylamide gels to glass fiber (Aebersold *et al.* 1986) does not appear to be used any more.

Alimi *et al.* (1993) transferred proteins to carboxymethylcellulose membranes purchased from Schleicher and Schuell (NA-49). Most of the proteins studied were basic but alcohol dehydrogenase, serum albumin, ovalbumin and catalase also transferred to the membranes with "high yields". A major advantage of CMC membranes is ease of eluting proteins, with 50 mM HCl

Proteins transferred to nitrocellulose cannot be loaded into a sequenator because the membrane disintegrates under the conditions used. Although some prefer to use nitrocellulose, (Aebersold *et al.*, 1987; Aebersold, 1989), PVDF appears more popular.

When proteins are blotted on to nitrocellulose, up to 90% of small proteins can be eluted with acetonitrile solutions (Parekh, *et al.*, 1985). Triton X-100, at 0.1% to 1% also may elute proteins; Polyethylene glycol under low salt conditions is a second choice. (Gelman Sciences, usenet message, 1994) . Lui *et al.* (1996) investigated elution of proteins from PVDF, and found that high temperatures or high TFA partially degrade proteins; alkaline solutions of acetonitrile are moderately successful; 1% piperidine in 40% acetonitrile is up to 90% effective. SDS and some other ionic detergents were ineffective, whereas several non-ionic detergents werw effective, and the favoured detergent was

3-16 (3%); adding organic solvent to detergents did not enhance elution. Zwittergent 3-16 gave greater recoveries (up to 90%) of peptides from digestion of proteins on nitrocellulose. Staining

proteins reduces elution efficiency, so the authors suggest staining a narrow strip of nitrocellulose to locate the band of nitrocellulose to be eluted. Aebersold (1989) suggests that extensive drying of nitrocellulose causes irreversible binding of proteins to nitrocellulose and that desorption is favoured at basic pH. Another procedure for removing proteins from nitrocellulose is to dissolve the membrane in acetone and precipitate the proteins with ammonium bicarbonate (Anderson, 1985).

A theoretical study with observations of the binding of proteins to nitrocellulose shows that salt and 20% methanol promote binding of proteins (van Oss *et al.*, 1987).

Electroblotting from IEF and 2-D gels

It has been reported that proteins can be blotted from isoelectric focusing gels after removal of ampholytes by perchloric acid (Hsieh *et al.*, 1988). Proteins have also been blotted from immobilized pH gradients (Knierim *et al.*, 1988). A recent two dimensional electrophoresis system, using acid urea gels reportedly can handle relatively large amounts of protein, and separation by hydrophobic properties can follow using a third electrophoresis (Vanfleteren, 1989). A two dimensional mapping system using acid urea minigels has been described earlier (Davie, 1985). Blotting from two dimensional gels on to derivatized glass fibre has been reported (Bauw *et al.*, 1987).

Proteins separated on two-dimensional gels have been analyzed by Kennedy *et al.* (1988a, 1988b) by performing digestions of the protein in the gel, then separating fragments by electrophoresis, transferring to PVDF and sequencing. It is claimed that data was obtained from 1-10 μg of protein, which may require pooling of material from several gels. To analyze proteins separated on 2-D gels, Bauw *et al.* (1990) cut out bands from multiple gels and loaded up to 20 gel segments on top of another gel which concentrated the protein. Jahnen *et al.* (1990) also describe sequencing of proteins separated on 2-D gels. Immobilized pH gradients have been used; ten gels give sufficient amounts of the more abundant proteins to enable sequencing (Hanash *et al.*, 1991).

Hochstrasser *et al.* (1988) claim that a modified 2-D electrophoresis system increases resolution over O'Farrel's original system.

Sequencing peptides produced by digestion of a protein.

Often a protein must be digested to peptides to obtain sequence data, most often because of a blocked N-terminal, and also because sequencing from the N-terminal may not give enough data. The steps of digestion and separation of peptides cause some sample loss, so more protein is needed than when sequencing the N-terminal of a protein.

Choice of digestion.

There are enzymatic and chemical methods for digesting proteins, but enzymatic digestions are more common. An ideal digestion cuts only at a specific amino acid, but cuts at all occurences of that amino acid. The number of digestion sites should not produce too many peptides because separation of peptides becomes too difficult. On the other can, too few digestions produces peptides too large for complete analysis.

The most common digestions are with trypsin and lysine specific proteinases, because these enzymes are reliable, specific and produce a suitable number of peptides. The next most common digestion is at aspartate or glutamate using endoproteinase Glu-C or endoproteinase Asp-N. Chymotrypsin is sometimes used, although it does not have a well defined specificity. Proteinases of broad speficity may generate too many peptides to separate, and the peptides may be too short to give useful data. Of the chemical cleavages, cyanogen bromide is the most common. All the chemical digestions are less efficient that a good enzymatic digest. However they do produce only a few peptides, which can ease the purification problem.

Sequencing a few large peptides may be more efficient than sequencing several smaller peptides. With a long peptide, there is a better chance of having a long sequence stretch for primer design, and there are less gaps between peptides. Also the actual sequencing can be faster because of the elimination of the startup steps of the sequencer operation.

Often standard analytical reagents are used for the digestions described above. If there is concern that the proteinase being used is contaminated with proteinases of different specificity, sequencing grade proteinases are available from Boehringer-Mannheim and Promega. It has been reported that some proteinases are inactive; one group prefers sequencing grade trypsin and endoproteinase Asp-N from Boehringer, endoproteinase Lys-C from Wako Chemicals, and chymotrypsin and subtilisin from virtually any supplier (Tempst *et al.*, 1990). These enzymes may be stored frozen in 0.1 M ammonium bicarbonate. We have not noted inactive proteinases.

Digestion conditions

Digestions with some proteinases may be performed in 20% organic solvent, effect on digestion which appears to improve the specificity of thermolysin (Welinder, 1988). Fernandez *et al.* (1992) find that 10% acetonitrile with or without 1% reduced Triton X-100, use in digestions allows tryptic digestion of proteins. This group also showed that carboxymethylation of a protein allows more complete .

Overnight digestions are common, to ensure that a protein is fully digested. However, such times may be longer than necessary, although only if there is sufficient protein for multiple conditions can the necessary digestion time be investigated. When subtilisin was denatured, with a trypsin:substrate ratio of 1:100, 10 minutes was sufficient for digestion, and longer digestions gave a more complex peptide pattern, possibly due to the formation and then the action of pseudotrypsin (Christianson and Paech, 1994).

Use of controls during digestion of a protein

When performing an enzymatic digestion, it is very important to perform a control digestion containing no substrate; also include a positive control, like cytochrome c. We have more than once sequenced fragments of the proteinases used to perform a digestion. Using proteinases of known sequence -trypsin, endoproteinase Glu-C, chymotrypsin, thermolysin- makes finding autolytic fragments easier.

It is also vital to ensure that the protein of interest is digested. We have seen incubations of proteins with proteinases in which no significant digestion occurs. After performing a digestion, some of the digested protein should be run on a gel to see if there is intact protein left, or if the digestion products were separated by chromatography, some of the intact protein should be run

to see that it is not present in the digestion mixture. Running a control digestion of the proteinase alone is not sufficient to be sure that there are no products from the proteinase.

To digest proteinase resistant proteins, Riviere *et al.* (1991) suggest either heating for 30 min at 50oC in 6 M guanidine-HCl, which can be diluted to 2 M; chymotrypsin, *Achromobacter* protease or subtilisin can be used for digestion but trypsin works poorly. Another procedure is heating at 37°C for 30 min in 8 M urea followed by digestion with *Achromobacter* protease or subtilisin.

Urea can form cyanates which will react with amino groups. Boehringer Mannheim recommend 20 mM methylamine to block this reaction. 0.1M ammonium bicarbonate also is adequate to prevent N-terminal modification (Stone and Williams, 1993: Fernandez *et al.*, 1992) Proteolysis may be performed in detergents, although may cause problems because it must be removed prior to HPLC separation and cannot be tolerated by all proteinases (Stone *et al.*, 1989c). Proteolysis in acetonitrile also worked well.

When preparing proteins for digestion, excess detergent should be removed to prevent inhibition of enzymatic digestion and interference with reverse phase separation of peptides (Shannon, *unpublished observations*, Stone *et al.*, 1989b). However, low levels of Tween 20 have been suggested to reduce losses of peptide during digestions (Simpson *et al.*, 1989a,b). This grouphas abandoned use of Tween 20 for digestions, possibly because it interferes with with mass spectrometry; they have not said anything about sample losses (Reid *et al.*, 1995).

A complication to proteolytic digestions is the formation of new peptide bonds by the proteinase under sub-optimal pH. transpeptidation One study found 10% formation of a new peptide bond with trypsin, and 1% levels of products using endoproteinase Glu-C (Canova-Davis *et al.*, 1991).

To separate peptides produced by a digestion of a protein, the most common method is reverse phase HPLC. When performing such separations, it is wise to chromatograph peptides in two solvent systems to reduce the chances of having two or more peptides in one peak. When collecting peptides, Aebersold (1989) recommends immediate freezing to avoid reduced yields after serine residues.

Moritz *et al.* (1996) state that broad peaks in the chromatogram of peptides from a digest are disulfide linked peptides, which can be avoided by reducing and alkylating the protein before digestion, which also gave more peptides, in greater quantity.

For digests of less than 50 pmole, Kenny *et al.* (1994) suggest that electrophoresis on 150 µm diameter capillaries gives better recovery than chromatography on 2.1 mm diameter columns. However use of electrophoresis for separation of peptides from digests has not been widely used.

Two dimensional electrophoresis can separate peptides also; its main use is separation of fragments which are not eluted from reverse phase columns (Nokihara *et al.*, 1994).

One scheme to reduce sample handling when obtaining proteolytic fragments of a protein is to load the protein on a reverse phase column to concentrate and desalt it, and possibly alkylate it. Proteinase introduced into the column digests the protein and the fragments can then be eluted and separated. This procedure has been described for the Hewlett-Packard sequencer with its sample cartridge. 20% acetonitrile is essential to obtain digestion, causing hydrophilic peptides

to be washed off the support with starting buffer. Some bonds are not cleaved with the protein on the column although they are cleaved in solution (Burkhardt, 1993). A similar technique is used by the laboratory of Don Hunt to generate peptides for mass spectrometry.

An example of successfully obtaining data from a protein is work done by Joel Hockensmith and Larry Mesner in the Department of Biochemistry. Intact ATPase, molecular weight 90,000 was blotted on PVDF but neither the intact protein nor proteolysis products gave sequence data. 500 pmole of protein was isolated, digested with CNBr and peptides separated on Tricine gels and blotted on to PVDF. Approximately 20 pmole of peptide was sequenced off PVDF. Two samples contained more than one sequence and one had only one peptide which was sequenced for over 25 amino acids.

Jones *et al.* (1994) optimized digestion of one protein by altering temperature and urea concentration, thereby increasing yields of peptides and reducing non-specific digestion. Another digestion with endoproteinase Glu-C gave either incomplete digestion, or artifacts of non-specific cleavage and transpeptidation (Hara *et al.* 1996).

Michrom BioResources sell small columns which trap protein while reduction, alkylation and digestion are performed (Baldwin *et al.*, 1995). The procedure is confusing as written, and recoveries are not given. Using a conventional C18 cartridge for alkylation of lysozyme followed by elution gave disappointing yields in my hands.

Immobilized proteinasescan be used for protein digestion. Davis *et al.* (1995) have described the production and use of capillary columns of trypsin for preparation of samples for mass spectrometry. Raising the temperature of digestion to 37°C increases autolysis, as does the presence of acetonitrile. Ronnenberg *et al.* (1994) immobilized proteinases on soft gels with diisocyanate chemistry, which has less bleeding than CNBr; the sensitivity of the technique was not demonstrated, but an absence of autolysis products was claimed.

Digestion of proteins separated by electrophoresis

After isolating a protein by electrophoresis, there are several options for digesting the protein to obtain internal sequence data. The gel may be dissolved, as described above; this option seems rarely used. Electroelution or passive elution can remove the protein from the gel; these techniques can work, and have been used successfully here. Digestion in the gel or after transferring to PVDF are commonly used, and in the hands of experienced laboratories, are equally efficient (Erdjument-Bromage *et al.*, 1995); transfer to nitrocellulose was less efficient.

For both in gel digestion and on PVDF digestions, electrophoresis was performed using high purity reagents (Bio-Rad), and the gels, 1.5 mm thick, aged for 24-48 hours prior to use. Fresh buffers for the gel are suggested. Samples were heated at 37°C for 15 minutes in sample buffer. A control piece of gel or PVDF is advised to identify artifacts from reagents. Using chromatographically repurified Coomassie blue (Sigma or Aldrich) reduces the number of artifact peaks seen during the peptide separation; dye from other suppliers may have a higher dye content but is not recommended. Coomassie Blue R-250 is about 50% pure, with about 45% of the impurities being less polar compounds, and the remainder tri and tetra-sulfonated molecules, blue-green colour, that inhibit binding by Coomassie Blue (Kundu *et al.*, 1996; the authors describe a purification scheme.

iltering through Nalgene disposable filterware may extract something which inhibits the nzyme by 50%. If trypsin is used, which is common, modified trypsin from Promega or oehringer will give less autolysis products.

n all cases, a high protein concentration is necessary for successful analysis, which means 2 µg r more of the protein of interest per lane (0.05 µg/mm3). Williams and Stone (1995) found that he concentration of protein in the gel is more important than the total amount, mentioning one rotein which gave data when redigested at a lower amount but higher concentration than an nsuccessful first digestion.

n gel digestions give greater yields of peptides than digestion on PVDF, especially of large or lycosylated peptides (Merewether *et al.*, 1995). In contrast to digestions in solution, methionine ppeared to be oxidised. On the other hand, Mørtz *et al.* (1994) found that in gel digestions gave arge unidentified peaks and were not reproducible and digestions on PVDF gave slightly greater ields than nitrocellulose.

Moritz *et al.* (1996) state that broad peptide peaks in a digest are disulfide linked peptides, which an be avoided by reducing and alkylating the protein before digestion.

n gel digestion

The bands were cut in small pieces and 1 ml of 50% methanol added and the supernatant discarded after 20 minutes at room temperature; the destaining was repeated. The gel pieces were dried for 2 minutes in a Speed-Vac, then add 0.5 µg of trypsin in 200 µl of 0.1M NH4CO3 0.1% Tween 20 and incubate for 20 hours at 30°C. Extract peptides twice with 0.1% TFA, 50% acetonitrile (volume not specified) for 30 minutes at 4°C and concentrate supernatants in Speed-Vac to < 100 µl. Lysyl endopeptidase has been used for this procedure.

The overall yield of this process has been estimated at 12% (Williams and Stone, 1996).

digestion on PVDF

The procedure here is from Fernandez *et al.* (1994). Blotting uses 10 mM Tris 100 mM glycine pH 8, 10% methanol to high retention membranes, namely ProBlott, Trans-Blot or Immobilon Psq. Transfer in a tank rather than semi-dry apparatus is believed to give higher yields. After cutting the PVDF in small pieces, add 50 µl of 1% reduced Triton X-100, 10% acetonitrile, 0.1M Tris pH 8. After 30 minutes at room temperature, add 0.2 µg of trypsin and incubate 24 hours at 37°C, then vortex, sonicate for 5 minutes, spin and remove supernatant. Add another 50 µl of above buffer and repeat, then use 100 µl of 0.1% TFA and pool supernatants. The suggested enzyme:substrate ratio is 1:10 but 5-fold variations work.

If MALDI-TOF data is to be obtained, octyl or decyl glucopyranoside or decyl or dodecyl maltopyranoside should be used to eliminate detergent clusters during mass spectrometry (Kirchner *et al.*, 1996); 1% octylglucopyranoside in 25mM ammonium bicarbonate 10% acetonitrile allows MALDI-TOF with saturated cyanohydroxycinnamic acid in 50% acetonitrile 0.1% TFA ((Gharahdaghi *et al.*, 1996).

Sutton *et al.* destain Coomassie stained PVDF with 70% acetonitrile and digest in 4 µl of 25 mM ammonium bicarbonate, 1% octyl glucoside, 10% methanol, 20 µg/ml modified trypsin; extract

peptides with 10 µl of formic acid:ethanol =1:1 for MALDI-TOF. Henzel *et al.* (1994) find that Coomassie does not interfere with alkylation, but does interfere with mass spectrometry or capillary LC. They remove Coomassie with the following, mentioning that direct extraction with methanol can cause significant protein loss: wet with 1-2 µl methanol, add 100 µl water, mix, add 400 µl methanol, mix; add 100 µl chloroform, mix, then remove solvents. To remove peptides> 3,000, incubate blot with 20 µl DMSO and shake for 30 minutes.

Iwamatsu and Yoshida-Kubomura (1996) exploit the retention of hydrophobic digestion products on PVDF by performing sequential digestions on electroblotted proteins. Their suggested order of digestion is lysyl peptidase, endoproteinase Asp-N, trypsin. The reasoning is that hydrophobic peptides not released by the first digestion may be digested and released by a later digestion.

Lui *et al.* (1996) claim that only 50% of peptides are recovered from a endoproteinase Lys-C digest, and 70% from a tryptic digest on membranes (PVDF or nitrocellulose not specified), better recoveries using Zwittergent 3-16 when digesting on either nitrocellulose or PVDF; recoveries were greater off nitrocellulose than PVDF. Thus this grouup recommends use of Zwittergent 3-16 when digesting proteins on nitrocellulose or PVDF membranes.

Digestion of proteins in gels

This process has been long used for peptide mapping (Cleveland *et al.*, 1977).

To digest proteins in a gel with cyanogen bromide, the gel slices were dried, then treated with 0.5 ml of cyanogen bromide in 70% formic acid (protein:CNBr 1:20 or 100 by weight), then the slices were dried again (Jahnen *et al.*, 1990). The resultant peptides could be separated by extraction with 1% trifluoroacetic acid followed by reverse phase separation (Jahnen *et al.*, 1990). To reduce peaks from Coomassie Blue staining, staining was only for 15 minutes or less, and then Coomassie was removed by treating the gel pieces with 2 1 ml aliquots of 3% SDS 50% n-propanol for 3 hours or until the gel was clear followed by 24 hours of washing with water. Extraction of Coomassie Blue increased the yield of peptides and eliminated the Coomassie Blue peaks on chromatography (Ward *et al.*, 1991). The extraction may also remove which can interfere with reverse phase separations of peptides (Bosserhoff *et al.*, 1989).

Limited acid hydrolysis of proteins in gels, after removal of Ponceau S or Coomassie blue stain, appears successful in generating internal sequence data and avoids problems with enzymatic reactions (Vanfleteren *et al.*, 1992).

