Why chaperone vectors?

A protein folding initiative

An open discussion with structural biologists

Protein Structure Initiative: Pilot Phase

"Whether the pilot phase achieved its goal depends on how we measure success" **(Editorial, PSI-phase 1 and beyond, Nature Structural & Molecular Biology, 11, p. 201, March 2004)**

- **Significant increase in throughput of determination of novel protein structures**
- **PSI structures are dominated by structures of single domains primarily from procaryotic proteins**
- **Proteins which "misbehave" are left aside, which leaves out nearly all important proteins for understanding human diseases and for developing cures**

"…it is still unclear how the bottlenecks for eucaryotic and membrane protein structure determination will be overcome" (ibid)

Protein Structure Initiative: Update

- *"The availability of suitable recombinant protein is still a major bottleneck in protein structure analysis."* **(Bussow K. et al., Microb Cell Fact., 4: 21, 2005)**

- *"Contrary to the popular assumption, the rate of growth of structural data has slowed, and the Protein Data Bank (PDB) has not been growing exponentially since 1995."* **(Levitt M., Proc Natl Acad Sci USA, 104: 3183-3188, 2007)**

- **PDB growth rates are steadily decreasing since 1997.**

- **Proteins which "misbehave" are left aside, which leaves out nearly all important proteins for understanding human diseases and for developing cures**

- **The number of novel structures is growing largely through computation.**

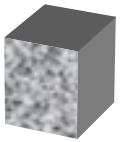
- **Such structures do not provide sufficient resolution to allow drug discovery through "docking" of potential drug candidates**

Inherent PSI problems

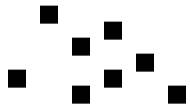
- **Many eucaryotic proteins require specific sets of molecular chaperones for their folding**

- **Chaperones are required at compatible concentrations with the overexpressed proteins, however the recombinant protein synthesis exceeds the host cell protein folding capacity**

Protein overexpression → **Misfolded protein**



Inclusion bodies



Protein degradation products

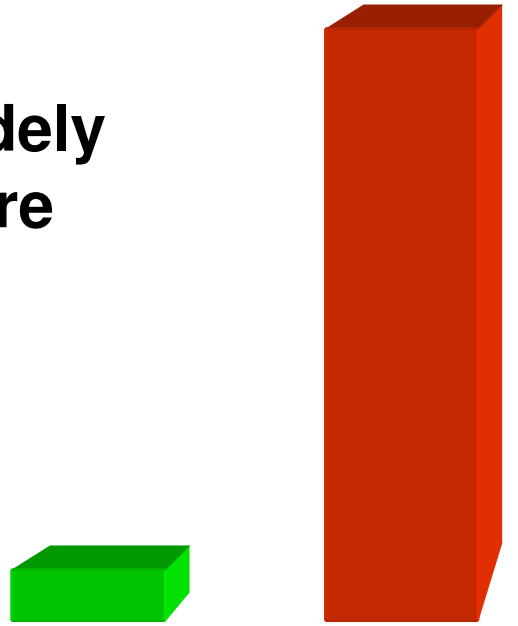
- *Proteins need to be overexpressed, since significant amounts of pure proteins are required for structural studies*

- *An overexpressed protein could be misfolded in any, even a mammalian system, unless the chaperones are co-expressed at comparable with the protein expression level*

Such problems are not unique to structural biology

★ **Only ~10% of human drug target proteins produced in the most widely used expression system, *E.coli,* are suitable for drug screening**

Protein Expression Consortium
(www.lifesensors.com/alliances/pxconsortium/pdf)



Common problems…

- Insolubility
- low expression level
- proteolysis
- poor antigenicity
- cytotoxicity
- poor biological activity

…The root cause is target protein misfolding

Solution to protein misfolding

Chaperone Vectors

Overexpression of both a recombinant protein and a specific set of human molecular chaperones

