

Kód předmětu: C8980



MASARYKOVA UNIVERZITA

Protein expression and purification

- V. Protein expression

Lubomír Janda, Blanka Pekárová,
Radka Dopitová, Jozef Hritz and Adam Norek

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

V. Protein expression

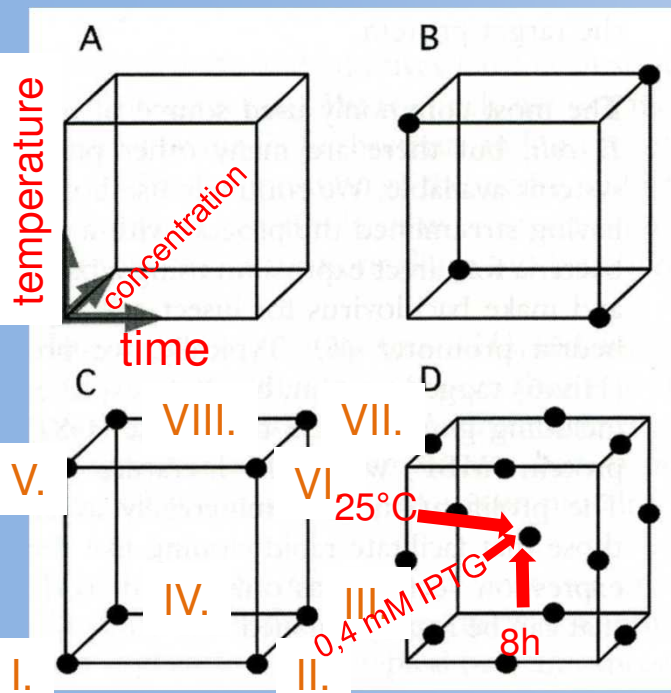
5.1. Designing experiments for high-throughput protein expression

High-throughput platform requires:

- Automation
- Miniaturization
- Quantitative management tools (to identify trends and relationships)

Experimental design:

- An ill-defined experiment will often produce ambiguous results and fail to reach any conclusion.
- Analysis of quantitative response allows the experimenter to optimize conditions critical to production of a soluble protein.
- Performing one-factor-at-a-time experiments raises the risk of locating a local maximum (missing the actual best conditions).



A: One factor at a time
B: Fractional factorial
C: Full factorial
D: Response surface model
(Box-Behnken design for three factors)

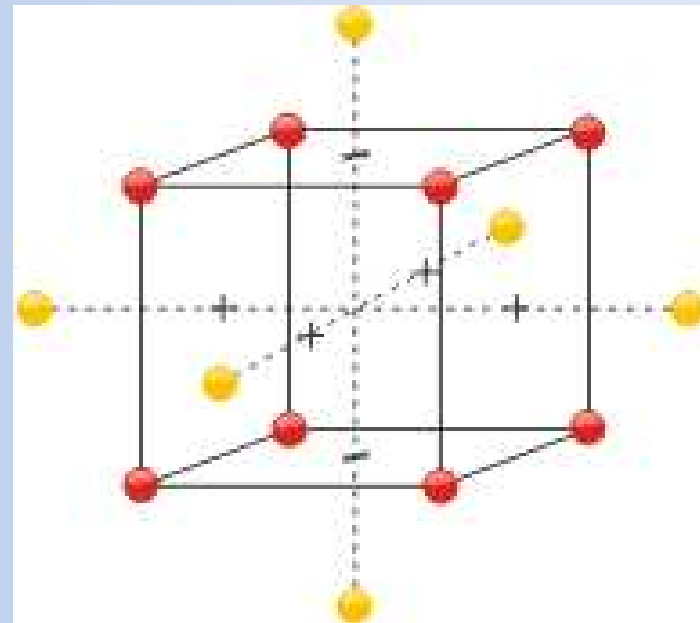
- I. 15°C/1h/0,1mM IPTG
- II. 15°C/24h/0,1 mM IPTG
- III. 15°C//24h/1mM IPTG
- IV. 15°C/1h/1mM IPTG
- V. 37°C/1h/0,1 mM IPTG
- VI. 37°C/24h/0,1 mM IPTG
- VII. 37°C/24h/1mM IPTG
- VIII. 37°C/1h/0,1 mM IPTG

Experimental design (commonly referred to as DOE) is a useful complement to [multivariate data analysis](#) because it generates “structured” data tables, i.e. data tables that contain an important amount of structured variation. This underlying structure will then be used as a basis for multivariate modeling, which will guarantee stable and robust models.