Stone and Williams destain proteins by adding 500 µl of cold 95% acetone at 4°C for 30 min, and washing twice with 0.1M ammonium bicarbonate, presumably to remove SDS.

A procedure claimed to be more efficient than earlier procedures specifies destaining the gel, incubating in 0.2M ammonium bicarbonate, 50% acetonitrile, and then, unlike other procedures, completely drying the gel, which is then rehydrated with ammonium bicarbonate, 0.02% Tween 20 and trypsin (Hellman *et al.*, 1995). The acetonitrile removes some Coomassie. Avoiding the use of Coomassie may be more efficient, but most people are used to using it, and the procedure works adequately when Coomassie stains is used. The proteins are alkyated before digestion; incubate in sample buffer containing 10 mM DTT, then add 20 mM iodoacetate (N-isopropyliodoacetamide not tested)(Hellman, ABRF posting, March 1997). The author (personal communication) suggests that by completely drying the gel, proteinase is better drawn into the

gel to digest all of the substrate. Reportedly some samples that did not give data with other in gel digestions gave peptides with this protocol. The rest of the procedure is common to other methods This procedure can be used on dried gels. Hellman (e-mail, 1997) says that the Tween has been omitted from this procedure successfully, which was used by many for the 1997 ABRF in gel digestion survey. Moritz *et al.* (1994) also omit Tween in their later procedure for in gel digestions.

A variation is to digest the protein and eluate with SDS and endoproteinase lys C, remove the SDS and then peform a complete digestion (Hwang *et al.*, 1996). A 60% yield is claimed, with success from 10 pmole protein; the authors report the Rosenfeld's procedure does not always work, and is sensitive to the amount of proteinase.

Another variation is a stress on the importance of reducing and alkylating proteins before digestion to prevent disulfide bonds forming (Jenö *et al.*, 1995), a point made by Moritz *et al.* (1996), whose procedure for alkylation in a gel is described in appendix A; this report also shows more efficient digestion after alkylating the heavily cross-linked BSA. The authors suggest that reduction and alkylation prior to electrophoresis may not work efficiently in mixtures but do not offer direct evidence. Instead they digest in the gel prior to electrophoresis and find for efficient production of peptides it is necessary to add SDS which is removed by precipitation with guanidinium chloride.

One group report more autolysis during in gel digestions than during digestions on PVDF, even using autolysis trypsin (Rasmussen *et al.*, 1994).

Digestion of electroblotted proteins

Digestion of proteins on PVDF is common. However the digestion may give different results from digestions in solution because of failure to extract hydrophobic proteins with the common aqueous digestion and extraction systems or failure to cleave in hydrophobic regions of the protein (Rasmussen *et al.*, 1991). However digestion on PVDF gives less autolysis products than in gel digestions, even with modified trypsin (Rasmussen *et al.*, 1994). Henzel *et al.* (1994) find that alkylation of proteins on PVDF is unaffected by Coomassie, but Coomassie must be removed prior to capillary LC and electrospray mass spectrometry. Methanol extraction removes proteins, but the chloroform/methanol extraction of Wessel and Flügge (1984) is suitable. After reduction and alkylatin, the blot is incubated with PVP-360 and trypsin; hydrophobic peptides are extracted with DMSO.

Applied Biosystems (Yuen *et al.* 1988, 1989) has described procedures for performing amino acid analyses, alkylation, and digestions of proteins that have been electroblotted on to PVDF. We prefer not to perform alkylations in the sequencer as described by Yuen *et al.* (1989) because we have found that the sequencer is contaminated by the reagents, so that extensive washing is needed before the sequencer can be used again. A method for with 4-vinylpyridine prior to electrophoresis described by Tempst *et al.* (1990) is to reduce the protein in Laemmli sample buffer, with 2-mercaptoethanol reduced to 0.5%, heated to 60°C for 10 min then 37°C for 20 min; then a 20% solution of 4-vinylpyridine in ethanol is added to give a final concentration of 1.5%. After 30 min at room temperature in the dark, the sample is loaded on to a gel. Drs. P. Glee and K. Hazen in the Department of Pathology have successfully used this procedure enabling identification of a cysteine during sequencing. Without alkylation, digestion,

complete of proteins may be less complete. Cysteine can be alkylated after electrophoresis on VDF (Ploug *et al.*, 1992, Iwamatsu, 1992) prior to digestion of the protein.

mpst *et al.* (1990) found that endoproteinase Glu-C was not successful for *in situ* digestion of oteins. In agreement with the report of Bauw *et al.* (1990), on nitrocellulose, not all possible oteolytic cleavages of substrates occurred. This group advises keeping the reaction volume low 25 µl. Up to 25 mm2 of nitrocellulose can be handled, with enzyme concentrations of 0.04 g/µl or more. The membrane should never be allowed to dry.

nother method of obtaining sequence data from a cyanogen bromide digestion is to electroblot e protein on to PVDF and digest with 150 µl of 0.15 M CNBr in 70% formic acid, then elute e fragments with 75 µl of 2% SDS, 1% Triton X-100, 50 mM Tris pH 9.5 for 90 min at room mperature . The fragments are then separated by electrophoresis, blotted and sequenced. The ield of sequenceable fragments was 2-5% of the material loaded on the gel. (Scott *et al.* 1988). has been reported that CNBr digestion on PVDF leaves about 19% of protein, an amount ncreased to 73% by reaction with PITC (Wadsworth *et al.*, 1992).

BNPS-skatole can digest proteins at on PVDF; the fragments can be separated by running a econd gel, or presumably by chromatography (Crimmins *et al.*, 1990).

A modified procedure for digesting protein blotted on to nitrocellulose has been described by Tempst *et al.* (1990). In studies on digestion of proteins on Fischer *et al.* (1991) used TES buffer pH 8 with 5% acetonitrile, 1 µg of enzyme/100 pmole of substrate (less enzyme gives less digestion, more gives autolysis products) for 8 hours after which no further digestion occurred. Yields of sequenceable peptides were 10-30%; again, not all possible peptides were seen. Proteinases that were used successfully were endoproteinase Glu-C, endoproteinase Asp-N, endoproteinase Lys-C and trypsin. Luo *et al.* (1991) have also described digestion of proteins blotted on to nitrocellulose.

To digest blotted proteins, Wong *et al.* (1992) applied 5 µl of methanol to PVDF, then added cyanogen bromide in 0.1 M HCl and incubated for 3 h. at 45°C. Before extracting electroblotted and digested proteins, Wong *et al.* (1992) used a modified extraction (Wessel and Flügge, 1984) to remove . The methanol wetted blots were placed in 100 µl of water, 400 µl of methanol added, the tube was vortexed for 2 min and 100 µl of chloroform added. After another 2 min, the liquid was removed prior to extraction with DMSO. Methanol or acetone have also been used to remove Coomassie from PVDF. Digestion was enhanced by alkylation of cysteine with iodoacetate (Henzel *et al.*, 1994). To digest with endoproteinase Lys-C, this groups dried off all but 20-50 µl of DMSO, added 0.1 M ammonium bicarbonate to give 10% DMSO, then digested with enzyme at a ratio of 1:20 substrate for 17 h at 37°C. Complete removal of DMSO causes difficulty in dissolving the peptides . DMSO did not alter the pattern of digestion, while a mixture of SDS, proteolysis and DMSO enhanced proteolysis. Similarly Fischer (1993) increased yields of hydrophobic peptides by digesting with trypsin in the presence of 50% DMSO. Another procedure to help extract peptides is the use of 10% trifluoroacetic acid (Rasmussen *et al.*, 1994).

Millipore have developed Immobilon-CD membrane for digestion of electroblotted proteins. The surface is cationic and peptides can be eluted readily after digestion. Common dyes cannot be used, but Promega's ChromaPhor is suitable. Millipore now supply a stain for the membrane.

Immobilon-CD appears to bind proteins as efficiently as unmodified PVDF. Patterson *et al.* (1992) eluted over 70% of bound proteins with 4 M guanidine-HCl and 0.1% Triton X-100, about 70% of CNBr peptides with 70% formic acid and 50% of peptic fragments using 2 M acetic acid. Enzymatic digestions can be performed on the membrane or after eluting the proteins. However it appears that digestion requires an enzyme active at either low pH or in guanidine, precluding the use of trypsin.

Cyanogen bromide may give a different digestion pattern when the substrate is on PVDF rather than in solution (Zhang *et al.*, 1994).

Elution of electroblotted proteins from PVDF

Yuen *et al.* (1988, 1989) also reported that 25% of the protein applied to the gel can be extracted from the membrane. The most efficient solvent for lactoglobulin was 70% 2-propanol, 5% TFA. Szewczyk and Summers (1988) report higher recoveries of protein when mixtures of Triton X-100 and SDS are used to extract proteins from the membranes. Simpson *et al.* (1989b) reported recoveries of 57-78% using 2% SDS, 1% Triton X-100, 0.1% DTT in 50 mM Tris-Cl pH 9. Detergent can be removed by a precipitation technique described above (p.2, Stone *et al.* (1989)) or extraction with heptane: isoamyl alcohol 4:1 (Bosserhoff *et al.*, 1989). Detergent can also be removed by "inverse gradient" reverse phase chromatography as described by Simpson *et al.* (1987, 1989a,b), the protein is digested and the digestion products separated by microbore HPLC (Simpson *et al.*, 1989a,b). SDS may also be removed by precipitation with guanidine-HCl, although recoveries of peptides are variable, especially for large or hydrophobic peptides (Riviere *et al.*, 1991)

Montelaro (1987) has reported extracting 75% of the protein A bound to Immobilon-P membranes with 40% acetonitrile in 0.1 M ammonium acetate pH 8.9 at 37°C for 3 hr. An application of removal of protein from PVDF was reported by Stone *et al.* (1989a) who applied a protein to PVDF in 70% formic acid and then performed a cyanogen bromide digestion by adding 10 µl of 70 mg/ml cyanogen bromide and incubating in the dark for 24 hours. To obtain the peptides, the supernatant was removed, then the PVDF was incubated in 200 µl of 40% acetonitrile for 3 h at 37°C. This supernatant was removed and the PVDF was incubated with 200 µl of 40% acetonitrile, 0.05% trifluoroacetic acid for 20 min at 50oC. The three supernatants were combined with 60 µl of water and dried. Recoveries in this procedure are reported to be 15-100% and protein dependent (Stone *et al.*, 1989b). Another procedure uses 20% acetonitrile in the digestion mixture to enhance digestion of the protein and elution of fragments (Choli *et al.*, 1989).

Although the procedure of Szewczyk and Summers (1988) gives higher recoveries, losses during detergent removal make this latter process less efficient.

Wong *et al.* (1992) have extracted proteins from PVDF with 2 consecutive 2 hour extractions with 200 µl of DMSO with shaking. This procedure is more efficient in extracting proteins from Immobilon-P than from ProBlott, with efficiencies of 32% of BSA and 70% of carbonic anhydrase extracted from ProBlott.

Another technique is to digest a protein in the presence of SDS and separate the peptides by reverse phase HPLC, using a DEAE pre-column to remove the SDS . This technique has been successfully used by Kawasaki *et al* (1990). Using a DEAE column ahead of the reverse phase column can disturb the baseline early in the gradient (Shannon, 1992, unpublished observations). Columns designed specifically for the removal of SDS are now available.

SDS can be quantitated by its interaction with methylene blue (Hayashi, 1975).

As described above, proteins can be electroblotted to nitrocellulose, eluted, digested and sequenced, or digested on nitrocellulose and the peptides eluted. Aebersold (1989) describes a procedure for such digestions.

One method of generating and isolating fragments is to use Promega's Probe-Design Peptide Separation System. The protein of interest is separated on a gel, transferred to an Immobilon PVDF membrane, digested with cyanogen bromide and the peptides are resolved on a gel system designed for peptides, which are then transferred to ProBlott PVDF membrane for sequencing. The necessary materials are supplied by the maker.

A different method for eluting proteins is to displace them with another protein. Chertov *et al.* (1992) eluted a hormone with a solution of 0.1 M ammonium bicarbonate pH 8, 25% acetonitrile, 0.5 mg/ml bovine serum albumin. This technique gives another separation problem if pure protein is needed.

While digesting proteins on PVDF, Zhang *et al.* (1994) observed that 30% acetonitrile with 2.5% TFA was more efficient than 60% acetonitrile-2.5% TFA in extracting hydrophobic peptides from PVDF and that 60% acetonitrile was more efficient for hydrophobic peptides; both solutions also contained 30 mM 4-hydroxycinnamic acid for mass spectrometry and peptide extraction. This group also observed that vacuum drying of a CNBr digestion on PVDF prevented extraction of peptides, but after drying in an open vial, peptides could be extracted with the solution mentioned above.

Another study showed that for the short peptide angiotensinogen, 2% formic or acetic acid removed 70% of bound peptide (Krishnamurty *et al.*, 1994).

After digestion, Mørtz *et al.* (1994) found the best procedure for eluting peptides from PVDF was 20 µl of 50% TFA for 10', and 2 x 50 µl of 0.1% TFA for 15'. Iwamatsu and Yoshida-Kubomura (1996) mention that more peptides are extracted with 20% acetonitrile than 10%.

Cyanogen Bromide

Cyanogen bromide cleaves on the carboxy side of methionine residues. Because there are usually only two methionine residues per 100 residues, several peptides, often large, are generated. Peptides produced by cyanogen bromide digestion have homoserine at their carboxy terminal, unless the peptide comes from the carboxy terminal of the intact protein. Thus the absence of homoserine from a peptide suggests that it is found at the C-terminal of a protein. Homoserine can be readily identified in an amino acid analysis, but is hard to distinguish from serine when PTH amino acids from the sequencer are analyzed. A cyanogen bromide digestion is simpler to perform than an enzymatic digestion. Performing the reaction in 70% formic acid usually avoids the protein solubility problems that often occur with reactions that are performed in aqueous

buffers. The reaction is carried out according to the procedure of Gross (1967). Although a ratio of cyanogen bromide to methionine is often quoted, achieving a precise ratio is difficult because the amount of methionine in 1 mg or less of protein requires an amount of cyanogen bromide that cannot be weighed easily. We have found that the excess of cyanogen bromide produced by adding the smallest amount of cyanogen bromide that can be taken from a large piece does not produce harm. Strydom *et al.* (1985) also used a large excess of cyanogen bromide. Other examples are Asano *et al.* (1986). Handle the cyanogen bromide in the hood; if you wish to weigh it, put the cyanogen bromide in a preweighed, sealable container for weighing. After digestion has taken place, the formic acid is removed on a Speed-Vac, equipped with an acid trap to reduce the amount of formic acid reaching the pump and to reduce the amount of cyanogen bromide which escapes. Dilute the sample 10 x with water first, and repeat the drying to reverse possible esterification of serine and threonine residues with formic acid; this process is repeated. Some references are: Reiman *et al.* (1984), Asano *et al.* (1986), Pastuszyn *et al.* (1987).

Tarr and Crabb (1983) report that exposure to formic acid appears to esterify serine and threonine residues and that aminoethanol reverses the esterification; however standard procedures do not use aminoethanol. Goodlett *et al.* (1990) confirmed formylation of serine and threonine by mass spectrometry and claimed that the formylation did not affect Edman sequencing but did not offer quantitative evidence. There are reports that Met-Cys, Met-Thr and Met-Ser bonds are not efficiently cleaved (Doyen and Lapresle, 1979; Schroeder *et al.*, 1969). With a Met-Cys sequence, methionine is converted to homoserine (loss of 30 mass units) with variable cleavage of the peptide bond. The latter authors achieved more efficient cleavage of Met-Thr and Met-Ser bonds by using 70% trifluoroacetic acid although the reaction proceeded more slowly than in formic acid. These authors also find that 4 hours at 50°C caused random cleavage in catalase. Shively *et al.* (1982) state that use of trifluoroacetic acid for cyanogen bromide digestions avoids blocking of terminal amino groups and Goodlett *et al.* (1990) showed that use of trifluoroacetic acid instead of formic acid reduced esterification of serine and threonine. Guandine hydrochloride may also be used and may give greater yields of peptides, and avoids possible reduction of disulfide bonds by formic acid, which was not directly shown (Villa *et al.*, 1989). Although pyridine formate buffers can block the N-terminus of peptides (Shively *et al.*, 1982; Levy *et al.*, 1981), formic acid is still commonly used for cyanogen bromide digestions.

A study of reaction conditions found that 1 mg/ml CNBr in formic acid for 12-18 hours, gave formylation, oxidation of tryptophan; acetic acid and 0.1M hydrochloric acid gave more acid hydrolysis of peptide bonds; 70% trifluoroacetic acid with 10 mg/ml CNBr for 6 hours gave little acid hydrolysis, no adducts and reduced oxidation of tryptophan (Andrews *et al.*, 1992). When digesting ovalbumin, the cleanest appearing peaks on chromatography were from cyanogen bromide in 70% TFA; 70% formic acid was next best, and 6M guanidine with 3M HCl gave the smallest and broadest peaks (Shannon, 1996, unpublished observations); the yield of peptide was about 25%.

Methionine sulfoxide and sulfone are unaffected by cyanogen bromide. The cyanogen bromide digestion converts tryptophan to monohydroxy tryptophan (+16 mass units).

An alternate procedure is to use cyanogen bromide and potassium iodide to cleave at both methionine and tryptophan residues (Huang and Huang, 1994). Yields of 80-100% are claimed.

Another procedure is to use a mixture of formic and heptafluorbutyric acids, with reported yields of 80% (Ozols *et al.*, 1977).

One reason for incomplete cleavage of proteins by cyanogen bromide is the formation of an unstable intermediate, which will give rise to some of the expected products (Zhang *et al.*, 1996).

Trypsin and Endoproteinase Lys-C

Trypsin cleaves on the carboxy side of lysine and arginine residues, although cleavage at Arg-Pro and Lys-Pro bonds is slow. The bond RXS(P) is not cleaved (Cohen *et al.*, 1991). The specificity of trypsin, its ability to work under denaturing conditions, and its low cost make tryptic digestions a favored digestion. Before performing a tryptic digest, it may be worth seeing how many arginine and lysine residues are present, and whether it might be better to produce fewer, larger fragments by digesting at lysine only with endoproteinase lys-C, or arginine only by citraconylation of lysine residues before digestion.