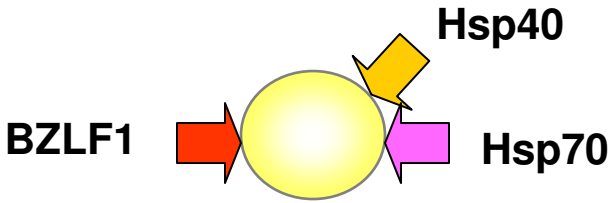
Chaperone-assisted protein folding

Correctly folded recombinant protein

Prior art: Improvement in protein folding using molecular chaperones

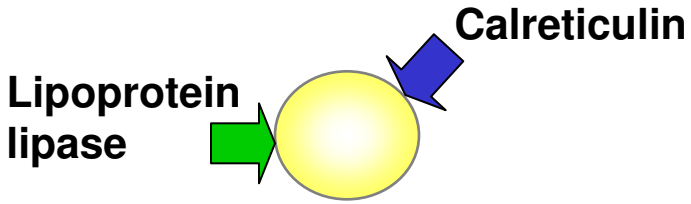
Experiments in insects cells co-infected with recombinant baculoviruses expressing human target proteins and human chaperones

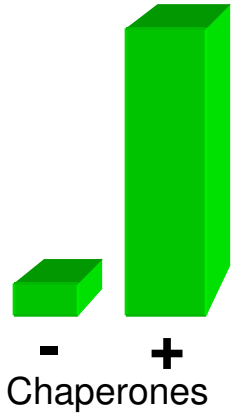
Effect of Hsp40 and Hsp70 chaperones on folding a target protein



8-fold increase
in yield of soluble cytoplasmic target protein BZLF (Yokohama et al., Biochim Biophys Acta, 1493: 119-124, 2000).

Effect of Calreticulin on folding target glycoproteins



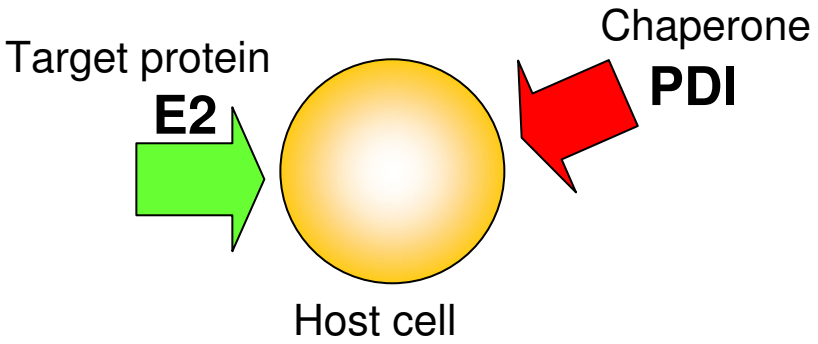


9-fold increase
in lipoprotein lipase enzymatic activity (Zhang et al., J. Biol. Chem., 278: 29344-51, 2003).

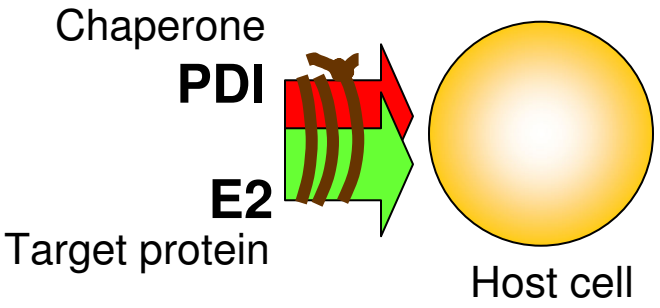
2-3 fold increase in production of secreted soluble HLA-DR4 tetramers (Fourneau et al., J. Immunol. Methods, 285: 253-64, 2004).

Chaperones' effect is enhanced if they are delivered in the same vector with target protein (chaperone vector)

2 vector system



Chaperone Vector

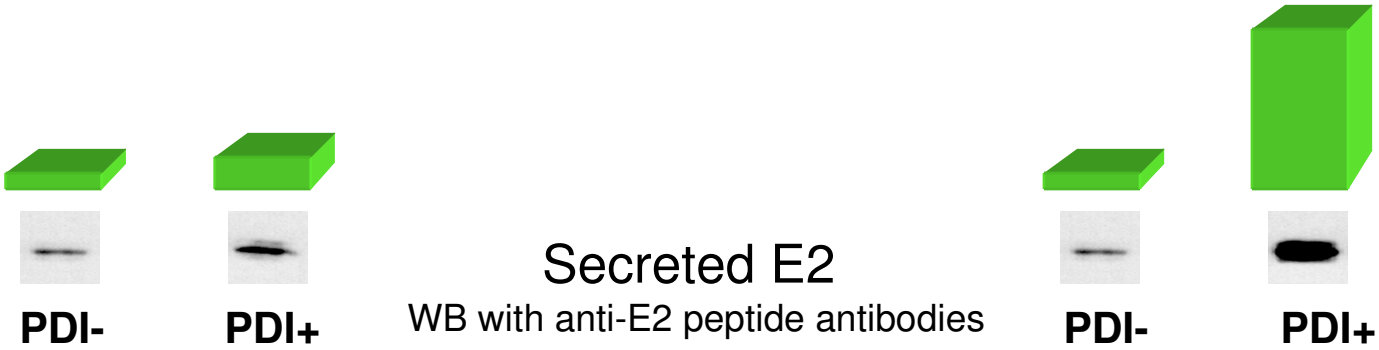


Modest positive effect (2-fold)

