Sample preparation and characterization

Rob Meijers
EMBL Hamburg

EMBO Global Lecture Course
Hyderabad 2012
Garbage in/garbage out
Sample preparation: under control?

PSI October 2011
Construct design

• ALWAYS try to make a model YOURSELF
• BLAST the sequence against PDB
• Align sequence
• Use programs to model:
  • SWISSMODEL
  • FFAS
  • Modeller
  • Rosetta

Models available for E.coli proteome

From Torsten Schwede
Protein modifications to stabilize proteins

- Remove flexible bits
- Mutate surface residues
- Change species
- Lysine methylation
- Use mFABs
- Coexpress with partner
Extreme engineering

- HIV envelope protein gp120:
  - Many clades tested
  - Truncated loop

- CD4 receptor:
  - Truncated construct
  - Deglycosylated

- +mAB!

Kwong et al. (1998) *Nature*. 393, 648-659
Cloning & Expression systems

- Synthetic genes:
  - No cDNA artifacts
  - No spurious restriction sites
  - Codon optimized

<table>
<thead>
<tr>
<th>Method</th>
<th>Source</th>
<th>Type of proteins</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>bacterial</td>
<td>Bacterial/membrane</td>
<td>$</td>
</tr>
<tr>
<td>Pichia</td>
<td>yeast</td>
<td>Cytosolic (but avoid post-translational modifications)</td>
<td>$</td>
</tr>
<tr>
<td>Baculo</td>
<td>Insect cells</td>
<td>any</td>
<td>$$$</td>
</tr>
<tr>
<td>HEK</td>
<td>human</td>
<td>mammalian</td>
<td>$$$</td>
</tr>
</tbody>
</table>
E. Coli expression at EMBL

- pETM vector series
- Restriction site cloning
- Different purification tags
Why HEK 293 cell expression?

- Glycosylation
- Human cell line
- pXLG vector
Standard Purification strategy

Dialysis + cleavage

IMAC

Elution

Flow through

3 hours

SEC

Dialysis + cleavage

3 hours
Affinity tags

- His tag
- GST-tag
  - (also solubility tag)
- STREP-tag
- Myc tag
- FLAG tag
Protease cleavage

- Thrombin
  - Natural source
  - Not very specific
- TEV
  - Cheap recombinant
  - Good for N-terminal tag
  - Cleaves at 4 degrees
- SUMO
  - Solubility domain
  - SUMO protease
  - Clear cut
- 3C Protease
- Enterokinase
  - Clean cut
  - Very expensive
The right stuff

- Molecular weight
- Oligomerization state
- Monodispersity
- Aggregation
- Protein concentration
- Protein folding state
Quality control

SDS Page

Mass spec

Circular Dichroism

Gel filtration

DLS/SLS
A gel of a heterodimeric receptor
A size exclusion profile
A second look

<table>
<thead>
<tr>
<th>Dilute</th>
<th>Concentrate 12</th>
<th>d12</th>
<th>Non-reduced 1</th>
</tr>
</thead>
</table>

[Image of a gel blot with bands indicated by the labels.]
Different gels

- SDS
- SDS non-reduced
- Native
- Iso-electric focussing
- Purity
- Disulphide bonds
- Post-translational modifications:
  - Phosphorylation
  - glycosylation
Another capricious sample

7XCYS ORIGINAL SEQUENCE, WITH HIS-TAG
MGSSHHHHHHHHSSGLVPRGSHMKICITVGHSLKSGACTSADGVVNEYQYNKSLAPVladTFRKEGHKVDVlICPEKQFKTKNeeKSyKIPRvNsgGYDLLIHELNLNASNGQQGKGSEvLYYSNKGLEYATRIICDKLGTVFKNRGAKLDKRLYILNSSKPTAVLIESFFCDNKEKDkAKKLGHEGIaKLIVEGVLNKnINNNEGvkQMkykHTIVYDGEvDKisATVVGwgyNdGkILICDIKDyVPGQtnLyvVGGGACeKiSSITKEKFIMIKGNDRFDTLYKALDFIN

SDS-PAGE 12%: Size Exclusion peaks 1 (P1) and 2 (P2) under reducing and non-reducing conditions.
Sizing profile