Trypsin is reported to be active in 6 M urea and 1 mg/ml SDS; our experience suggests that it is not active in 3 M guanidine hydrochloride. Aebersold (1989) reports that trypsin is more active in bicarbonate buffer than Tris-Cl and that trypsin is not inhibited by up to 40% acetonitrile. Suzuki and Terada (1988) reported that 0.05% SDS in 0.2M ammonium acetate pH 8 stimulates hydrolysis of BSA, but this concentration of SDS inhibits activity against a . No other proteins were tested to determine the specificity of the effect of SDS. Stone and Williams (1993) find that trypsin will digest suspensions of proteins unlike endoproteinase lys-C.

Prolonged incubation of trypsin appears to give rise to pseudo-trypsin (C-trypsin) which can hydrolyze bonds adjacent to aromatic residues (Keil-Dlouhá *et al.*, 1971). About one third of trypsin can autolyse to this form whose activity towards chymotrypsin substrates is low but apparently sufficient to give chymotrypsin like digestion (Smith and Shaw, 1969). Promega Corporation and Boehringer Mannheim sell trypsin which is claimed to be resistant to autolytic digestion because it has been alkylated. Methylation of trypsin makes the enzyme more resistant to breakdown and stabilizes the activity, but does not prevent all autolysis (Rice *et al.*, 1977). Although calcium also stabilizes trypsin, reductive methylation further stabilizes it.

Cleavage with endoproteinase lys-C, which cleaves on the carboxy side of lysine (Jekel *et al.*, 1983), may be performed at pH 8 in 0.1 M NH4CO3 and 5 M urea (Reiman *et al.* (1984). Steffens *et al.* (1982) used the same buffer but with 0.1% sodium dodecyl sulfate, proteinase:substrate 1:100 by weight, for 2 hours at 37°C, then added another batch of proteinase and repeated the incubation.

Kawasaki and Suzuki (1990) report that some bonds are cleaved more slowly in the presence of SDS but digestion still occurs even at 1% SDS. Later these workers reported that the digestion pattern in 0.05% SDS is the same as without SDS, but some bonds are not cleaved at higher concentrations, unless more proteinase is used; digestion will occur in up to 2% SDS (Kawasaki and Suzuki, 1992).

Achromobacter protease I is also reported to cleave only on the carboxy terminal side of lysine. Brenner *et al.* (1990) report that *Achromobacter* protease and endoproteinase Lys-C have different activity at some lysine residues, that endoproteinase Lys-C produces less nonspecific

cleavages but requires more enzyme and longer digestion times. With both enzymes, most non-specific cleavages are on the C-terminal side of arginine

Submaxillaris protease (Schenkein *et al.*, 1977) and clostripain or endoproteinase Arg-C are claimed to digest specifically at arginine. However, with a preparation of submaxillaris protease from Pierce, we found that digestion was not specific, and we have had no other experience with these enzymes. One report showed useful digetss by this enzyme, but not at arginine residues (Proudfoot *et al.*, 1995). This enzyme has been used by Simpson *et al.* (1987). Boehringer Mannheim suggest denaturing the substrate in 5M urea but conducting the digestion in 0.1M urea at pH 8, or using 1 mg/ml SDS. Our mass spectrometry lab has also found this enzyme ineffective (M. Kinter, personal communication, 1997), as did Iwamatsu and Yoshida-Kubomura (1996).

Instead we suggest modification of lysine residues with citraconic anhydride before digestion with trypsin (see Reiman *et al.*, (1984); Lau *et al.*, (1985); Atassi and Habeeb, (1972)).

Annan and Biemann (1993) peracetylate protein prior to digestion and analysis by mass spectrometry. Acetylation prevents tryptic digestion at arginine, if acetylation is complete, which it may not be. Amino acids besides lysine may be acetylated.

Trypsin also cleaves at S-2-aminoethyl cysteine (Raftery and Cole, 1966); the modification of cysteine is performed with 2-bromoethylamine (Okazaki *et al.*, 1985). Pierce sell the reagent aminoethyl-8ð (N-(iodoethyl)-trifluoroacetamide) for this reaction and Molecular Probes sell N-(2-iodoethyl)trifluoroacetamide (Schwartz *et al.*, 1980). This digestion is not often used.

We suggest performing digestions with a molar ratio of trypsin to substrate of 1:30 in a volume of 1 ml. A control of trypsin alone should be run to see if any peptides are derived from trypsin instead of the substrate. Digestions with trypsin are usually run at pH 8 and 37°C e.g. Strydom *et al.* (1985), Pastuszyn *et al.* (1987), Lau *et al.* (1985), 4 hr., proteinase :substrate 1:50 by weight.

Trypsin can release exposed peptides within one minute of digestion (Wettenhall and Cohen, 1982).

PerSeptive Biosystems claim that their immobilised trypsin columns allow digestion with a high ratio of enzyme to substrate and hence short digestion times, reducing oxidation of peptides and minimizing trypsin autolysis, which is lessened by immobilization.

Limited Acid hydrolysis

Limited heating with dilute acid can cleave a protein predominantly on the carboxy side of aspartyl residues. The following procedure of Haniu *et al.* (1986) has been used successfully in our laboratory. To dried sample, add 0.5 ml of 0.1% trifluoroacetic acid. Place the tube in a sealed vial which is heated at 110°C for 2 hours. Remove the acid by lyophilizing or in a Speed-Vac.

A published procedure is that of Titani *et al.* (1987). The sample is heated at 110°C for 4 hours with 2 M formic acid. Another procedure is to incubate the peptide at 37°C for 72 hours in 70% formic acid (Reiman *et al.* (1984)). Gerber *et al.* (1979) dissolved insoluble peptides first in 88% formic acid or 50% trifluoroacetic acid before hydrolyzing for 20-35 min. at 110°C in 6M HCl.

rifluoroacetic acid vapour hydrolyzes proteins on Polybrene treated glass fibre filters and VDF at serine, hydrolysis at and threonine, hydrolysis at residues (Hulmes and Pan, 1991). ypical conditions are six days at 22°C with 15-35% efficiency. Higher temperatures increased ydrolysis at , as does the use of liquid acid. Efficiency of digestion was higher on glass fibre lters than for samples in a tube and lowest for samples on PVDF

he lability of Asp-Pro bonds enables them to be hydrolyzed under mild conditions (Landon, 977).

n alternative procedure is to use endoproteinase Asp-N from Boehringer-Mannheim, which eaves on the amino side of aspartate residues (Drapeau, 1980; Ponstingl *et al.*, 1981). Simpson *et al.* (1989c) find that this enzyme does not cleave Asp-Asp bonds and slowly cleaved a Thr-sp bond. Some digestions at glutamate have been reported (Hara *et al.*, 1996).

***taphylococcus aureus* V8 proteinase (Endoproteinase Glu-C)**

his proteinase cleaves on the carboxy side of glutamyl and aspartyl residues in phosphate buffer pH 7.8, and on the carboxy side of only glutamyl residues in ammonium acetate at pH 4. owever, it has been reported that this enzyme will also cleave after glutaminyl, asparaginyl and eryl residues (McWherter *et al.*, 1984). Other workers also find this specificity and that the nzyme does not hydrolyze all bonds adjacent to glutamate (Dognin and Wittman-Liebold, 1977; Valker *et al.*, 1977; Moonen *et al.*, 1980). non-specific cleavage Conditions for using this roteinase are described by Asano *et al.* (1986), proteinase:substrate 1:27 by weight in 50 mM mmonium acetate pH 4 at 37°C for 18 hours; Pastuszyn *et al.* (1987), proteinase:substrate 1:41 y weight, 0.1 M ammonium acetate pH 4 overnight at room temperature; Titani *et al.* (1986), roteinase:substrate 1:100 molar ratio, 0.1 M ammonium bicarbonate pH 8 for 48 hours at 25°C. his enzyme retains activity in 6 M urea, 5.5 M guanidine-HCl, 5 mg/ml SDS. Aebersold (1989) eports that this proteinase is not inhibited by up to 40% acetonitrile. Kawasaki and Suzuki 1990) report that endoproteinase Glu-C shows low activity in 0.05% SDS and gives different igestion in SDS (Kawasaki and Suzuki, 1992).

ones *et al.* (1994) describe a number of non-specific cleavages by this proteinase and describe ptimized conditions for digestion of one substrate. Another response to non-specific cleavages nd transpeptidation by this enzyme was to add endoproteinase Asp-N (Hara *et al.*, 1996).

ome groups find that this proteinase does not digest as efficiently as others (Iwamatsu and oshida-Kubomura, 1996).

leavage at Tryptophan

rocedures for chemical cleavage at tryptophan have been described (Asano *et al.* ,1986; Mahoney *et al.*, 1981; Huang *et al.*,1983; Fontana *et al.*, 1983; Reiman *et al.*, 1984; Houghten nd Li, 1978), Crimmins *et al.* (1990). However, these procedures have not been overly uccessful in our hands. Crimmins *et al.* (1990) report 10-60% yield with BNPS-skatole. xidation of methionine can occur. The major side reaction is bromination of tyrosine. Vestling *et al.* (1994) found 6 minutes digestion was adequate and mention that concentrations of BNPS-katole and protein must be chosen carefully to produce clean peptides. Oxidation of both ryptophan and methionine was found.

An alternate procedure is to use cyanogen bromide and potassium iodide to cleave at both methionine and tryptophan residues (Huang and Huang, 1994). Yields of 80-100% are claimed.

Chymotrypsin cleavage

Some examples of the use of chymotrypsin are: Pastuszyn *et al.* (1987), 0.12 M Tris-Cl pH 7.8, 10 mM calcium chloride at room temperature for 2.5 hours with proteinase:substrate 1:140, Reiman *et al.* (1984), 0.1 M ammonium bicarbonate pH 8, 10 mM *p*-aminobenzoic acid at 37°C for 16 hours (proteinase:substrate ratio not given).

The primary specificity of chymotrypsin A is cutting after tyrosine, phenylalanine or tryptophan, but it may also cut after leucine, methionine and alanine (Keesey, 1987) and Boehringer list it as a broad specificity proteinase. An early survey of chymotrypsin A action shows a high proportion of bonds adjacent to leucine being cut, as well as many others (Neil *et al.*, 1966). Chymotrypsin C is more prone than chymotrypsin A to cut after leucine (Folk and Cole, 1965). A study of hydrolysis of dipeptides by chymotrypsin shows that hydrolysis after methionine may be comparable to that after aromatic amino acids and also that chymotrypsin does not have a clear cut specificity (Duan and Laursen, 1994). In addition, Keil (1992) compiled data showing that aspartate in position P1' reduces cleavage, proline in P1' usually prevents cleavage whereas arginine in P1' enhances cleavage.

Cleavage at Asparaginyl-Glycine Bonds

Bornstein and Balian (1977) have described a procedure to cleave proteins between asparaginyl and glycyl residues. An example is in Strydom *et al.* (1985). Modifications are described by Titani *et al.* (1987).

Asparaginylendopeptidase

Asparaginylendopeptidase from jack bean is reported to cleave specifically on the carboxy side of asparagine (Ishii *et al.*, 1990). This information is from the supplier, Takara Biochemicals, 719 Allston Way, Berkeley, CA, 94710, ph (415) 649-9895, or Panvera, Madison, WI (800) 791-1400, whose catalogue shows a digestion by this enzyme and peptide bonds not cleaved. This product is also available from ClonTech (Palo Alto). One group found this enzyme ineffective (Iwamatsu and Yoshida-Kubomura, 1996).

A collection of digestion procedures is presented by Matsuduira (1989) and Stone *et al.* (1989b).

Cleavage at Cysteine

Trypsin also cleaves at S-2-aminoethyl cysteine (Raftery and Cole, 1966); the modification of cysteine is performed with 2-bromoethylamine (Okazaki *et al.*, 1985). This digestion is not often used.

Another method to cleave at cysteine only is to modify aspartate and glutamate residues with a glycineamide-carbodiimide reaction, oxidize cysteine to cysteic acid with performic acid and digest with endoproteinase Asp-N (Wilson *et al.*, 1989). The modification reaction produces modified aspartate and glutamate and the oxidation reaction also modifies methionine, tyrosine and tryptophan. Minor non-specific cleavages were also seen.

)ther digestions

)ucumisin digests on the carboxyl side of digestion at charged amino acids. Clontech is a)pplier. Enzymes that cleave after proline have been identified (Szwajcer-Dey *et al.*, 1992; ;alwant and Porter, 1991), although they also cleave after alanine (Yoshimoto *et al.*, 1978). roline specific endopeptidase (prolyl endopeptidase, post-proline cleaving enzyme) is available rom ICN

leferences

)ebersold, R.H., Teplow, D.B., Hood, L.E. and Kent, S.B.H. (1986) *J. Biol. Chem.* 262 4229-238

)ebersold, R.H., Leavitt, J., Saavedra, R.A., Hood, L.E. and Kent, S. B.H. (1987) *Proc. Natl. lcad. Sci.* 84 6970-6974

)ebersold, R. (1989) pp.71-88 in *A Practical Guide to Protein and Peptide Purification for Microsequencing* Matsuidaira, P.T. ed. (Academic Press, New York)

)kiyama, T.A., Sasagawa, T., Suzuki, M. and Titani, K. (1994) *Anal. Biochem.* 222 210-216

)lba, F.J., Bermudez, A., Bartolome, S. and Daban, J.-R. (1996) *BioTechniques* 21 624-626

)limi, E., Martinage, A., Sautière, P. and Chevaillier, P. (1993) *BioTechniques* 15 912-917

)llen, G. (1989) Sequencing of Proteins and Peptides, Laboratory Techniques in Biochemistry)nd Molecular Biology v. 9 (R.H. Burdon and P.H. van Knippenberg eds) (Elsevier)

)nderson, P.J. (1985) *Anal. Biochem.* 148 105-110

)ndrews, P.C. and Dixon, J.E. (1987) *Anal. Biochem.* 161 524-528

)ndrews, P.C., Allen, M.H., Vestal, M.L. and Nelson, R.W. (1992) pp. 515-523 in Techniques)n Protein Chemistry III Ed. R.H. Angeletti (Academic Press)

)nnan, R.S. and Biemann, K. (1993) *J. Am. Soc. Mass Spectrom.* 4 87-96

)sano, K., Svensson, B., Svendsen, I. and Poulsen, F.M. (1986) *Carlsberg. Res. Commun.* 51 129-141

)shani, Y. and Catravas, G.N. (1980) *Anal. Biochem.* 109 55-62

)squith, T.N., Keough, T.W., Takigiku, R., Lacey, M.P., Purdon, M.P. and Gauggel, D.L. (1992) pp. 99-106 in *Techniques in Protein Chemistry* R.H. Angeletti ed. (Academic Press, San Diego)

)tassi, M.Z. and Habeeb, A.F.S.A. (1972) *Methods in Enzymology* 25 546-553

)ailey, J.M. and Shively, J.E. (1990) *Biochemistry* 29 3145-3156

)ailey, J.M., Shenoy, N.R. and Shively, J.E. (1992) pp.11-21 in *Techniques in Protein Chemistry III* R.H. Angeletti ed. (Academic Press, San Diego)

Bailey, J.M., Nikfarjam, F., Shenoy, N.R. and Shively, J.E. (1992) *Protein Science* 1 1622-1633

Baldwin, S.W., Stoney, K. and Hail, M.E. (1995) *Protein Science* 4 suppl. 2 151

Baker, C.S., Dunn, M.J. and Yacoub, M.H. (1991) *Electrophoresis* 12 342-348

Baumann, M. (1990) *Anal. Biochem.* 190 198-208

Bauw, G., De Loose, M., Inzé, D., van Montagu, M. and Vandekerckhove, J. (1987) *Proc. Natl. Acad. Sci.* 84 4806-4810

Bauw, G., Rasmussen, H.H., Van Den Bulcke, M., Van Damme, J., Puype, M., Gesser, B., Celis, J.E. and Vandekerckhove, J. (1990) *Electrophoresis* 11 528-536

Bergman, T., Gheorghe, M.T., Hjelmqvist, L. and J-rnvall, H. (1996) *FEBS Letts.* 390 199-202

Bermudez, A., Daban, J.-R., Garcia, J.R. and Mendez, E. (1994) *BioTechniques* 16 621-624

Bhavsar, J.H., Remmler, J. and Lobel, P. (1994) *Anal. Biochem.* 221 234-242

Bickar, D. and Reid, P.D. (1992) *Anal. Biochem.* 203 109-115

Bizzozero, O.A., Good, L.K. and Evans, J.E. (1990) *Biochem. Biophys. Res. Commun.* 170 375-382

Black, S.D. and Coon, M.J. (1986) pp. 161-216 in "Cytochrome P-450" (P.R. Ortiz de Montellano ed.) (Plenum Publ., NY)

Blombäck, B. (1967) *Methods in Enzymology* 11 398-411

Bornstein, P. and Balian, G. (1977) *Methods in Enzymology* 47 132-145

Bosisio, A.B., Loeherlein, C., Snyder, R.S. and Righetti, P.G. (1980) *J. Chromatogr.* 189 317-330

Bosserhoff, A., Wallach, J. and Frank, R.W. (1989) *J. Chromatog.* 473 71-77

Brandt, W.F., Henschen, A. and von Holt, C. (1980) *Hoppe-Seyler's Z. Physiol. Chem.* 361 943-952

Bray, D. and Brownlee, S.M. (1973) *Anal. Biochem.* 55 213-221

Breitlow, S. (1992) *Promega Notes* 35 15-17

Brot, N. and Weissbach, H. (1983) *Arch. Biochem. Biophys.* 223 271-281

Brown, J.L. and Roberts, W.K. (1976) *J. Biol. Chem.* 251 1009-1014

Bull, H.B. (1957) *Arch. Biochem. Biophys.* 68 102-111

Burkhardt, W. (1993) pp. 399-406 in *Techniques in Protein Chemistry IV* R.H. Angeletti ed. (Academic Press)

Busby, W.H., Jr., Quackenbush, G.E., Humm, J., Youngblood, W.W. and Kizer, J.S. (1987) *J. Biol. Chem.* 262 8532-8536

Canova-Davis, E., Kessler, T.J. and Ling, V.T. (1991) *Anal. Biochem.* 196 39-45

Cannon-Carlson, S. and Tang, J. (1997) *Anal. Biochem.* 246 146-148

Cheetham, P.S.J. (1979) *Anal. Biochem.* 92 447-452

Chertov, O., Krasnoselskii, A., Blischenko, E., Telezhinskaya, I. and Chertova, E. (1992) *J. Protein Chem.* 11 353-354

Chiari, M., Righetti, P.G., Negri, A., Ceciliani, F. and Ronchi, S. (1992) *Electrophoresis* 13 882-884

Chin, C.C.Q. and Wold, F. (1985) *Bioscience Reports* 5 847-854

Chin, C.C.Q. and Wold, F. (1986) pp.505-512 in "Methods in Protein Sequence Analysis" K.A. Walsh ed. Humana Press, Clifton N.J. 1986