Target protein and chaperone are delivered into host cells from different vectors

Strong positive effect (8-fold)

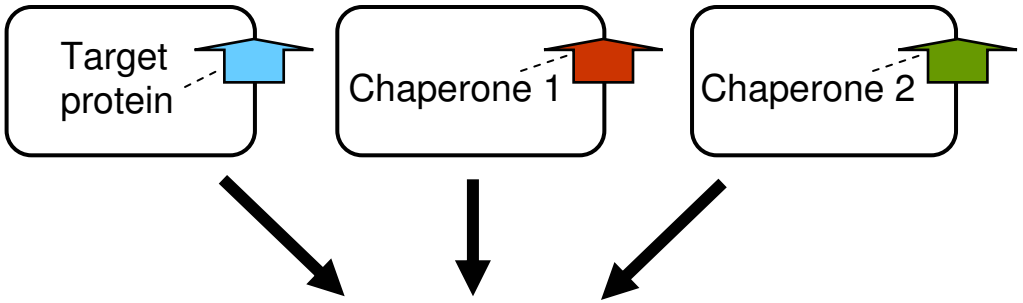
Target protein and chaperone are delivered into host cells from the same vector



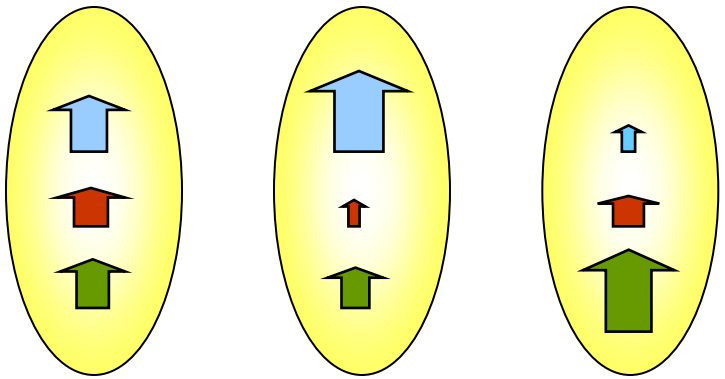
PDI was reported to have modest positive effect on a target protein folding in a 2 vector system (Hsu et al., Protein Expr. Purif., 7, 281-8, 1996). However, the positive effect becomes more pronounced if both PDI and the target protein are co-expressed using the same recombinant baculovirus (chaperone vector). In our experiments, improvement in secretion of cysteine-rich E2 glycoprotein was only about 2 times when insect cells were co-infected with 2 recombinant baculoviruses-one expressing E2 glycoprotein and another expressing human PDI. However, improvement was much more significant (about 8-fold) when both the target protein and the PDI were expressed from the same recombinant baculovirus (Belyaev A.S., unpulished).