Add DTT
Cysteine point mutations
Sizing profile
Protein concentration

- UV-Vis
- Bradford
- Refractive index
Quality control

SDS Page

Mass spec

Gel filtration
Mass spectrometry

- MALDI-TOF or ESI
- Integral sample
- Proteolytic digest:
  - Trypsin
  - Peptide mass more accurate
  - Sequencing
Mass spectrometry

- MS/MS + Ion mobility:
  - Detailed folding state
  - Protein-protein interactions
  - Whole protein size...
Ion mobility derived particle size

in combination with SAXS…

Limited proteolysis

- Incubate with protease
- SDS-PAGE
- Mass spec
- Edman degradation
Sample optimization

- Reduced Gel problematic: change purification protocol
- NR Gel problematic: check cysteines
- Protein aggregation, folding stability:
  - Size exclusion, light scattering, CD, NMR, **thermofluor**
Thermal stability

- Thermofluor
- Modified real-time PCR machine
  - Add hydrophobic fluorescent probe
  - When protein unfolds…
  - Fluorescence increases
Thermal stability

- Check protein stability
- Additive/ligand screen

Intramolecular interference produces a disymmetry in the scattered light.

Size of molecule/particle must be significant compared to wavelength of light.
Angular Dependence

small molecules
radius < 15 nm
\( P_\theta = 1 \) for all \( \theta \)

large molecules
radius > 15 nm
\( P_\theta = 1 \) for \( \theta = 0^\circ \)
Normalised LS signals show no angular dependence for proteins
- Molecular weight requires only RALS
- Can not measure size by light scattering alone
Combined particle analysis

- UV
- Refractive index
- Viscosity
- Right angle light scattering

Viscotek/Malvern
Laser Light Scattering Detector, Refractive Index Detector and UV-Cell
SEC + SAXS

• SWING beamline at Soleil
• HPLC in FPLC mode
Under construction:

- Automatic sample loading
- Cooling of the system
Integration of experimental data

- SAXS scattering curve
- Accurate protein concentration (UV-VIS +RI)
- Molecular weight (RALS)
- Radius of gyration? (MALS)
- Hydrodynamic radius (DLS)
- Dynamic complex formation (Thermophoresis)
Estimation of Particle Radius
Stokes-Einstein relation

- $D = \text{Diffusion coefficient}$
- $k = \text{Boltzmann’s coefficient}$
- $T = \text{Temperature}$
- $\eta = \text{Viscosity}$
- $R = \text{hydrodynamic radius}$

$$D = \frac{kT}{6\pi \eta R}$$
Dynamic light scattering
Dynamic Light Scattering scheme

Laser (λ=500nm) → Sample → Diffused Light → Detector → Transmitted light → Correlator → Graph
Information from the correlation curve

The time at which the correlation of the signal starts to decay gives information about the mean diameter.

The baseline gives information about the presence of large particles/aggregates.

The angle of decay gives information about the polydispersity of the distribution.
Integration of experimental data

- SAXS scattering curve
- Accurate protein concentration (UV-VIS +RI)
- Molecular weight (RALS)
- Radius of gyration? (MALS)
- Hydrodynamic radius (DLS)
- Dynamic complex formation (Thermophoresis)
Thermophoresis

- Heat sample by 2K
- Molecules dissipate
- When fluorescently labelled
- Equilibrate depending on size of hydration shell, charge distribution

Duhr and Braun (2006) PNAS 103, pp19678
Current technology

Jerabek-Willemsen et al. 2011
Sample optimization

- Reduced Gel problematic: change purification protocol
- NR Gel problematic: check cysteines
- Protein aggregation, folding stability:
  - Size exclusion, light scattering, CD, NMR, thermofluor, limited proteolysis
- Modify buffers, additives
- If nothing works: change construct
Using SAXS to test constructs

• Large protein (1160 residues)
• Check whether construct is globular
• Removed last 160 residues

Lamers et al. (2006) Cell 126, 881
Additives?

- DTT (will affect OD 280nm)
- Glycerol (no more than 5 %)
- Detergents at less than 2xCMC
  - 0.1% 1-s-Nonyl-β-D-thioglucoside
  - 0.2% n-Decanoylsucrose
  - 0.3% n-Nonyl-β-D-maltoside
  - 0.4% DDAO
  - 0.5% C8E5
  - 0.8% FOS-Choline®-10
  - 1.1% FOS-Choline®-9
Conclusions

• Quality Control is tedious but necessary
• It helps you to pick the right fight
• Some quality control methods can provide useful complementary data