Choli, T., Kapp, U. and Wittman-Liebold, B. (1989) *J. Chromatog.* 476 59-72

Christianson, T. and Paech, C. (1994) *Anal. Biochem.* 223 119-129

Clarke, S., Vogel, J.P., Deschens, R.J. and Stock, J. (1988) *J. Biol. Chem.* 85 4643-4647

Clauser, K.R. and Burlingame, A.L. (1993) *Protein Science* 2 suppl. 1. 105

Cleveland, D.W., Fischer, S.G., Kirschner, M.W. and Laemmli, U.K. (1977) *J. Biol. Chem.* 252 1102-1106

Cohen, P., Gibson, B.W. and Holmes, C.F.B. (1991) pp. 153-168 in *Methods in Enzymology 201* T. Hunter and B.M. Sefton eds. (Academic Press)

Correia, J.J., Lipscomb, L.D. and Lobert, S. (1993) *Arch. Biochem. Biophys.* 300 105-114

Crimmins, D.L., McCourt, D.W., Thoma, R.S., Scott, M.G., Macke, K. and Schwartz, B.D. (1990) *Anal. Biochem.* 187 27-39

Davagnino, J., Wong, C., Shelton, L. and Mankarious, S. (1995) *J. Immunol. Meth.* 185 177-180

Davis, M.T., Lee, T.D., Ronk, M. and Hefta, S.A. (1995) *Anal. Biochem.* 224 235-244

De Jongh, K.S., Rotman, E.I. and Murphy, B.J. (1992) pp. 179-186 in *Techniques in Protein Chemistry IV* R.H. Angeletti ed. (Academic Press, San Diego)

Dilone, E, Costello, M.A. and Bongers, J. (1994) *Protein Science* 3 suppl. 1 146

Dixon, J.D., Gildea, B.D. and Coull, J.M. (1993) *Millipore Bioforum* 4 7-9

Dognin, M.J. and Wittman-Liebold, B. (1977) *FEBS Lett.* 84 342-346

Dopheide, T.A. and Ward, C.W. (1981) *J. Gen. Virol.* 52 367-370

Doyen, N. and Lapresle, C. (1979) *Biochem. J.* 177 251-254

Drapeau, G.R. (1980) *J. Biol. Chem.* 255 839-840

Drapeau, G.R. (1977) *Methods in Enzymology* 47 189-191 C.H.W. Hirs and S.N. Timasheff eds. (Academic Press)

Driessen, H.P. (1985) *Critical Reviews in Biochemistry* 18 281-325

Duan, Y. and Laursen, R.A. (1994) *Anal. Biochem.* 216 431-438

Dunbar, B. and Wilson, S.B. (1994) *Anal. Biochem.* 216 227-228

Eckerskorn, C., Mewes,W., Goretzki, H. and Lottspeich, F. (1988) *Eur. J. Biochem.* 176 509-519

Eckerskorn, C. and Lottspeich, F. (1990) *Electrophoresis* 11 544-561

Eckerskorn, C. and Grimm, R. (1996) *Electrophoresis* 17 899-906

Erdjument-Bromage, H., Geromanos, S., Chodera, A. and Tempst, P. (1993) pp. 419-426 in *Techniques in Protein Chemistry IV* R.H. Angeletti ed. (Academic Press)

Erdjument-Bromage, H., Fernandez, J., Kobayashi, R., Lane, W., Mische, S., Poppito, N., Reim, D., Robinson, R., Speicher, D., Stone, K., Tempst, P. and Williams, K. (1995) *Association of Biomolecular Resource Facilities collaborative study*

Esch, F.S. (1984) *Anal. Biochem.* 136 39-47

Feick, R.G. and Shiozawa, J.A. (1990) *Anal. Biochem.* 187 205-211

Fernandez, J., DeMott, M., Atherton, D. and Mische, S.M. (1992) *Anal. Biochem.* 201 255-264

Fernandez, J., Andrews, L. and Mische, S.M. (1993) *Protein Science* 2 suppl. 1 103

Fernandez, J., Andrews, L. and Mische, S.M. (1994) *Anal. Biochem.* 218 112-117

Fernandez-Patron, C., Castellanos-Serra, L. and Rodriguez, P. (1995a) *Anal. Biochem.*224 203-211

Fernandez-Patron, C., Madrazo, J., Hardy, E., Mendez, E., Frank, R., Castellanos-Serra, L. (1995b) *Electrophoresis* 16 911-920

Fischer, W.H. and Spiess, J. (1987) *Proc. Natl. Acad. Sci.* 84 3628-3632

Fischer, W.H., Park, M., Lucero, L. and Vale, W. (1991) pp.163-170 in *Techniques in Protein Chemistry II* J.J. Villafranca (ed.) (Academic Press, San Diego)

Fischer, W.H. (1993) *Protein Science* 2 suppl. 2 104

Fischer, W.H. and Park, M. (1992) *J. Protein Chem.* 11 366

FMC *Resolutions* 7(3) (1991) (Company newsletter)

Folk, J.E. and Cole, P.W. (1965) *J. Biol. Chem.* 240 193-197

Fontana, A., Dalzoppo, D., Grandi, C. and Zambonin, M. (1983) *Methods in Enzymology* 91 311-318

Fordyce, A.M., Midwinter, G.G. and Moore, C.H. (1979) *Biochem. Soc. Trans.* 7 721-723

Furka, Á., Dibó, G., Kovács, J. and Sebestyén, F. (1983) *Anal. Biochem.* 129 14-21

Geiger, T. and Clarke, S. (1987) *J. Biol. Chem.* 262 785-794

Geisthardt, D. and Kruppa, J. (1987) *Anal. Biochem.* 160 184-191

Gelfi, C. and Righetti, P.G. (1981) *Electrophoresis* 2 213-219

Gerber, G.E., Anderegg, R.J., Herlihy, W.C., Gray, C.P., Biemann, K. and Khorana, H.G. (1979) *Proc. Natl. Acad. Sci.* 76 227-231

Ghaffari, S.H., Dumbar, T.S., Wallace, M.O. and Olson, M.O.J. (1988) *Anal. Biochem.* 171 352-359

Gharahdaghi, F., Kirchner, M., Fernandez, J. and Mische, S.M. (1996) *Anal. Biochem.* 233 94-99

Gillespie, P.G. and Gillespie, S.K.H. (1997) *Anal. Biochem.* 246 239-245

Goodlett, D.R., Armstrong, F.B., Creech, R.J. and van Breemen, R.B. (1990) *Anal. Biochem.* 186 116-120

Gorman, J.J. and Shiell, B.J. (1993) *J. Chromatogr.* 646 193-205

Grego, B., van Driel, I.R., Stearne, P.A., Goding, J.W., Nice, E.C. and Simpson, R.J. (1985) *Eur. J. Biochem.* 148 485-491

Gray, W.R. (1993) *Protein Sci.* 2 1732-1748

Gross, E. (1967) in *Methods in Enzymology* 11 238-255

Gültekin, H. and Heerman, K.H. (1988) *Anal. Biochem.* 172 320-329

Gurd, F.N.R. (1972) *Methods in Enzymology* 25 424-438

Hager, D.A. and Burgess, R.R. (1980) *Anal. Biochem.* 109 76-86

Hanash, S.M., Strahler, J.R., Neel, J.V., Hailat, N., Melhem, R., Keim, D., Zhu, X.X., Wagner, D., Gage, D.A. and Watson, J.T. (1991) *Proc. Natl. Acad. Sci.* 88 5709-5713

Haniu, M., Iyanagi, T., Miller, P., Lee,T.D. and Shively, J.E. (1986) *Biochemistry* 25 7906-7911

Hansen, J.N., Pheiffer, B.H. and Boehnert, J.A. (1980) *Anal. Biochem.* 105 192-201

Hansen, J.N. (1981) *Anal. Biochem.* 116 146-151

Hara, S., Rosenfeld, R. and Lu, H.S. (1996) *Anal. Biochem.* 243 74-79

Harrington, M.G. (1990) pp. 488-495 in *Methods in Enzymology 182* M.P. Deutscher ed. (Academic Press)

Hawke, D. and Yuan, P. (1987) Applied Biosystems protein sequencer User Bulletin 28

Hayahsi, K. (1975) *Anal. Biochem.* 67 503-506

Hayashi, R. (1977) *Methods in Enzymology* 47 84-93

Hayashi, T. and Sasagawa, T. (1993) *Anal. Biochem.* 209 163-168

Hellman, U., Wernstedt, C., Gónñez, J. and Heldin, C.-H. (1995) *Anal. Biochem.* 224 451-455

Hennessey, J.P. and Scarborough, G.A. (1989) *Anal. Biochem.* 176 284-289

Henzel, W.J., Billeci, T.M., Stults, J.T., Wong, S.C., Grimley, C. and Watanabe, C. (1994) pp. 3-9 in *Techniques in Protein Chemistry V* J.W. Crabb ed. (Academic Press)

Henzel, W.J., Grimley, C., Bourell, J.H., Billeci, T.M., Wong, S.C. and Stults, J.T. (1994) *METHODS: A Companion to Methods in Enzymology* 6 239-247

Hermansen, L.F., Pedersen, Ø. and Sletten, K. (1992) *J. Protein Chem.* 11 353

Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981) *J. Biol. Chem.* 256 7990-7997

Hirabayashi, J. and Kasai, K.-i. (1992) *J. Chromatogr.* 597 181-187

Hirano, H., Komatsu, S., Kajiwara, H., Takagi, Y. and Tsunawa, S. (1993) *Electrophoresis* 14 839-846

Hochstrasser, D.F., Harrington, M.G., Hochstrasser, A.-C., Miller, M.J. and Merril, C.R. (1988) *Anal. Biochemistry* 173 424-435

Houghten, R.A. and Li, C.H. (1978) *Int. J. Peptide Protein Res.* 11 49-58

Hsieh, J.-C., Lin, F.-P. and Tam, M.F. (1988) *Anal. Biochem.* 170 1-8

Huang, H.V., Bond, M.B., Hunkapiller, M.W. and Hood, L.E. (1983) *Methods in Enzymology* 91 318-324

Huang, S.S. and Huang, J.S. (1994) *J. Protein Chem.* 13 450-451

Hulmes, J.D. and Pan, Y-C.E. (1991) *Anal. Biochem.* 197 368-376

Hunkapiller, M.W., Lujan, E., Ostrander, F. and Hood, L.E. (1983) *Methods in Enzymology* 91 227-236

Hunkapiller, M.W., Hewick, R.M., Dreyer, W.J. and Hood, L.E. (1983) *Methods in Enzymology* 91 399-413

Hwang, B.J. and Chu, G. (1996) *Biotechniques* 20 982-984

Hwang, B.J., Smith, A.J. and Chu, G. (1996) *J. Chromatogr. B* 686 165-175

Ishii, S.-i., Abe, Y., Hatsushita, H. and Kato, I. (1990) *J. Protein Chem.* 9 294

Iwamatsu, A. (1992) *Electrophoresis* 13 142-147

Iwamatsu, A. and Yoshida-Kubomura, N. (1996) *J. Biochem.* 120 29-34

Jahnen, W., Ward, L.D., Reid, G.E., Moritz, R.L. and Simpson, R.J. (1990) *Biochem. Biophys. Res. Commun.* 166 139-145

Jenö, P., Mini, T., Moes, S., Hintermann, E. and Horst, M. (1995) *Anal. Biochem.* 224 75-82

Jones, W.M. and Manning, J.M. (1985) *Biochem. Biophys. Res. Commun.* 126 933-940

Jones, M.D., Merewether, L.A., Clogston, C.L. and Lu, H.S. (1994) *Anal. Biochem.* 216 135-146

Jungblut, P., Eckerskorn, C., Lottspeich, F. and Klose, J. (1990) *Electrophoresis* 11 581-588

Kain, S.R. and Henry, H.L. (1990) *Anal. Biochem.* 189 169-172

Kalwant, S. and Porter, A.G. (1991) *Biochem. J.* 276 237-244

Kapadia, G. and Chrambach, A. (1972) *Anal. Biochem.* 48 90-102

Kaplan, H. and Oda, G. (1983) *Anal. Biochem.* 132 384-388

Kaplan, J.M., Mardon, G., Bishop, J.M. and Varmus, H.E. (1988) *Molec. Cell. Biol.* 8 2435-2441

Kawasaki, H. and Suzuki, K. (1990) *Anal. Biochem.* 186 264-268

Kawasaki, H. Emori, Y. and Suzuki, K. (1990) *Anal. Biochem.* 191 332-336

Kawasaki, H. and Suzuki, K. (1992) *J. Protein Chem.* 11 371-372

Keesey, J. (1987) p. 98 in *Biochemica Information* (Boehringer Mannheim Biochemicals, Indianapolis)

Keil-Dlouhá, V., Zylber, N., Imhoff,J.-M., Tong, N.-T. and Keil, B. (1971) *FEBS Lett.* 16 291-295

Kennedy, T.E., Wager-Smith, K., Barzilai, A., Kandel, E.R. and Sweatt, J.D. (1988) *Nature* 336 499-500

Kennedy, T.E., Gawinowicz, Barzilai, A., Kandel, E.R. and Sweatt, J.D. (1988) *Proc. Natl. Acad, Sci. U.S.A.* 85 7008-7012

Kenny, J.W., Ohms, J.I. and Smith A.J. *Beckman literature*

Khandke, K.M., Fairwell, T., Chait, B.T. and Manjula, B.N. (1989) *Int. J. Peptide Protein Res.* 34 118-123

Kirchner, M., Fernandez, J., Shakey, Q.A., Gharahdaghi, F. and Mische, S.M. (1996) in *Techniques in Protein Chemistry VII* (D. Marshak ed.) (Academic Press, San Diego)

Konigsberg, W.H. and Henderson, L. (1983) pp. 254-259 in *Methods in Enzymology* 91 C.H.W. Hirs and S.N. Timasheff eds. (Academic Press

Kosiarz, J.T., K-hler, H. and Steck, J.L. (1978) *Anal. Biochem.* 86 78-89

Kost, J. Liu, L-S., Ferreira, J. and Langer, R. (1994) *Anal. Biochem.* 216 27-32

Knierim, M., Buchholz, J. and Pflug, W. (1988) *Anal. Biochem.* 172 139-144

Kowit, J.D. and Maloney, J. (1982) *Anal. Biochem.* 123 86-93

Krishna, R.G., Chin, C.C.Q. and Wold, F. (1991) *Anal. Biochem.* 199 45-50

Krishna, R.G. (1992) pp.77-84 in *Techniques in Protein Chemistry III* R.H. Angeletti ed. (Academic Press, San Diego)

Krishnamutry, T., Hauer, C.R., Prabhakaran, M., Freedy, J.G. and Hayashi, K. (1994) *Biol. Mass Spectrometry* 23 719-726

Krowczynska, A.M., Donoghue, K. and Hughes, L. (1995) *BioTechniques* 18 698-703

Krutzsch, H.C. and Inman, J.K. (1993) *Anal. Biochem.* 209 109-116

Kubo, K. (1995) *Anal. Biochem.* 225 351-353

Kundu, S.K., Robey, W.G., Nabors, P., Lopez, M.R. and Buko, A. (1996) *Anal. Biochem.* 235 134-140

Lahm, H.-W., Hawke, D.H., Shively, J.E. and Todd, C.W. (1986) pp. 359-364 in "Methods in Protein Sequence Analysis 1986" K.A. Walsh ed. (Humana Press, Clifton, N.J.)

Landon, M. (1977) pp. 145-149 in *Methods in Enzymology vol 47* C.H.W. Hirs and S.N. Timasheff eds (Academic Press, New York)

Larson, G.A. and Shultz, J.W. (1993) *BioTechniques* 15 316-323

Lau, S.Y.M., Sanders, C. and Smillie, L.B. (1985) *J. Biol. Chem.* 260 7257-7263

Laurière, M. (1995) *Anal. Biochem.* 212 206-211

Lee, F.-J. S., Lin, L.-W. and Smith, J.A. (190) *J. Biol. Chem.* 265 11576-11580

LeGendre, N. and Matsudaira, P. (1988) *Biotechniques* 6 154-159

LeGendre, N. and Matsudaira, P. (1989) pp. 52-69 in *A Practical Guide to Protein and Peptide Purification for Microsequencing* P.T. Matsudaira ed. (Academic Press)

LeGendre, N. Mansfield, M., Weiss, A. and Matsudaira, P. (1993) pp. 71-101 in *A Practical Guide to Protein and Peptide Purification for Microsequencing 2nd ed.* P.T. Matsudaira ed. (Academic Press)

Lever, M. (1977) *Anal. Biochem.* 83 274-284

Levy, W.P., Rubinstein, M., Shively, J., Del Valle, U., Lai, C.-Y., Moschera, J., Brink, L., Gerber, L., Stein, S. and Pestka, S. (1981) *Proc. Natl. Acad. Sci. USA* 78 6186-6190

Lissilour, S. and Godinot, C. (1990) *BioTechniques* 9 397-401

Lombard-Platet, G. and Jalinot, P. (1993) *BioTechniques* 15 668-672

Lottspeich, F. (1990) *J. Protein Chem.* 9 268-269

Lu, H.-S., Lai, P.-H. (1986) *J. Chromatography* 368 215-231

Lui, M., Tempst, P. and Erdjument-Bromage, H. (1996) *Anal. Biochem.* 241 156-166

Luo, K., Hurley, T.R. and Sefton, B.M. (1991) pp.149-152 in *Methods in Enzymology 201* T. Hunter and B.M. Sefton eds. (Academic Press)

Mahoney, W.C., Smith, P.K. and Hermodson, M.A. (1981) *Biochemistry* 20 443-448

Matsudaira, P. (1987) *J. Biol. Chem.* 262 10035-10038

Matsudaira, P. (1989) pp.1-13 in *"A Practical Guide to Protein and Peptide Purification for Microsequencing"* (P. Matsuduira ed.) (Academic Press)

Matsuidaira, P. (1989b) pp. 234-239 in *Methods in Protein Sequence Analysis* B. Wittman-Liebold ed. Springer Verlag

McWherter, C.A., Thannhauser, T.W., Fredrickson, R.A., Zagotta, M.T. and Scherage, H.A. (1984) *Anal. Biochem.* 141 523-537

Merewether, L.A., Clogston, C.L., Patterson, S.D. and Lu, H.S. (1995) pp. 153-160 in *Techniques in Protein Chemistry VI* J.W. Crabb ed. (Academic Press)