Prior Art

Multi-vector system



Co-introduce vectors into host cells



Good protein yield and folding in some of the cells

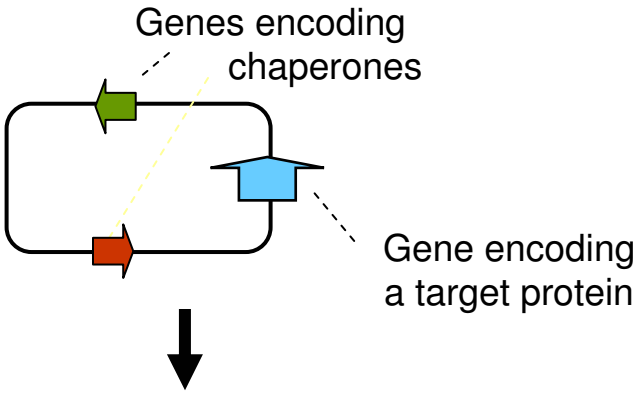
Poor folding

Poor yield

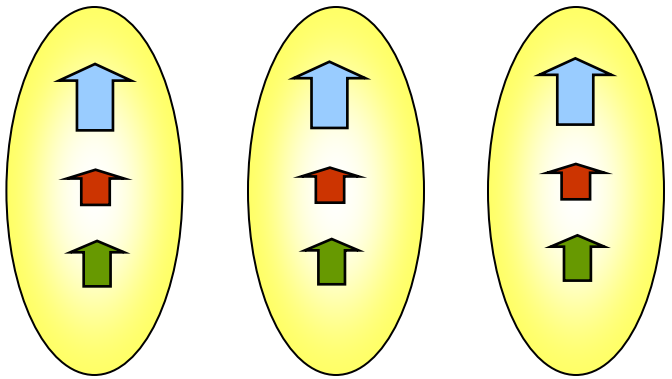
AB Vector

U.S. Patent No. 7,226,781

Chaperone vector



Introduce vector into host cells



Good yield and folding of a target protein in every cell of the population

Why chaperone vectors are more efficient

Multi-vector system

★ *Unbalanced*

Some host cells provide for good target protein yield and folding

Variable production of chaperons and target protein in individual cells as:

a) Individual host cells receive variable number of vectors
b) The vectors which enter the cells earlier are replicating faster than the vectors, which enter the cells later

★ *Wasteful*

Host cell resources are diverted for replication of several vectors

★ *Complex*

A combination of several vectors are used to express molecular chaperones and target protein

Chaperone vector

★ *Balanced*

All host cells provide for good target protein yield and folding

Guaranteed synthesis of target protein and molecular chaperones at the defined ratio in all the cells of the population

★ *Economical*

Host cell resources are more efficiently employed for the synthesis of target protein and chaperones. The waste is minimized as there is only one vector

★ *Simple*

One vector is used to express molecular chaperones and target protein

Chaperone vector design

Vector selection:

Baculovirus vectors

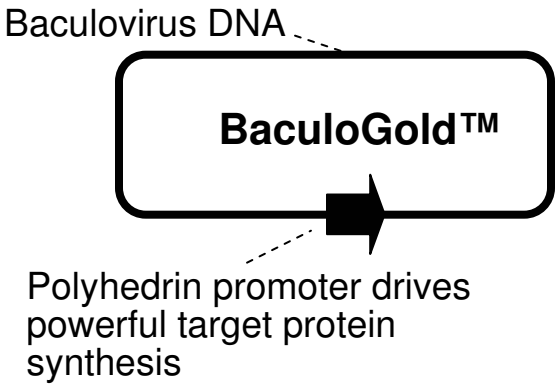
- Nearly as powerful as *E.coli*
- Higher eucaryotic system (insect cells)
- Well developed technology

Chaperone selection:

Major human chaperones with demonstrated capacity for target protein folding in insect cells

- **Hsp40** and **Hsc70** — Major cytoplasmic chaperones, work in concert
- **Calreticulin-** glycoprotein folding
- **Bip-** non-glycoprotein folding in the ER
- **PDI-** correct formation of disulfide bonds

Conventional vectors

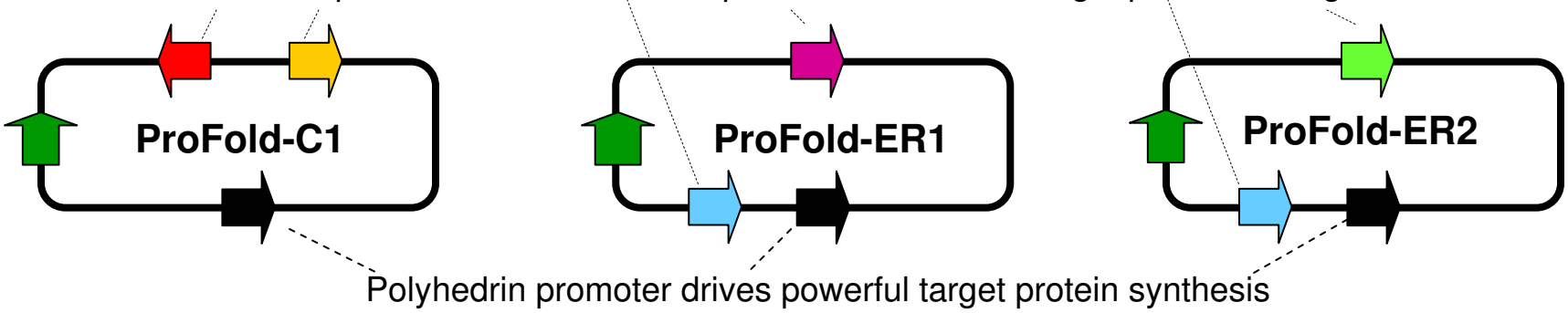


Marker selection:

GFP- convenient operation

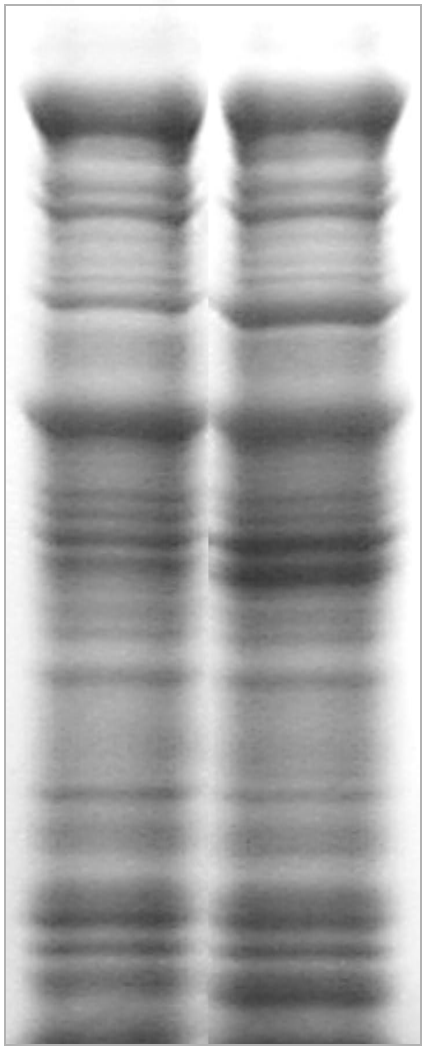
Humanized chaperone vectors

Powerful expression of molecular chaperones to enhance target protein folding



ProFold™-C1 – for protein folding in cytoplasm

BacPAK6 ProFold-C1



Target protein

Hsc70

Hsp40F

Hsp40 (dj2)

GFP

Target protein is expressed as efficiently as using parent BaculoGold-BacPAK6 vectors.

Optimal human Hsp40/Hsp70 pair for enhanced folding of cytoplasmic proteins.

Farnesylated Hsp40 located at the cytosolic surface of the ER. Folding of endoplasmic domains of transmembrane proteins (ion channels, GPCRs).

Non-farnesylated Hsp40 located in the cytoplasm.

GFP marker allows easy monitoring of virus propagation, titration, cell sorting.

Protein expression profiles obtained using conventional BacPAK6 vector and ProFold™-C1 chaperone vector. SDS-PAGE of cell extracts infected with recombinant baculoviruses. Coomassie blue staining.

ProFold™-ER1 – vector for glycoproteins

BacPAK6 ProFold-ER1

Target protein is expressed as efficiently as using parent BaculoGold-BacPAK6 vectors.