Messing, R.A. (1969) *J. Am. Chem. Soc.* 91 2370-2371

Meyer, H.E., Hoffmann-Posorske, E., Korte, H. and Heilmeyer, L.M.G. jr. *FEBS Letts.* 204 61-66

Meuth, J.L., Harris, D.E., Dwulet, F.E., Crowl-Powers, M.L. and Gurd, F.R.N. (1982) *Biochemistry* 21 3750-3757

Molnar, A., Kulcsáár, P. and Prestwich, G.D. (1990) *BioTechniques* 9 534-536

Montelaro, R.C. (1987) *Electrophoresis* 8 432-438

Moonen, P., Akeroyd, R., Westerman, J., Puijk, W.C., Smits, P. and Wirtz, K.W.A. (1980) *Eur. J. Biochem.* 106 279-290

Moos, Jr., M., Nguyen, N.Y. and Liu, T.-Y. (1988) *J. Biol. Chem.* 263 6005-6008

Morgan, J.H., Curtis, F.P. and Nochumson, S. (1991) *BioTechniques* 11 256-261

Moritz, R.L., Eddes, J., Ji, H., Reid, G.E. and Simpson, R.J. (1994) *Protein Science* 3 suppl. 1 95

Moritz, R.L., Eddes, J.S., Reid, G.E. and Simpson, R.L. (1996) *Electrophoresis* 17 907-917

Morrissey, B.W. and Stromberg, R.R. (1974) *J. Colloid Interface Sci.* 46 152

Mørtz, E., Vorm, O., Mann, M. and Roepstorff, P. (1994) *Biol. Mass Spectrom.* 23 249-261

Moyer, M., Harper, A., Payne, G. Ryals, J. and Fowler, E. *J. Protein Chemistry* 9 282-283

Mozdzanowski, J. and Speicher, D., (1990) pp. 87-93 in "Current Research in Protein Chemistry: Techniques, Structure, and Function" J.J. Villafranca ed. (Academic Press, New York)

Mozdzanowski, J., Hembach, P. and Speicher, D.W. (1992) *Electrophoresis* 13 59-64

Mozdzanowksi, J. and Speicher, D.W. (1992a) *Anal. Biochem.* 207 11-18

Nakagawa, S. and Fukuda, T. (1989) *Anal. Biochem.* 181 75-78

Nakamura, S., Iwanaga, S. and Suzuki, T. (1974) *Biochem. Biophys. Res. Commun.* 58 250-256

Neil, G.L., Niemann, C. and Hein, G.E. (1966) *Nature* 210 903-907

Nokihara, K., Kuriki, T. and Morita, N. (1994) J. *Chromatogr. A* 676 233-239

Ochs, D. (1983) *Anal. Biochem.* 135 470-474

Ohhashi, T., Moritani, C., Andoh, H., Satoh, S. Ohmori, S., Lottspeich, F. and Ikeda, M. (1991) *J. Chromatogr.* 585 153-159

Okazaki, K., Yamada, H. and Imoto, T. (1985) *Anal. Biochem.* 149 516-520

Ortiz, M.L., Calero, M., Fernandez-Patron, C., Castellanos, L. and Mendez, E. (1992) *FEBS Letts.* 296 300-304

Otter, T., King, S.M. and Witman, G.B. (1987) *Anal. Biochem.* 162 370-377

Ozols, J. and Gerard, C. (1977) *J. Biol. Chem.* 252 5986-5989

Ozols, J. (1989) *Biochim. Biophys. Acta* 997 121-130

Pardue, K. and Williams, D. (1993) *BioTechniques* 14 580-583

Parekh, B.S., Mehta, H.B., West, M.D. and Montelaro, R.C. (1985) *Anal. Biochem.* 148 87-92

Parmelee, D.C., Benjamin, T., Hoang, T.N. and Thorgeirsson, S.S. (1989) abstr. S175, Third Symposium of the Protein Society, Seattle, WA

Parrado, J., Ayala, A. and Machado, A. (1993) *Peptide Research* 6 13-16

Pastuszyn, A., Noland, B.J., Bazan, J.F., Fletterick, R.J. and Scallen, T.J. (1987) *J. Biol. Chem.* 262 13219-13227

Patterson, S.D., Hess, D., Yungwirth, T. and Aebersold, R. (1992) *Anal. Biochem.* 202 193-203

Patton, W.F., Lam, L., Lui, M., Erdjument-Bromage, H. and Tempst, P. (1994) *Anal. Biochem.* 220 324-335

Phang, T., Ji, I. and Ji, T.H. (1996) *Anal. Biochem.* 234 96-97

Pirie, A. (1971) *Biochem. J.* 125 203-208

Ploug, M., Jensen, A.L. and Barkholt, V. (1989) *Anal. Biochem.* 181 33-39

Ploug, M., Stoffer, B., and Jensen, A.L. (1992) *Electrophoresis* 13 148-153

Pluskal, M.G., Przekop, M.B., Kavonian, M.R., Vecoli, C. and Hicks, D.A. (1986) *Biotechniques* 4 272-282

Podell, D.N. and Abraham, G.N. (1978) *Biochem. Biophys. Res. Commun.* 81 176-185

Pong, A.S. and Wells, J.N. (1994) *Anal. Biochem.* 217 163-165

Ponstingl, H., Krauhs, E., Little, M. and Kempf, T. (1981) *Proc. Natl. Acad. Sci. USA* 78 2757-2761

Proudfoot, A.E.I., Magnenat, E., Haley, T.M., Maione, T.E. and Wells, T.N.C (1995) *Eur. J. Biochem.* 228 658-664

Prussak, C.E., Almazan, M.T. and Tseng, B.Y. (1989) *Anal. Biochem.* 178 233-238

Purcell, C.J. and Speicher, D. (1989) abstr. S184, Third Symposium of the Protein Society, Seattle, WA

Raftery, M.A. and Cole, R.D. (1966) *J. Biol. Chem.* 241 3457-3461

Rasmussen, H.H., Van Damme, J., Bauw, G., Puype, M., Gesser, B., Celis, J.E. and Vanderkerckhove, J. (1991) pp. 103-114 in *Methods in Protein Sequence Analysis*

Rasmussen, H.H., Van Damme, J., Puype, M., Gesser, B., Celis, J.E. and Vanderkerckhove, J. (1991) *Electrophoresis* 12 873-882

Rasmussen, H.H., Mørtz, E., Mann, M., Roepstorff, P. and Celis, J.E. (1994) *Electrophoresis* 15 406-416

Reid, G.E., Ji, H., Eddes, J.S., Moritz, R.L. and Simpson, R.J. (1995) *Electrophoresis* 16 1120-1130

Reim, D.F. and Speicher, D.W. (1992) *Anal. Biochem.* 207 19-23

Reiman, E.M., Titani, K., Ericsson, L.H., Wade, R.D., Fischer, E.H., and Walsh, K.A. (1984) *Biochemistry* 23 4185-4192

Rice, R.H., Means, G.E. and Brown, W.D. (1977) *Biochim. Biophys. Acta* 492 316-321

Rider, M.H., Puype, M., Van Damme, J., Gevaert, K., De Boeck, S., D'Alayer, J., Rasmussen, H.H., Celis, J.E. and Vandekerckhove, J. (1995) *Eur. J. Biochem.* 230 258-265

Rittenhouse, J. and Marcus, F. (1984) *Anal. Biochem.* 138 442-448

Riviere, L.R., Fleming, M., Elicone, C. and Tempst, P. (1991) pp. 171-179 in *Techniques in Protein Chemistry II* J.J. Villafranca (ed.) (Academic Press, San Diego)

Ronnenberg, J., Preitz, B., W-stmeier, G and Diekmann, S. (1994) *J. Chromatogr. B* 656 169-177

Rosenfeld, J., Capdevielle, J., Guillemot, J.C. and Ferrara, P. (1992) *Anal. Biochem.* 203 173-179

Sanchez, J.C., Ravier, F., Pasquali, C., Frutiger, S., Paquet, N., Bjellqvist, B., Hochstrasser, D.F. and Hughes, G.J. (1992) *Electrophoresis* 13 715-717

Sauvé, D.M., Ho, D.T. and Roberge, M. (1995) *Anal. Biochem.* 226 382-383

Schägger, H. and von Jagow, G. (1987) *Anal. Biochem.* 166 368-379

Schenkein, I., Levy, M., Franklin, E.C. and Frangione, B. (1977) *Arch. Biochem. Biophys.* 182 64-70

Schinina, M.E., Carlini, P., Polticelli, F., Zappacosta, F., Bossa, F. and Calabrese, L. (1996) *Eur. J. Biochem.* 237 433-439

Schroeder, W.A., Shelton, J.B. and Shelton, J.R. (1969) *Arch. Biochem. Biophys.* 130 551-556

Schwartz, W.E., Smith, P.K. and Royer, G.E. (1980) *Anal. Biochem.* 106 43-48

Scott, M.G., Crimmins, D.L., McCourt, D.W., Tarrand, J.J., Eyerman, M.C. and Nahm, M.H. (1988) *Biochem. Biophys. Res. Commun.* 155 1353-1359

Scott-Ennis, R.J. and Noltmann, E.A. (1985) *Arch. Biochem. Biophys.* 239 1-11

Sefton, B.M. and Buss, J.E. (1987) *J. Cell. Biol.* 104 1449-1453

Sheer, D. (1990) *Anal. Biochem* 187 76-83

Sheer, D. (1994) pp. 243-248 in *Techniques in Protein Chemistry V* J.W. Crabb ed. (Academic Press)

Shively, J.E., Hawke, D. and Jones, B.N. (1982) *Anal. Biochem.* 120 312-322

Shultz, J. (1992) *Promega Notes* 36 20-21

Simpson, R.J. and Nice, E.C. (1984) *Biochem. International* 8 787-791

Simpson, R.J., Moritz, R.L., Nice, E.C. and Grego, B. (1987) *Eur. J. Biochem.* 165 21-29

Simpson, R.J., Moritz, R.L., Lloyd, C.J., Fabri, L.J., Nice, E.C., Rubira, M.R. and Burgess, A.W. (1987) *FEBS Letts.* 224 128-132

Simpson, R.J., Moritz, R.L., Begg, G.S., Rubira, M.R. and Nice, E.C. (1989a) *Anal. Biochem.* 177 221-236

Simpson, R.J., Ward, L.D., Reid, G.E., Batterham, M.P. and Moritz, R.L. (1989b) *J. Chromatog.* 476 345-361

Simpson, R.J., Moritz, R.L., Rubira, M.R., Gorman, J.J. and van Snick, J. (1989c) *Eur. J. Biochem.* 183 715-722

Smejkal, G. and Gallagher, S. (1994) *BioTechniques* 16 196-202

Smith, R.L. and Shaw, E. (1969) *J. Biol. Chem.* 244 4704-4712

Soderling, T.R. and Walsh, K. (1982) *J. Chromatog.* 253 243-251

Speicher, D.W. (1989) pp.24-35 in Techniques in Protein Chemistry (T.E. Hugli ed. Academic Press)

Speicher, D.W., Mozdzanowski, Beam, K. and Chen, D.-h. *J. Protein Chemistry* 9 254-255

Steffens, G.J., Günzler, W.A., Ötting, F., Frankus, E. and Flohe, L. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* 363 1043-1058

Stevenson, B., DiSpirito, A.A. and Schmidt, T.M. (1994) *BioTechniques* 17 1048-1050

Stone, K., LoPresti, M., McNulty, D., Crawford, M., DeAngelis, R. and Williams, K. (1989) abstr. S187, Third Symposium of the Protein Society, Seattle, WA

Stone, K.L., LoPresti, M.B., Crawford, J.M., DeAngelis, R. and Williams, K.R. (1989b), pp. 31-47 in *A Practical Guide to Protein and Peptide Purification for Microsequencing"* (P. Matsuduira ed.) (Academic Press)

Stone, K.L., LoPresti, M.B., Williams, N.D., Crawford, J.M., DeAngelis, R. and Williams, K.R. (1989c) pp.377-391 in *Techniques in Protein Chemistry* T.E. Hugli ed. (Academic Press, New York)

Strydom, D.J., Fett, J.W., Lobb, R.R., Alderman, E.M., Bethune, J.L., Riordan, J.F. and Vallee, B.L. (1985) *Biochemistry* 24 5486-5494

Sutton, C.W., Pemberton, K.S., Cottrell, J.S., Corbett, J.M., Wheeler, C.H., Dunn, M.J. and Pappin, D.J. (1995) *Electrophoresis* 16 308-316

Suzuki, H. and Terada, T. (1988) *Anal. Biochem.* 172 259-263

Szewczyk, B. and Summers, D.F. (1988) *Anal. Biochem.* 168 48-53

Szwajcer-Dey, E., Rasmussen, J., Meldal, M. and Breddam, K. (1992) *J. Bacteriol.* 174 2454-2459

Tarr, G.E. and Crabb, J.W. (1983) *Anal. Biochem.* 131 99-107

Tempst, P., Link, A.J., Riviere, L.R., Fleming, M. and Elicone, C. (1990) *Electrophoresis* 11 537-553

Teshima, G., Stults, J.T., Ling, V. and Canova-Davis, E. (1991) *J. Biol. Chem.* 266 13544-12547

Tetaz, T., Bozas, E., Kanellos, J., Walker, I. and Smith, I. (1993) pp. 389-397 in *Techniques in Protein Chemistry IV* R.H. Angeletti ed. (Academic Press)

Thompson, D., and Larson, G. (1992) *BioTechniques* 12 656-658

Thomsen, J., Bucher, D., Brunfeldt, K., Nexø, E. and Olesen, H. (1976) *Eur. J. Biochem.* 69 87-96

Titani, K., Torff, H.-J., Hormel, S., Kumar, S., Walsh, K.A., R-dl, J., Neurath, H. and Zwilling, R. (1987) *Biochemistry* 26 222-226

Tovey, E.R. and Baldo, B.A. (1989) *J. Biochem. Biophys. Meth.* 19 169-184

Tous, G.I., Fausnaugh, J.L., Akinyosoye, O., Lackland, H., Winter-Cash, P., Vitorica, F.J. and Stein, S. (1989) *Anal. Biochem.* 179 50-55

Towler, D.A., Adams, S.P., Eubanks, S.R., Towery, D.S., Jackson-Machelski, E., Glaser, L. and Gordon, J.I. (1988) *J. Biol. Chem.* 263 1784-1790

Trudel, J. and Asselin, A. (1994) *Anal. Biochem.* 221 214-216

Tsunasawa, S. and Sakiyama, F. (1984) *Methods in Enzymology* 106 165-170

Tsunasawa, S., Takakura, H. and Sakiyama, F. (1990) *J. Protein Chemistry* 9 265-266

van der Zee, R., Welling-Wester, S. and Welling, G.W. (1983) *J. Chromatogr.* 266 577-584

Vanfleteren, J.R. (1989) *Anal. Biochem.* 177 388-391

van Oss, C.J., Good, R.J. and Chaudhury, M.K. (1987) *J. Chromatog.* 391 53-65

Vanfleteren, J.R., Raymackers, J.G., Van Bun, S.M. and Heheus, L.A. (1992) *BioTechniques* 12 550-557

Van-Seuningen, I. and Davril, M. (1990) *Anal. Biochem.* 186 306-311

Vestling, M.M., Kelly, M.A. and Fenselau, C. (1994) *Rapid Commun. Mass Spectrom.* 8 786-790

Villa, S. De Fazio, G., Canosi, U. (1989) *Anal. Biochem.* 177 161-164

Vinogradov, S.N. and Kapp, O.H. (1983) pp.259-263 in *Methods in Enzymology* 91 C.H.W. Hirs and S.N. Timasheff eds. (Academic Press).

Violand, B.N., Schlittler, M.R., Toren, P.C. and Siegel, N.R. (1990) *J. Protein Chemistry* 9 109-117

Vissers, J.P.C., Chervet, J.-P., and Salzmann, J.-P. (1996) *J. Mass Spectrometry* 31 1021-1027

Wadsworth, C.L., Knuth, M.W., Burrus, L.W., Olwin, B.B. and Niece, R.L. (1992) pp. 61-68 in *Techniques in Protein Chemistry* R.H. Angeletti ed. (Academic Press, San Diego)

Walker, J.M., Hastings, J.R.B. and Johns, E.W. *Eur. J. Biochem.* 76 461-468

Wang, Y., Filo, C.J., DePaoli-Roach, A.A., Bell, A.W., Hermodson, M.A. and Roach, P.J. (1988) *Anal. Biochem.* 174 537-547

Ward, L.D., Reid, G.E., Moritz, R.L. and Simpson, R.J. (1990) *J. Chromatogr.* 519 199-216

Ward, L.D., Reid, G.E., Moritz, R.L. and Simpson, R.J. (1990) pp.179-190 in *Current Research in Protein Chemistry: Technioques, Structure, and Function* J.J. Villafranca ed. (Academic Press)

Ward, L.D., Hong, J. and Simpson, R.J. (1991) pp. 21-30 in *Techniques in Protein Chemistry II* Villafranca, J.J. ed. (Academic Press)

Warlow, R.S., Gooley, A., Rajasekariah, P., Oszarac, N. and Walls, R.S. (1995) *Electrophoresis* 16 84-91

Welinder, K.G. (1988) *Anal. Biochem.* 174 54-6

Wellner, D., Panneerselvam, C. and Horecker, B.L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87 1947-1949

Werner, W.E., Grimley, C. and Yuan, P.-M. (1996) *Anal. Biochem.* 237 146-149

Wessel, D. and Flügge, U.I. (1984) *Anal. Biochem.* 138 141-143

Wettenahll, R.E.H. and Cohen, P. (1982) *FEBS letts.* 140 263-269

Williams, K.R. and Stone, K.L. (1995) pp. 143-152 in *Techniques in Protein Chemistry VI* J.W. Crabb ed. (Academic Press)

Williams, K.R. and Stone, K.L. (1996) abstr S2 ABRF'96 Biomolecular Techniques, San Francisco, CA

Wilson, K.J., Fischer, S. and Yuan, P.M. (1989) pp. 310-314 in *Methods in Protein Sequence Analysis* B. Wittman-Liebold ed. (Springer-Verlag)

Wilson, K.J., Hong, A.L., Brasseur, M.M. and Yuan, P.M. (1986) *Biochromatography* 1 106-112

Wold, F. (1981) *Ann. Rev. Biochem.* 50 783-814

Wong, S., Padua, A. and Henzel, W.J. (1992) pp.3-9 in *Techniques in Protein Chemistry III* R.H. Angeletti (ed.) (Academic Press, San Diego)

Xu, Q.-Y. and Shively, J.E. (1988) *Anal. Biochem.* 170 19-30

Yamada, R. and Bradshaw, R.A. (1991) *Biochemistry* 30 1017-1021

Yoshimoto, T., Fischl, M., Olowski, R.C. and Walter, R. (1978) *J. Biol. Chem.* 253 3708

Yuen, S.W., Chui, A.H., Wilson, K.J. and Yuan, P.M. (1988) Applied Biosystems protein sequencer User Bulletin 36

Yuen, S.W., Chui, A.H., Wilson, K.J. and Yuan, P.M. (1989) *Biotechniques* 7 74-83

Zalut, C., Henzel, W.J. and Harris, H.W. (1980) *J. Biochem. Biophys. Meth.* 3 11-30

Zardoya, R., Diez, A., Mason, P.J., Luzzato, L., Garrido-Pertierra, A. and Bautista, J.M. (1994) *BioTechniques* 16 270-276

Zhang, W., Czernik, A.J., Yungwirth, T., Aebersold, R. and Chait, B.T. (1994) *Protein Science* 3 677-686

Zhang, X., Dillen, L., Vanhoutte, K., Van Dongen, W., Esmans, E. and Claeys, M. (1996) *Anal. Chem.* 68 3422-3430

Note: A

Alkylation of cysteine residues with 4-vinylpyridine

The procedure described here is that of Hawke and Yuan (1987) of Applied Biosystems.
1. Prepare the following solutions: 1 M Tris HCl, pH 8.5 containing 4 mM EDTA; 8 M guanidine-HCl; 10% 2-mercaptoethanol in water.
2. Mix 1 volume of Tris solution with 3 volumes of guanidine solution to give 6 M guanidine buffered with 0.25 M Tris.
3. Dissolve or dilute the sample (1-10 μg) in 50 μl of the above buffer. Add 2.5 μl of 10% 2-mercaptoethanol (3.6 μmole) and incubate at room temperature in the dark under argon for 2 hours.
4. Add 2 μl of 4-vinylpyridine (18.5 μmole) and incubate as above. Store 4-vinylpyridine under argon at -20 ° C to reduce polymerization. If the 4-vinylpyridine is brown, it may be partially polymerized and cause the chromatographic problems noted below.
5. Desalt by reverse phase chromatography.

We have scaled up the above reaction. During desalting of scaled up reactions, we find many peaks that appear to come from the 4-vinylpyridine. Thus separation of 4-vinylpyridine from small peptides may be difficult although cytochrome c can be separated satisfactorily from the other reactants. Tempst *et al.* (1990) report that vinylpyridine must be removed promptly to avoid side reactions and damage to pyridylethyl cysteine. They also find that when desalting by reverse phase chromatography, a 15 to 20 minute wash at 100 μl/min with 3.5% acetonitrile is necessary to remove alkylation reagents; they suggest that the gradient used is advantageous at other times, but for hydrophobic proteins, increased time on a column may decrease yields.

After storage for a year, one batch of 4-vinylpyridine became viscous and formed an insoluble glob in the presence of acetonitrile. This material clogged a column. Some vacuum distill 4-

/inylpyridine and discard after seeing colouration because storage and exposure to light may ;ause polymerization (Gray, 1993).

)esalting by filtration using a device like an Amicon Centricon may lead to an insoluble protein.

Alkylation in gel (Moritz *et al.*, 1996)
Wash gel with water 3 changes over 1.5 hours, then immerse in 50 ml of 10 mM DTT, 0.2M Tris-Cl pH 8.4, 2 mM EDTA for 2 hours at 40°C. Then add 4 vinylpyridine, 2% v/v and incubate for 1 hour at 25°C in dark. Add 2-mercaptoethanol at 2% v/v to stop alkylation and wash gel with water for 1 hour. Addition of guanidine-HCl to improve solubility of vinylpyridine does not improve alkylation but does reduce recovery of peptides.

Note: B

Removal of detergent from samples

To remove SDS, Stone *et al.* (1989) suggest the following procedure: Dialyse 1 ml or less of protein against 0.05% SDS, 5 mM NH4HCO3. Dry. Add 50 µl of water, then 450 µl of cold acetone with 1 mM HCL. Incubate at -20oC for 3 h. Centrifuge saving supernatant in case the protein did not precipitate. Wash the pellet twice with 100 µl of cold acetone and dry.

Bosserhoff *et al.* (1989) have extracted SDS from peptides using heptane:isoamyl alcohol 4:1. The starting solution is 100 µl of 0.1% TFA and 0.05% SDS; add 5 µl of TFA and vortex with 100 µl of heptane:isoamyl alcohol 4:1; centrifuge and remove heptane layer with a pipet and then drying. Repeat if SDS >0.05%.

Konigsberg and Henderson (1983) extracted ion-pairs of dodecyl sulphate in organic solvents.

SDS can be removed from samples loaded on reverse phase columns by precolumns. One type is a DEAE-precolumn described by Kawasaki *et al.* (1990) which may disturb the baseline early in the gradient (Shannon, 1992, unpublished observations). Another precolumn is the SDS precolumn sold by Poly LC and the Nest Group, from which SDS is eluted by high concentration of acetonitrile.

Michrom BioResources sell cartridges for removal of ionic and non-ionic detergents. Although most conveniently used with an HPLC equipped with a 10 port injector, they can be used with a syringe. A capillary LC setup to remove SDS and concentrate a sample for mass spectrometry has been described (Vissers *et al.*, 1996).

Pierce and BioRad sell detergent removal columns, and Supelco recommends Amberlite XAD-2 and XAD-4 resins for detergent removal. The use of Amberlite resins for removing Triton X-100 has been described but the recovery of protein was not measured (Cheetham, 1979; van der Zee *et al.*, 1983).Vinogradovand Kapp (1983) have described this technique.

Another method of removing SDS is to add KCl to 50 mM on ice, leave 10 minutes and remove the KDS precipitate by centrifugation (Prussak *et al.*, 1989). Suzuki and Terada (1988) found that 20 mM KCl was adequate to precipitate dodecyl sulphate. After I added KCl to a solution of peptide in SDS, the solution still bubbled as though detergent was present. Guanidine Hcl also

precipitates SDS but causes variable recovery of peptides, especially large or hydrophobic peptides (Riviere *et al.*, 1991).

A procedure for precipitating proteins from solution, including detergent solutions, is from Wessel and Flügge (1984). To 0.1 ml of solution add 0.4 ml of methanol, mix and centrifuge at 9000g for 10 s; add 0.1 ml of chloroform, mix and spin at 9000g for 10 s. Add 0.3 ml of water, mix vigorously and centrifuge at 9000g for 1 min. Remove upper phase. Add 0.3 ml of methanol to lower phase and interphase with precipitated protein, mix and centrifuge at 9000g for 2 min to pellet protein. A procedure using phenol-ether works with larger volumes, with claimed 80% recovery even from 10 ng/ml solutions (Sauvé *et al.*, 1995)

This procedure was tested with · 40 µg of protein, although the authors claim quantitative recovery with 2 µg of protein. Proteins down to 18,000 MW are claimed to give > 90% recovery. For 1 ml samples, at 4 ml methanol, 1 ml chloroform and 3 ml water. After removal of the upper phase, subsequent steps are as above.

Detergent can also be removed by "inverse gradient" reverse phase chromatography as described by Simpson *et al.* (1987, 1989a,b). Thorough equilibration of the column is necessary for this technique to work; BSA may be a poor test protein because it gives a broad peak (my observation, confirmed by Robert Moritz).

Detergents can also be removed by filtration devices such as Centricon and Ultrafree. Removal of detergent is inefficient above the critical micelle concentration, where the detergent forms high molecular weight micelles. Amicon document 2302 gives a guide to the effectiveness of these devices for different detergents.

Another chromatographic technique is hydrophilic chromatography. Clauser *et al.* (1993) eluted proteins in SDS from polyhydroxyethyl aspartimide column with a gradient of acetonitrile from 90-50% and 10-100 mM ammonium TFA.

BioRad suggest their hydroxyapatite columns to remove detergents from protein solutions.

Non-ionic detergents in protein solutions can be quantitated by passing the solution through a strong cation or anion exchange column to which the protein binds while detergent passes through and is quantitated by UV absorption; there may be no recovery of protein (Pardue and Williams, 1993).

Note: C

Alkylation of cysteine in proteins on PVDF.

Ploug *et al.* (1992) and Iwamatsu (1992) have described procedures, both of which have been tested successfully on lysozyme. After performing the alkylations, the lysozyme was sequenced and amount of alkylated cysteine at position 6 measured. Because of uncertaintities in the quantitation of these derivatives, the measurements are uncertain. Alkylation with 4-vinylpyridine based on the method of Ploug *et al.* (1992) is 90% efficient as measured by area of PTH-pyridylethyl-cysteine, 50% efficient as measured by amounts based on a standard (which may be subject to differential loss during storage of the standard and selective concentration during storage), 60% by peak height. The method followed was to wet the PVDF with methanol

and then incubate in 1 ml of 0.1 M phosphate pH 8, 5 mM EDTA, 10% methanol, 20 mM DTT at room temperature for 1 hour under argon, then incubate for another hour with the above buffer except containing 20 mM 4-vinylpyridine instead of DTT, then rinse with 1 ml of 20% methanol. Using 1 mM DTT and 5 mM 4-vinylpyridine appears to be equally efficient. Alkylation with 20 mM N-isopropyliodoacetamide based on the method of Iwamatsu (1992) appears 90% efficient or more; use of 1 mM reductant and alkylating agent is almost as efficient. The method was to wet the PVDF with methanol, incubate 1 hour at room temperature in 1 ml of 6 M GuHCl, 0.4 M Tris-Cl pH 8.5, 5 mM EDTA, 2% acetonitrile, 20 mM DTT (FW 154) followed by a similar incubation using a smaller volume of 20 mM N-isopropyliodoacetamide (FW 227) in the above buffer instead of DTT.

In a native protein, some cysteine residues do not react with iodoacetate and other alkylating agents, but do react after denaturation (Scott-Ennis and Noltmann, 1985).

Note: D
Materials which interfere with Edman chemistry

1. Primary amines, which compete with the peptide for PITC and also generate UV absorbing peaks which interfere with identification of PTH amino acids.
2. Oxidizing agents which will destroy PITC.
3. Buffers which neutralize the trifluoroacetic acid used for cleavage or the trimethylamine used for the coupling reaction (Krishna *et al.*, 1991).
4. Materials which react with the amino terminus of a peptide. Some of these materials are; cyanates, which may be generated from urea; aldehydes and peroxides which are often found in Triton X-100, Brij and Tween solutions (Lever, 1977; Ashani and Catravas, 1980).
5. SDS, precipitation of protein which may cause sample loss from the filter support and bubbling in the instrument and require service, although reportedly 50 µg can be tolerated. See appendix B for procedures to remove detergent.
6. Buffers containing formic acid which may block the N-terminal of peptides and modify lysine (Shively *et al.*, 1982; Levy *et al.*, 1981).

Note: E
Methods to concentrate protein

Two methods of concentrating proteins by precipitation are the methanol/chloroform technique of Wessel and Flügge (1984), which we have successfully used, and a phenol-ether precipitation method of Sauvé *et al.* (1995), which can handle larger volumes; 80-90% recovery was reported.

Topic 19. Proteomic Technique

Proteomics is the study of proteome, used for analysis of complete complements of proteins. Proteomics includes identification, quantification and determination of proteins localization, modifications, interactions, activities, and their functions. Proteomics is much more complicated than genomics. The genome of an organism is generally constant, whereas the proteome, the functional products of genes, may vary in different forms within a particular cell or between different cells. Moreover, many proteins from every gene are alternative splicing and posttranslational modifications. This machinery affects protein structure, localization, function and turnover. Thus, several protein forms cannot be characterized by genomic analysis alone, providing proteomics is a powerful tool for proteins analysis (1). In proteomic technologies rely on protein separation, protein identification and protein characterization. The technology used in proteomics such as two dimensional (2D) gel electrophoresis, mass spectrometry and protein array are developed.

2D gel electrophoresis or 2DGE is a method for separating complex mixtures of proteins. This method allows the sample to separate in two dimensions for higher-resolution visualization. It works by separation of proteins in two steps: the first-dimension is isoelectric focusing (IEF), which separates proteins according to their isoelectric points (pI); the second-dimension is sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), which separates proteins according to their molecular weights (MW). This provides maximum separation and allows thousands of proteins to be resolved. The resulting gel is then Coomassie Blue, silver and amido black stained for viewing the protein spots. Spots on the gel are proteins that have migrated to specific locations. These can be selected and subjected to analyze by mass spectrometry (1,2,3).

For proteome analysis, two dimensional polyacrylamide gel electrophoresis (2D PAGE) is used to separate mixtures of proteins allowing comparative of related samples - such as healthy and disease tissues. The methodology includes 3 steps (fig. 1-3.) (4).

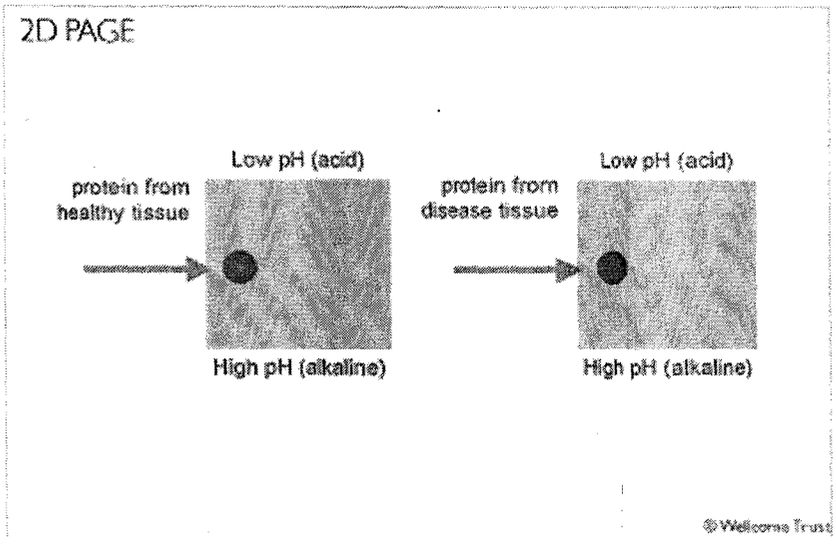


Figure 1. The proteins from healthy and disease tissues are load onto the middle of the left side of the gel, where the pH is neutral. The gel has a pH gradient from top to bottom, which the top is more acidic than the bottom (4).

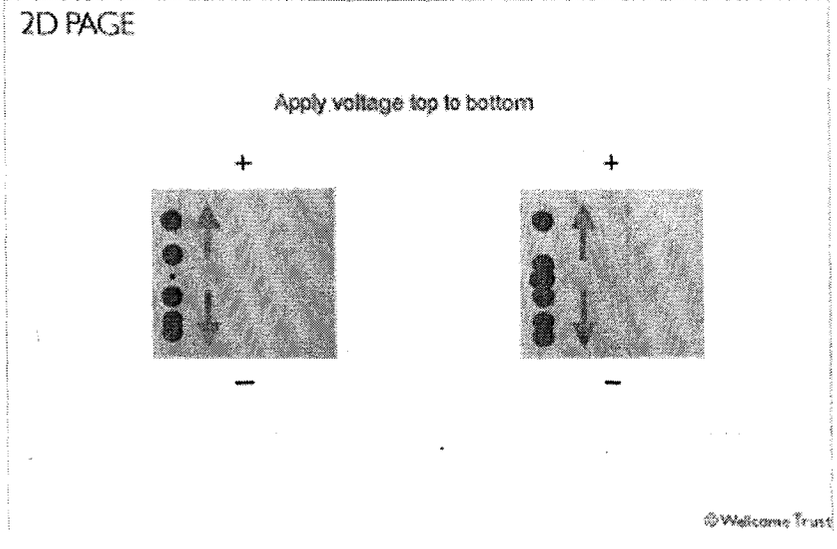


Figure 2. An electric potential is applied across the gel, making one end more positive than the other. If proteins are positively charged, they will be move towards the negative end of the gel and if they are negatively charged, they will be move to the positive end of the gel. The proteins are reaches at their isoelectric point, which the overall on the protein is zero (i.e. a neutral charge). This separation is called isoelectric focusing (4).

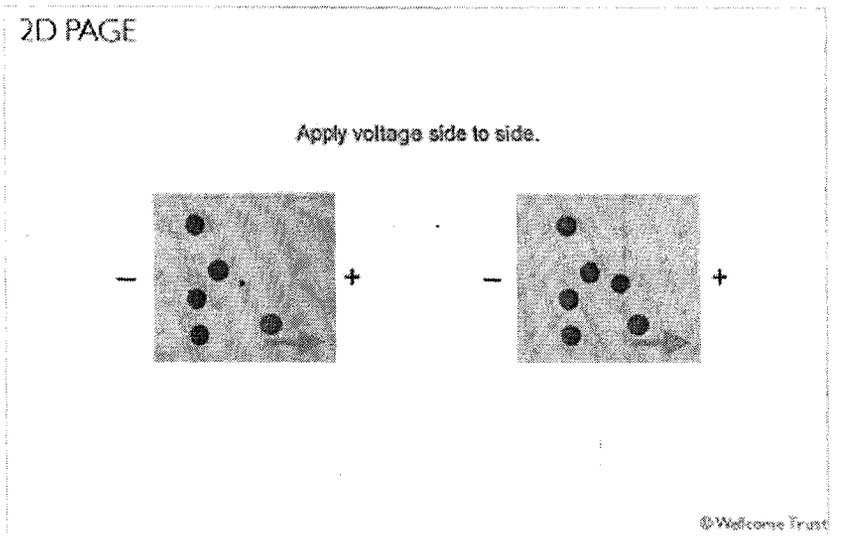


Figure 3. The second step for separated the proteins by charge, according their mass in the perpendicular dimension. By comparing the two gels reveal that proteins are expressed at different levels. In the disease tissue, the red protein is more abundant than in the healthy tissue and could be represent novel drug targets or diagnostic disease markers (4).

The 2DGE permitting the concomitant analysis of hundreds or even thousands of gene products in a single run. However, the reproducibility of 2DGE data has the difficulty to perform because of pH gradient instability with prolonged focusing time as the pH gradient moves towards the cathode (cathodic drift) and flattens in the centre (plateau phenomenon). The comparative of two samples by comparing the images of the stained gels and the exchange of 2DGE data between laboratories has been a major problem because of irreproducibility between 2DGE generated by the conventional method. This problem has overcome by using immobilized pH gradient (IPG) strips. IPG strips are based on the use of immobilized pH gradients, made with buffering acrylamide derivatives that contain a free carboxylic acid or a tertiary amino group which are co-polymerized with acrylamide and bis-acrylamide. The pH gradient is precast into the gel and cannot shift during electrophoresis. Thus cathodic drift is eliminated, enhanced reproducibility and pattern matching, and the comparison of 2DE data between laboratories can be simplified. The strategy in an IEF or IPG combined with SDS-PAGE for proteins separations are showed in figure 4 (1).