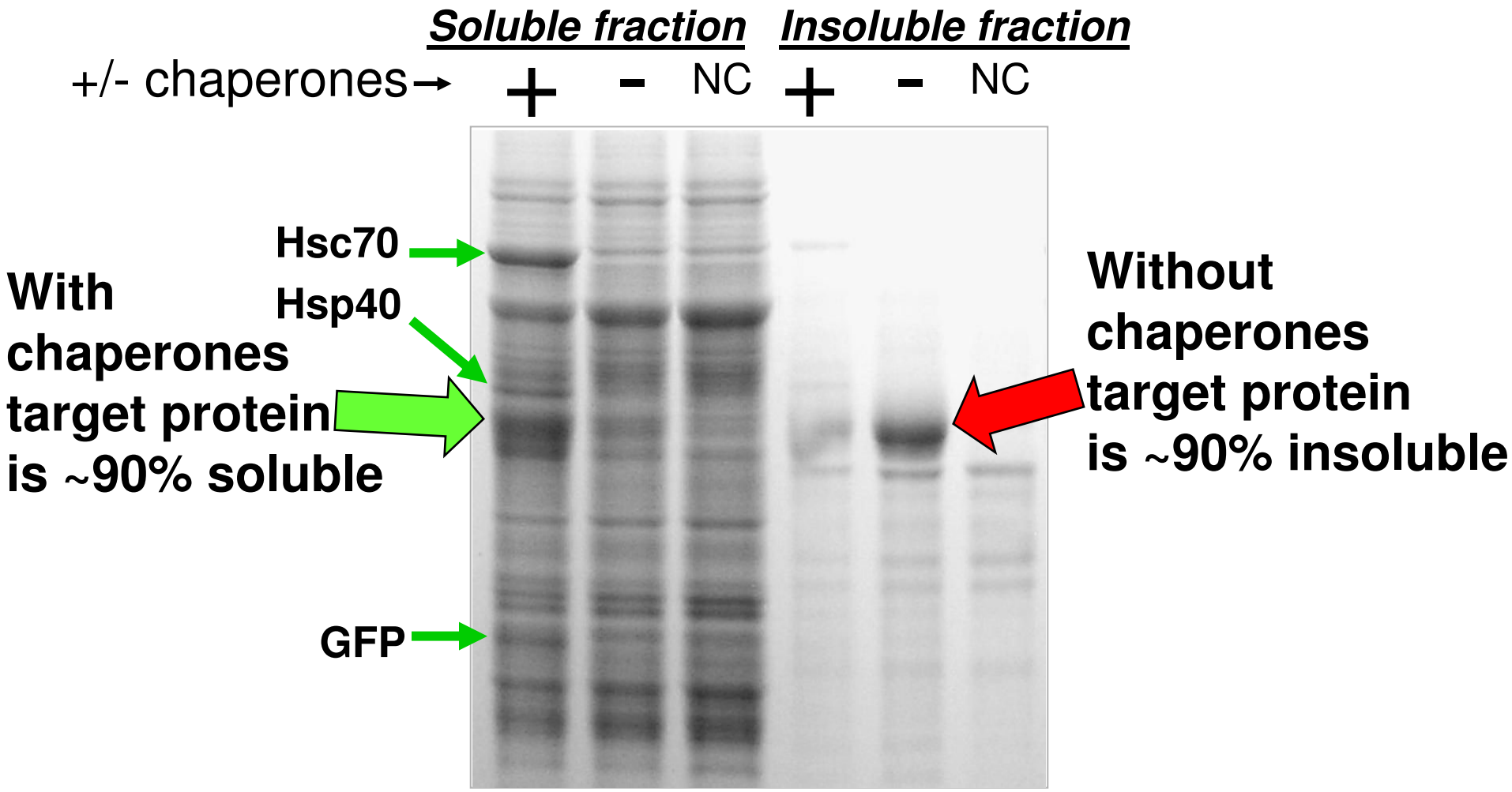
PDI drives correct formation of disulfide bonds and assists protein folding. Human PDI can improve yield of secreted target proteins from mammalian, insect, yeast and even *E.coli* cells.

Calreticulin is a major chaperone assisting membrane or secreted glycoprotein folding. It is secreted from insect cells reaching up to 200 mg/l of the protein in the cell culture media, thus driving production of a co-expressed secreted glycoprotein (Fourneau et al. 2004).

GFP marker for convenient operation.

Protein expression profiles obtained using conventional BacPAK6 vector and ProFold™-ER1 chaperone vector. SDS-PAGE of insect cell extracts infected with recombinant baculoviruses. Coomassie blue staining.

Improvement in target protein folding using a chaperone vector



Extracts of insects cells infected with recombinant baculoviruses. SDS-PAGE, Coomassie blue staining. A cytoplasmic target protein was expressed using ProFold™-C1 (+ chaperones) or using a conventional vector BacPAK6 (-chaprerones). NC-cells infected with negative control recombinant baculovirus, which does not express target proteins, chaperones or markers.

How chaperone vectors compare to conventional vectors?

- The same powerful expression of target proteins as conventional vectors
- The same convenience of making recombinant clones using standard kits
- The same tags for protein purification and detection are available
- More convenient to operate due to the GFP marker

★ ***Improved quality of your target protein and even more convenience***

★ ***Any reason not to switch to the chaperone vectors?***

Concluding remarks

- ★ **Chaperone vectors are commercially available**
  (www.abvector.com)

- ★ **Ready to go, just insert ORF encoding target protein into convenient cloning sites**

- ★ **Collaborations are welcome with the focus on "misbehaving" novel eucaryotic proteins**