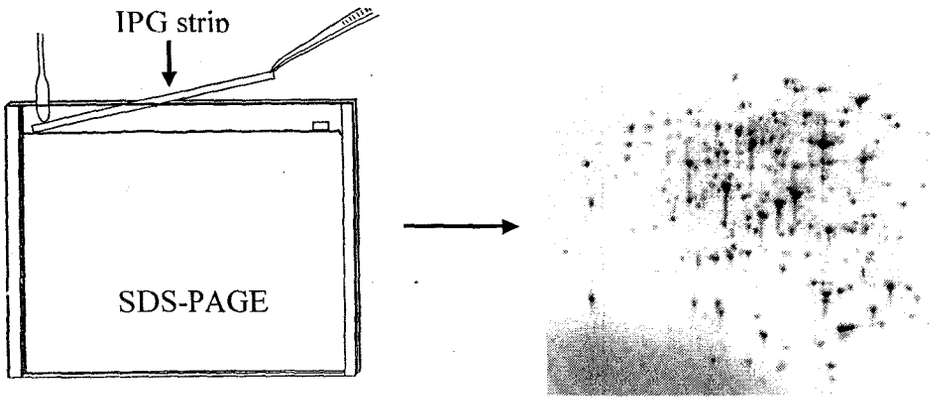


Figure 4. Strategy for proteome analysis by two dimensional gel electrophoresis (1).

The IPG strips for IEF are now commercially available, such as ReadyStrip IPG Strips (Bio-Rad). This commercially system has a variety in the strip length and the linear pH gradient range. The use of IPG strip length and the pH ranges for protein loads per strip are under the product's guidelines information. The samples (cell, tissue or fluid) will be subjected to 2DGE must be prepared by lyses the cell of interest and minimizing damage from protease. The method of lysis will vary greatly depending on the sources of the cells (5).

In previous study, Fatoohi *et al.* have improved the identification of the parasite proteins that induce cellular immunity in naturally infected woman. By using the ReadyStrip IPG strips (Bio-Rad) and SDS-PAGE for separated the presence of soluble *Toxoplasma gondii* antigen (ST-Ag) or fractionate proteins from ST-Ag. Two dimensional separation of ST-Ag showed that all fractions were made up of many different proteins (more than 200 spots) (Fig. 4). However, other proteins in the 20-40 kD range were absence, probably due to the much lower protein concentration (6).

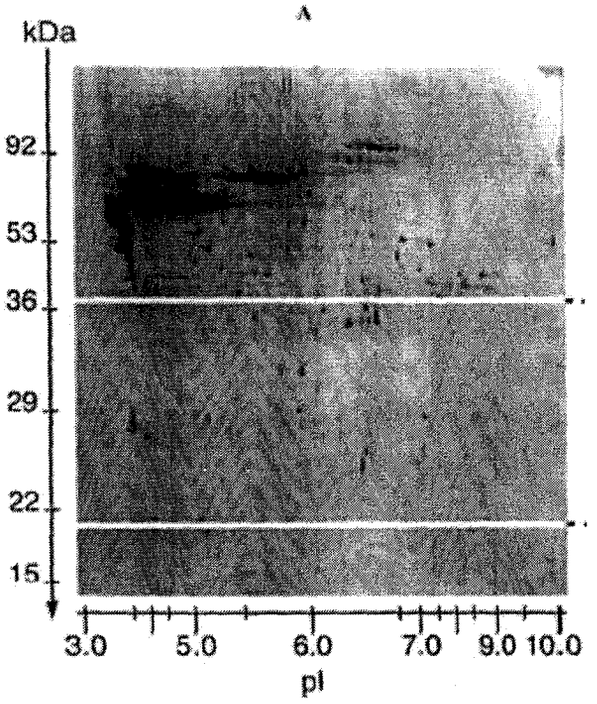


Figure 5. Silver stained two dimensional electrophoresis pattern of tachyzoite ST-Ag (6).

Proteins, which are separated by 2DGE can be determined the specific characteristics of interest by mass spectrometry (MS). The integration technology advances in 2DGE and MS analysis, have increased sensitivity, reproducibility, and throughput of proteome analysis.

Mass spectrometry (MS) provides the ability to accurately measure the mass of almost any molecule that can be ionized to the gas phase. A mass spectrometer consists of three essential components (Fig. 6.); an ion source, a mass analyzer and a detector. An ion source converts molecules into gas-phase ions. Once these ions are created, they are separated in the mass analyzer by their mass (m) to charge (z) ratio and detected by an electron multiplier. MS data are recorded as 'spectra' which display ion intensities versus their m/z value (1).

MS is a powerful analytical technique that is used to identify unknown compounds, to quantify known compounds, and to elucidate the structure and chemical properties of molecules.

For analysis of biomolecules, MS is used to identify proteins and to determine the molecular weight and/or structure of species ranging in size from diatomic molecules to large proteins. MS instruments that have a role in proteomic analysis are matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) and electrospray ionization tandem mass spectrometry (ESI tandem MS or ESI-MS-MS) (1).

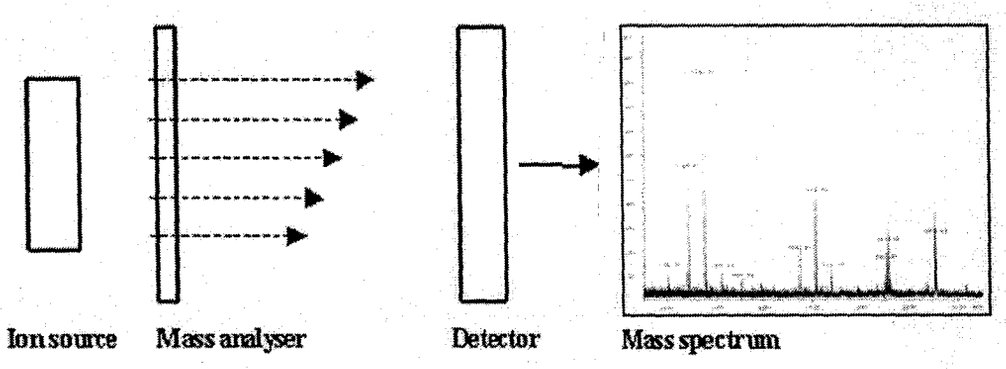


Figure 6. General mass spectrometry scheme (1).

MALDI-TOF MS is composed of two parts, the first part "MALDI" mentions to a method of ionization, while the "TOF" mentions to the mass analyzer. This technology is suitable for the peptide-mass fingerprinting. The samples to be analyzed are first mixed with so called matrix molecules and spotted onto a small plate or slide, and then allowed to evaporate in the air. Upon drying, the matrix molecules crystallize and solid sample/matrix co-crystals are finally formed. The target is then inserted into the ion source of the mass spectrometer which is under a high vacuum. A laser (e.g., a pulsed nitrogen laser at 337 nm) is fired onto the sample, resulting in the desorption event due to absorbance of the laser energy by the matrix molecules. The matrix crystals results in the modulation of the absorbed energy into heat. The rapid heating causes sublimation of the matrix crystals into the gas phase. The ions are ejected from the target surface and accelerated into the mass analyzer (Fig. 7.). At the time that the laser is pulsed a

voltage is applied to the target plate to accelerate the ionized sample towards a time-of-flight mass analyzer. The TOF can be operated in either linear or reflectron mode and can easily separate proteins up to 500 kDa. The precision is considered to be approximately 10 ppm (1,7).

ESI tandem MS instruments, which ESI mentions to the ions are produced in the source of the instrument. Tandem mass spectrometry mentions to mass analyzers that are able to perform two-stage (or multistage) mass analyses of ions. There are different types of mass analyzers are used in ESI tandem MS such as quadrupole, ion trap, and TOF mass analyzers. From different tandem mass analyzers with ESI sources provides instrumental adaptability in proteomics research. ESI tandem MS are performed to acquire partial amino acid sequence or peptide fragmentation. In contrast to MALDI, the peptides or proteins to be analyzed by ESI are in aqueous solution. Typically, peptide mixtures are first separated by liquid chromatography (LC) or high performance liquid chromatography (HPLC). Prior to mass spectral analysis, the protein of interest is either chemically or enzymatically cleaved, the resultant mixture of peptides is fractionated by an online-HPLC from which sample peptides are directly injected into the first mass analyzer. In two quadrupoles series, the first mass spectrometer (MS1) is used to select, from the primary ions, those of a particular m/z value which then pass into the fragmentation region. The ion selected by the MS1 is the parent ion and can be a molecular ion resulting from the primary fragmentation. Dissociation occurs in the fragmentation region. The daughter ions are analyzed in the second spectrometer (MS2) (Fig.8). In fact, the MS1 can be viewed as an ion source for MS2. Tandem mass spectral data interpretation is automated and performed by computer software. The fragmentation of peptides under conditions results in the specific cleavage of the peptide bonds. Consequently the product ion spectrum is a representation of each amino acid in the peptide chain. In the procedure involves the use of a triple-quadrupole mass analyzer (three quadrupoles in series) equipped with an electrospray ionization source that may be interfaced with a high-resolution separation method (Fig. 9). A precursor ion (single peptide) is then selected by the first quadrupole and transferred into the second quadrupole where it is induced to undergo fragmentation. The fragment ions are then analyzed by the third quadrupole mass analyzer with subsequent acquisition of the product ion mass spectrum (1,7).

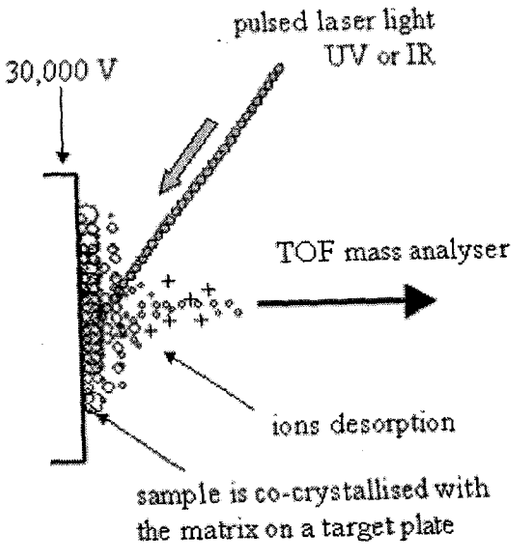


Figure 7. The MALDI ionization process (1).

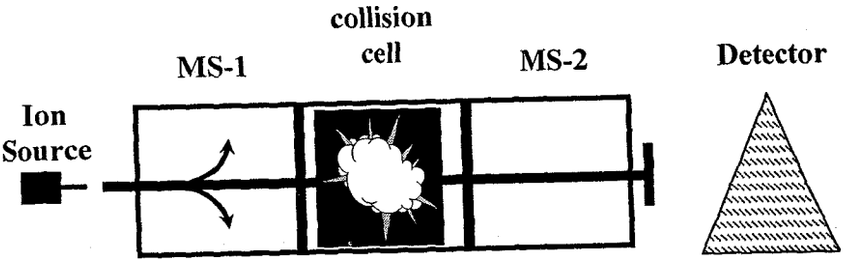


Figure 8. The operation of the tandem MS instrument (1).

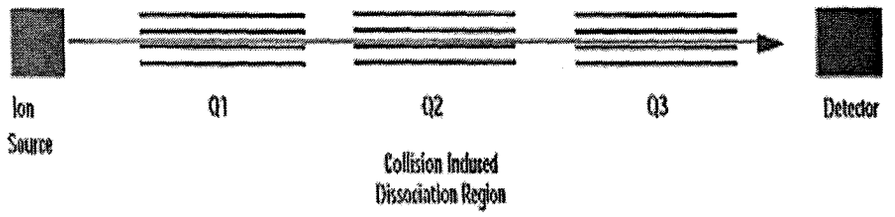


Figure 9. The operation of a triple quadrupole MS instrument (1).

An ion-trap mass analyzer, the design and operation is very different from triple quadrupoles. Ion traps collect and store ions in order to perform MS-MS analyses on them. The analyzer is very simple in design. The ions from the source are directed into the ion trap, which consists of a top and bottom electrode (end caps) and a ring electrode around the middle (fig. 10). Ions collected in the trap are maintained in orbits within the trap by a combination of DC and radiofrequency voltages. A small amount of helium is used as a "cooling gas" to help control the distribution of energies of the ions. To perform MS-MS analyses, the trap fills with ions from the source. Then a particular ion of interest is selected and the trap voltages are adjusted to eject ions of all other m/z values. The voltages on the trap then are quickly increased to increase the energies of the remaining ions, which results in energetic collisions of the peptide ions with the helium gas atoms in the trap and induce fragmentation of the ions. The fragments then are caught in the trap and scanned out in according to their *m/z* values (1,7).

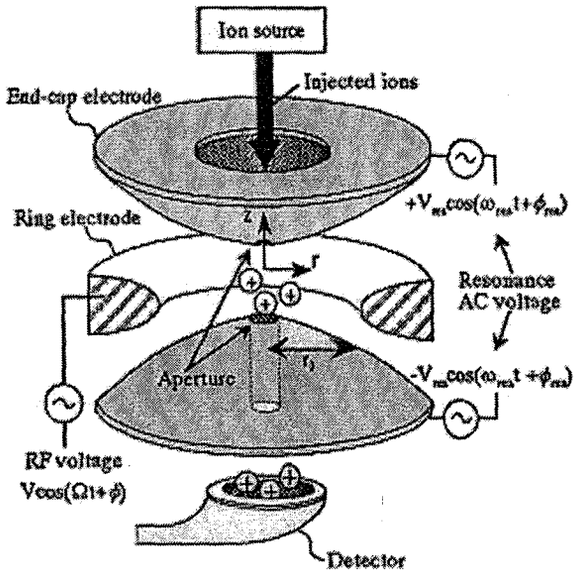


Figure 10. An ion-trap MS instrument (1).

Some instrument control software allows the automated switching of the triple quad or ion trap between full-scan and tandem MS modes to acquire peptide MS-MS spectra. In this approach, the instrument is set by default in full-scan mode to detect peptide ions as they emerge from the source. When peptide ions are detected, the instrument selects the most intense ion and subjects it to collision-induced dissociation (CID) to obtain an MS-MS spectrum. The instrument then switches back to full-scan mode and selects the next most intense peptide ion and subjects it to CID. This switching cycle is repeated to obtain MS-MS spectra of multiple peptide ions automatically.

Peptide-mass fingerprints or product-ion data are used to search a protein-sequence database to identify the protein of interest. The identification is made by comparing the experimentally-generated data with theoretical data calculated for database entry. Database search programs are often included in commercial software packages that are provided with mass spectrometers. One such example is the SEQUEST program that is used for database searching with uninterpreted product-ion spectra. A number of search engines can also be accessed free-of-charge over the Internet, for example the PeptIdent and MultIdent programs at the ExPASy

Molecular Biology server, MS-Fit andMS-Tag at the Protein Prospector server, or MASCOT at the Matrix Science server. These web sites also provide additional proteomics software tools, technical information, and links to other resources (7).

For proteome analysis of human-prostate tissue that are studied by Sarka Beranova-Giorgianni, which is used 2DGE to separate proteins and identification of selected protein spots by MALDI-TOF MS and ESI-Q-IT-MS. Proteins from a whole prostate-tissue specimen were solubilized, separated by 2DGE, and visualized by staining with Coomassie Blue (Fig. 11.). There are approximately 400 protein spots of various intensities are detected in the 2D map. These proteins spots were identified based on peptide-mass fingerprinting or product-ion data. By MALDI-TOF MS analysis, the spectrum contains 24 peptide ions with masses of 0.5-2.6 kDa that represent the peptide-mass fingerprint of the protein spot is shown in figure 12 – a single spectrum was produced from all peptides in the tryptic digest. Moreover, a typical product-ion spectrum obtained by ESI-Q-IT-MS (Fig. 13.). It should be noted that the tryptic digest was first separated by capillary HPLC, and each eluted peptide was introduced online into the mass spectrometer and analyzed. Thus, a product-ion spectrum was recorded for each of the tryptic peptides. (In contrast, in the MALDI-TOF experiment described above, a single spectrum was produced from all peptides in the tryptic digest). The spectrum in Figure 13 displays series of product ions that arise from gas-phase fragmentations of the particular peptide and that are diagnostic of the peptide's sequence. All product-ion spectra generated in the ESI-Q-IT analysis of the digest spot were used in a database search to pinpoint the sequences of the peptides and to identify the protein (7).

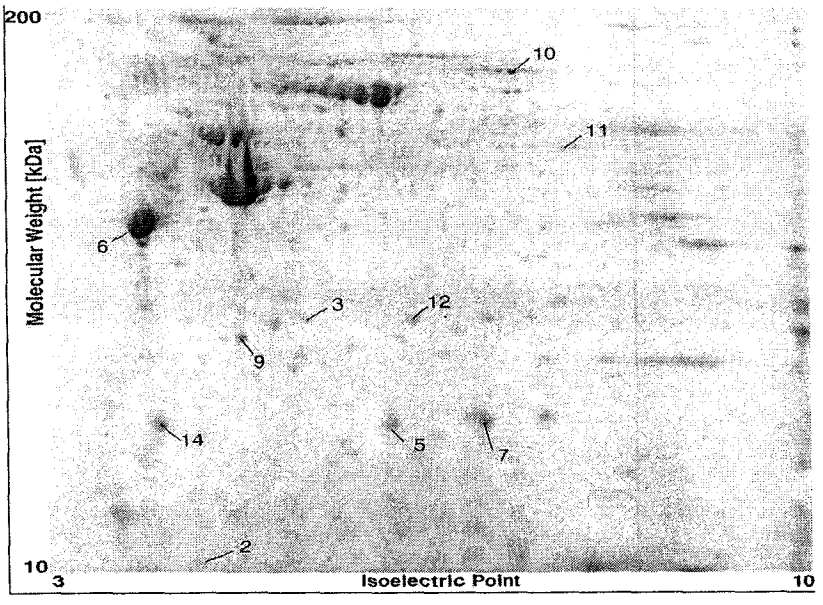


Figure 11. The 2D map of the human-prostate proteome (7).

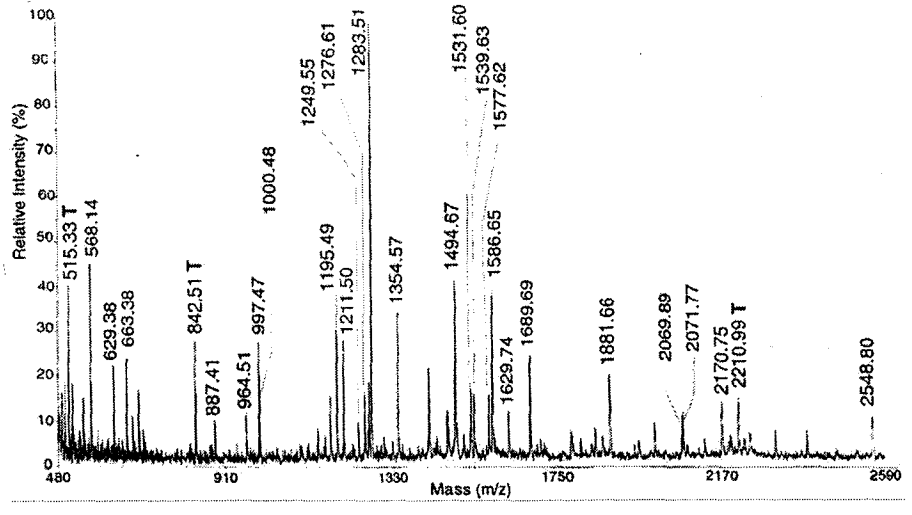


Figure 12. MALDI-TOF mass spectrum of the trypic digest of protein spot 10 from the human-prostate proteome. The letter T indicates trypsin-auto-digestion peptides (7).

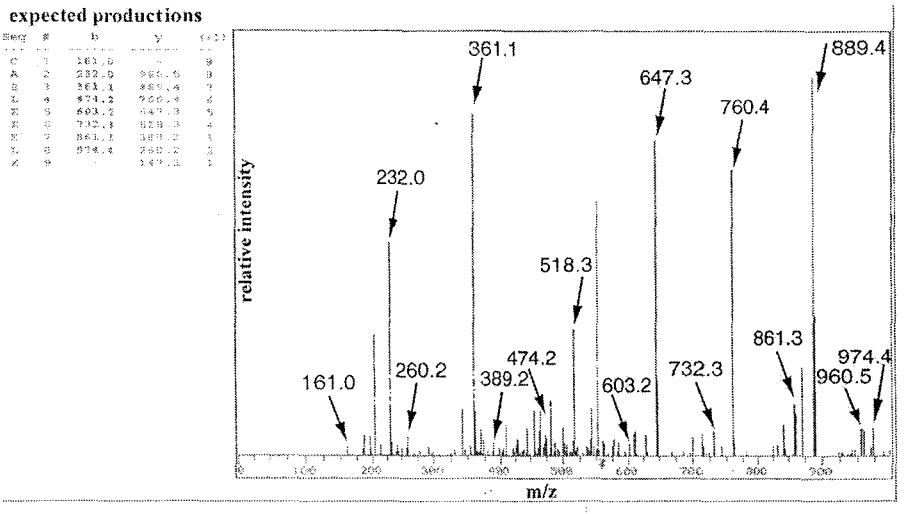


Figure13. Product-ion spectrum of the peptide C*AELEEEK (where C* denotes carbamidomethylcysteine) from the tryptic digest of protein spot 6 from the human-prostate proteome. The data were obtained by ESI-Q-IT-MS (7).

References

1. Liebler DC. Introduction to proteomics: tools for the new biology. Humana Press: Totowa, NJ, 2005.
2. Two dimensional gel electrophoresis of proteins. Nature Methods 2005;2:83-84.
3. http://www.aber.ac.uk/parasitology/Proteome/Tut_2D.html#Section%201
4. http://genome.wellcome.ac.uk/doc_wtd021045.html
5. http://www.bio-rad.com
6. Fatoohi AF, Cozon GJN, Gonzalo P, Mayencon M, Greenland T, Picot S, Peyron F. Heterogeneity in cellular and humoral immune responses against Toxoplasma gondii antigen in humans. Clin Exp Immunol 2004;136:535-541.
7. Beranova-Giorgianni S. Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations. Trends Anal. Chem. 2003;22:273-281.

Topic 20. Protein microarray Technique

Protein microarrays are tools that can be used in many different areas of research, including basic and translational research. Protein arrays are solid-phase ligand binding assay systems using immobilized proteins on surfaces which include glass, membranes, microtiter wells, mass spectrometer plates, and beads or other particles. The assays are highly parallel (multiplexed) and often miniaturized (microarrays, protein chips). Their advantages include being rapid and automatable, capable of high sensitivity, economical on reagents, and giving an abundance of data for a single experiment. Bioinformatics support is important; the data handling demands sophisticated software and data comparison analysis. Fortunately some of the software can be adapted from that used for DNA arrays, as can much of the hardware and detection systems (1).

There are two general types of protein microarray: analytical and functional protein microarrays. Analytical protein microarrays involve a high-density array of affinity reagents (e.g. antibodies or antigens) that are used for detecting proteins in a complex mixture. Functional protein chips are constructed by immobilizing large numbers of purified proteins on a solid surface. Unlike the antibody antigen chips, protein chips have an enormous potential in assaying for a wide range of biochemical activities (protein-protein, protein-lipid, protein-nucleic acid, and enzyme-substrate interactions), as well as drug and drug target identification (2). There are many variations in protein array formats to date (Table 1). A protein array is defined as a microarray in which protein interaction is detected with the aid of miniaturized 'ligand' spotting technology. However, it is important to design studies carefully based on the design of the experiments, the kind of output needed, and the kind of samples that are expected (3).

Table 1. Miniaturized protein array technologies (10).

Printed on matrix	Detection method	Matrix substrate
Antibody	Antibody (ELISA)	NHS[a]
Antibody	Analyte labeling	Nitrocellulose
Antibody	Antibody (ELISA)	Scillicon
Antibody	Antibody (ELISA)	Aminosilanated
Antibody/antigen	Analyte labeling	Poly-L-lysine
Antibody/lysozyme	Direct imaging	Gold
Antibody (microcantellever)	Energy transition	Gold
Purified protein (human)	Analyte labeling	Nitrocellulose
Purified protein	Analyte labeling	Aldehyde
Purified protien (yeast)	Analyte labeling	Aldehyde
Purified protein (yeast)	Antibody	Nitrocellulose
Fractionated protein	Antibody	Various[b]
Whole cell lysate (human)	Antibody	Nitrocellulose
Recombinant protein	Analyte labeling	Aldehyde/Nickel
Peptide	Multiple methods[c]	Gold
Various substrates	Analyte labeling	Poly-L-lysine

[a] N-hydroxysuccinimide.

[b] Poly-amine derivatized, poly-aldehyde, and nitrocellulose slides.

[c] Surface plasmon resonance, fluorescence, and phosphoimaging.

Protein arrays have been designed as a miniaturization of familiar immunoassay methods such as ELISA and dot blotting, often utilizing fluorescent readout, and facilitated by robotics and high throughput detection systems to enable multiple assays to be carried out in parallel. Variables in immobilization of proteins include both the coupling reagent and the nature of the surface being coupled to. The properties of a good protein array support surface are that it should be chemically stable before and after the coupling procedures, allow good spot morphology, display minimal nonspecific binding, not contribute a background in detection systems, and be compatible with different detection systems. The immobilization method used should be reproducible, applicable to proteins of different properties (size, hydrophilic, hydrophobic), amenable to high throughput and automation, and compatible with retention of fully functional protein activity. Orientation of the surface-bound protein is recognized as an important factor in presenting it to ligand or substrate in an active state; for capture arrays the most efficient binding results are obtained with orientated capture reagents, which generally require site-specific labelling of the protein (1).

The most important investment for production of protein microarrays is the arraying robot. The reliability and reproducibility of the fabricated arrays are entirely dependent on the quality and morphology of the microspots produced by the arrayer. All arrayers have three basic components: (4) a source plate location for placing the samples to be arrayed, usually for 384-well plates; (5) a device or head for printing the samples; and (6) a destination zone for placement of the supports to be arrayed, usually 1 x 3-inch (25 x 75-mm) slides. There are basically two types of arraying robot, those that use pin or contact printing and those that use Piezo-electric non-contact printing. Non-contact printing is faster at multi-dispensing a single sample but slower at changing between samples, compared to contact printing. Contact printing has the major advantage that it can produce protein spots of the highest quality. The choice of arrayer type will thus be influenced by whether a few samples are to be arrayed on each slide (2).

Fluorescence labelling and detection methods are widely used. The same instrumentation as used for reading DNA microarrays is applicable to protein arrays. For differential display, capture (e.g. antibody) arrays can be probed with fluorescently labelled proteins from two different cell states, in which cell lysates are directly conjugated with different fluorophores (e.g. Cy-3, Cy-5) and mixed, such that the colour acts as a readout for changes in target abundance. Capture arrays form the basis of diagnostic chips and arrays for expression profiling. They

employ high affinity capture reagents, such as conventional antibodies, single domains, engineered scaffolds, peptides or nucleic acid aptamers, to bind and detect specific target ligands in high throughput manner. Antibody arrays have the required properties of specificity and acceptable background. Antibodies for capture arrays are made either by conventional immunization (polyclonal sera and hybridomas), or as recombinant fragments, usually expressed in E. coli, after selection from phage or ribosome display libraries. The term 'scaffold' refers to ligand-binding domains of proteins, which are engineered into multiple variants capable of binding diverse target molecules with antibody-like properties of specificity and affinity. The variants can be produced in a genetic library format and selected against individual targets by phage, bacterial or ribosome display (1).

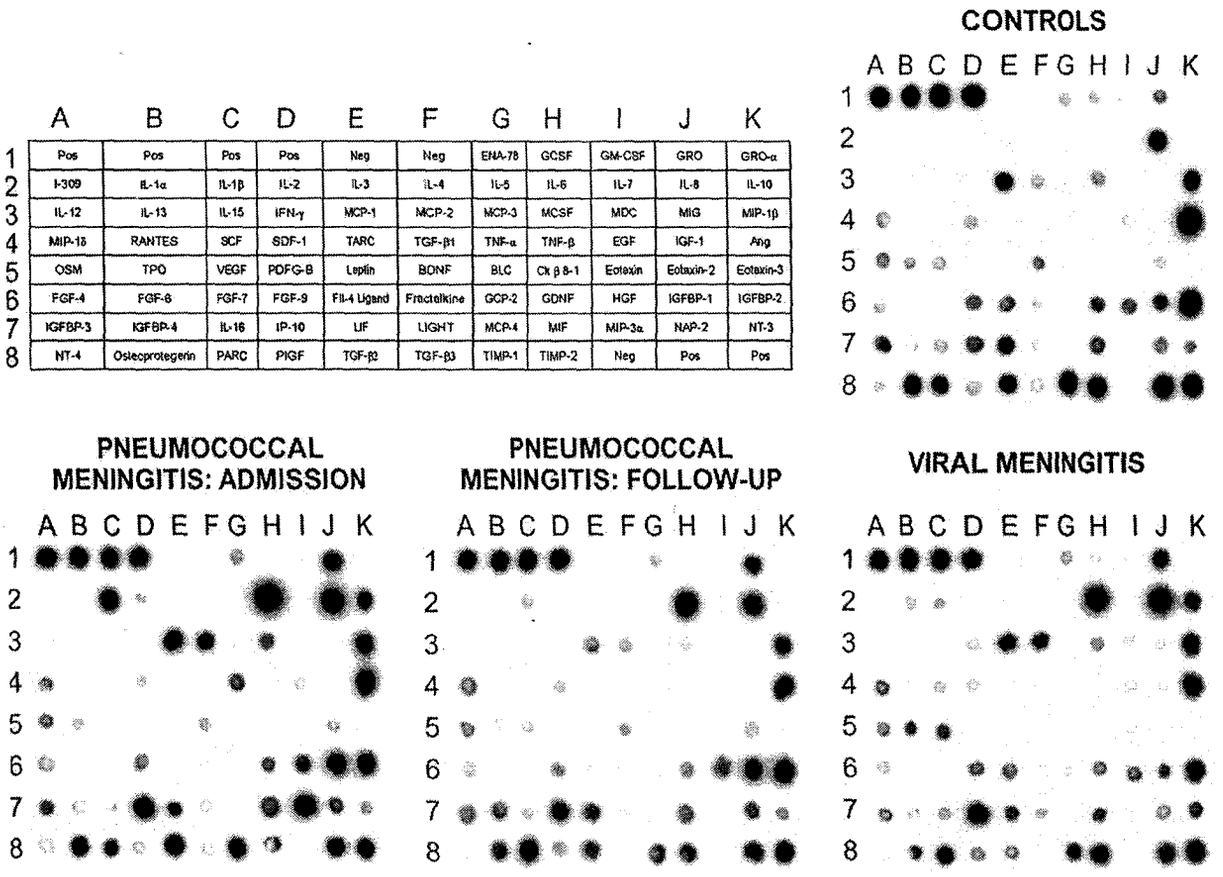
Nonprotein capture molecules, notably the single-stranded nucleic acid aptamers which bind protein ligands with high specificity and affinity, are also used in arrays. Aptamers are selected from libraries of oligonucleotides and their interaction with protein can be enhanced by covalent attachment, through incorporation of brominated deoxyuridine and UV-activated crosslinking (photoaptamers). Photocrosslinking to ligand reduces the crossreactivity of aptamers due to the specific steric requirements. Aptamers have the advantages of ease of production by automated oligonucleotide synthesis and the stability and robustness of DNA; on photoaptamer arrays, universal fluorescent protein stains can be used to detect binding (1).

Protein analyzes binding to antibody arrays may be detected directly or via a secondary antibody in a sandwich assay. Direct labelling is used for comparison of different samples with different colours. Where pairs of antibodies directed at the same protein ligand are available, sandwich immunoassays provide high specificity and sensitivity and are therefore the method of choice for low abundance proteins such as cytokines; they also give the possibility of detection of protein modifications. Label- free detection methods, including mass spectrometry, surface plasmon resonance and atomic force microscopy, avoid alteration of ligand (1).

The question of cross-reactivity is an important one which applies to all ligand binders and particularly to antibodies, being the most popular reagents. While antibodies are thought of as being highly specific, monoclonals can show unpredictable cross-reactions which will be revealed by thorough screening. The ultimate usefulness of individual reagents then depends on the relative level of cross-reaction and specific reaction. The use of sandwich assays, in which antibody pairs are used to bind and detect ligand, may go a long way towards eliminating the

problem, since it is unlikely that both members of the sandwich will exhibit the same cross-reactivity. Polyclonal antibodies are emerging as array reagents for protein expression studies; although they require affinity purification, rabbit sera are easier to produce than monoclonal, and cross-reactions may be reduced as a result of heterogeneity. There are ambitious projects to raise monoclonal antibodies and antisera against the entire human proteome. An important general principle is that, for optimal specificity where assays are highly multiplexed, it is essential to provide dual level target recognition, i.e. two levels of specificity for each locus in the array. Sandwich assays achieve this with two antibodies, photocrosslinking reduces the cross-reactivity of aptamers and MS provides definitive label-free protein identification (1).

The new technology of protein array analysis allows for the simultaneous quantification of a large number of proteins in one sample. Kastenbauer *et al.* have used this method in order to describe patterns of protein expression in the cerebrospinal fluid (CSF) of patients with viral and pneumococcal meningitis. The RayBioi Human Cytokine Array V (http://www.raybiotech.com) is used which detects 79 human cytokines, chemokines, and growth factors. The array membrane contains dots of antigen-specific immobilized antibodies, arranged in 11 columns and 8 rows (Table in Fig. 14). Six dots are coated with biotin-conjugated IgG (positive controls) and three are uncoated (negative controls). The antigen-specific immunoreactivity is detected with biotin-conjugated soluble antibodies and horseradish peroxidase-conjugated streptavidin. Densitometry of chemiluminescence-exposed X-ray films was used for quantification. In order to normalize the results, the optical densities of each dot were then expressed as percentage of the average optical densities of the 6 positive controls contained on each array membrane. Because the protein array detects only relative expression levels and not absolute values, there is no defined lower detection limit. The distinct pattern of protein expression in viral meningitis and in the acute stage and during recovery from pneumococcal meningitis are demonstrated in figure 14 (7).



	A	B	C	D	E	F	G	H	I	J	K
1	Pos	Pos	Pos	Pos	Neg	Neg	ENA-78	GCSF	GM-CSF	GRO	GRO-α
2	I-309	IL-1α	IL-1β	IL-2	IL-3	IL-4	IL-5	IL-6	IL-7	IL-8	IL-10
3	IL-12	IL-13	IL-15	IFN-γ	MCP-1	MCP-2	MCP-3	MCSF	MDC	MIG	MIP-1β
4	MIP-1δ	RANTES	SCF	SDF-1	TARC	TGF-β1	TNF-α	TNF-β	EGF	IGF-1	Ang
5	OSM	TPO	VEGF	PDFG-B	Leptin	BDNF	BLC	Ck β 8-1	Eotaxin	Eotaxin-2	Eotaxin-3
6	FGF-4	FGF-6	FGF-7	FGF-9	Flt-4 Ligand	Fractalkine	GCP-2	GDNF	HGF	IGFBP-1	IGFBP-2
7	IGFBP-3	IGFBP-4	IL-16	IP-10	LIF	LIGHT	MCP-4	MIF	MIP-3α	NAP-2	NT-3
8	NT-4	Osteoprotegerin	PARC	PIGF	TGF-β2	TGF-β3	TIMP-1	TIMP-2	Neg	Pos	Pos

Fig. 14. Cerebrospinal fluid protein array analysis of 79 cytokines, chemokines, and growth factors in pooled samples (10 patients in every diagnostic group) of controls (patients with non-inflammatory diseases), patients with pneumococcal meningitis on admission and during follow-up, and in patients with viral meningitis (7).

References

1. Beranova-Giorgianni S. Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations. Trends Anal. Chem. 2003;22:273-281.
2. http://www.functionalgenomics.org.uk/sections/resources/protein_arrays.htm#definig
3. Kambhampati D. Protein microarray technology. Weinheim: Wiley-VCH, 2004.
4. Liebler DC. Introduction to proteomics: tools for the new biology. Humana Press: Totowa, NJ, 2005.
5. Two dimensional gel electrophoresis of proteins. Nature Methods 2005;2:83-84.
6. http://www.aber.ac.uk/parasitology/Proteome/Tut_2D.html#Section%201
7. Kastenbauer S, Angele B, Sporer B, Pfister HW, Koedel U. Patterns of protein expression in infectious meningitis: A cerebrospinal fluid protein array analysis. J Neuroimm 2005;164:134-139.