More generally, **careful sample selection** increases the chances of extracting useful information from the data. When one has the possibility to actively perturb the system (experiment with the variables), these chances become even greater. The **critical part** is to decide 1) **which variables to change**, 2) **the intervals for this variation**, and 3) **the pattern of the experimental points**.

- 1) Temperature
- 2) 15°C-37°C
- 3) About 1°C; 5°C or 10°C

Temperature	15°C-37°C~23 culture conditions
IPTG	0,1 – 1,0 mM ~10 culture conditions
Time	1h-24h ~25 culture conditions
	23x10x25=5750 culture conditions



V. Protein expression

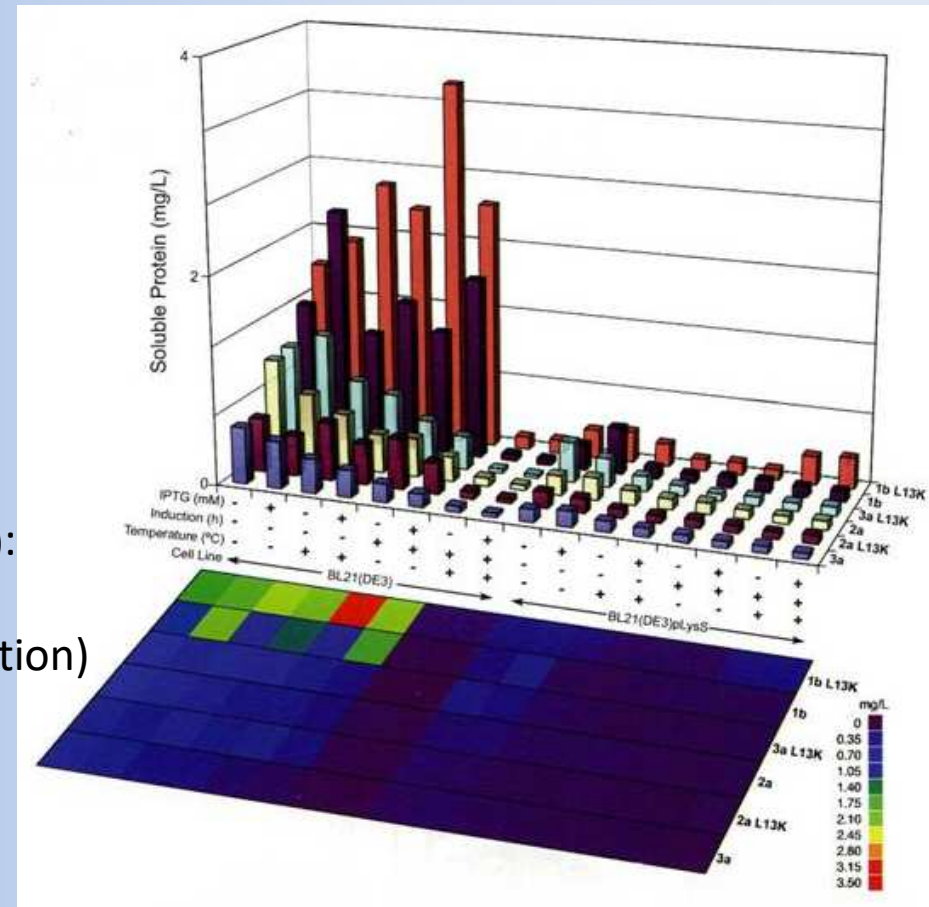
5.1. Designing experiments for high-throughput protein expression

Factors affecting expression:

- Construct - Cat
- Expression system and vector - Cat
- Cell line (host strain) - Cat
- Temperature and time - Cont
- Media - Cat
- Additives - Cat

Full factorial design (16 conditions per construct):

- three continuous factors (temperature, time, IPTG concentration)
- one categorical (host strain)

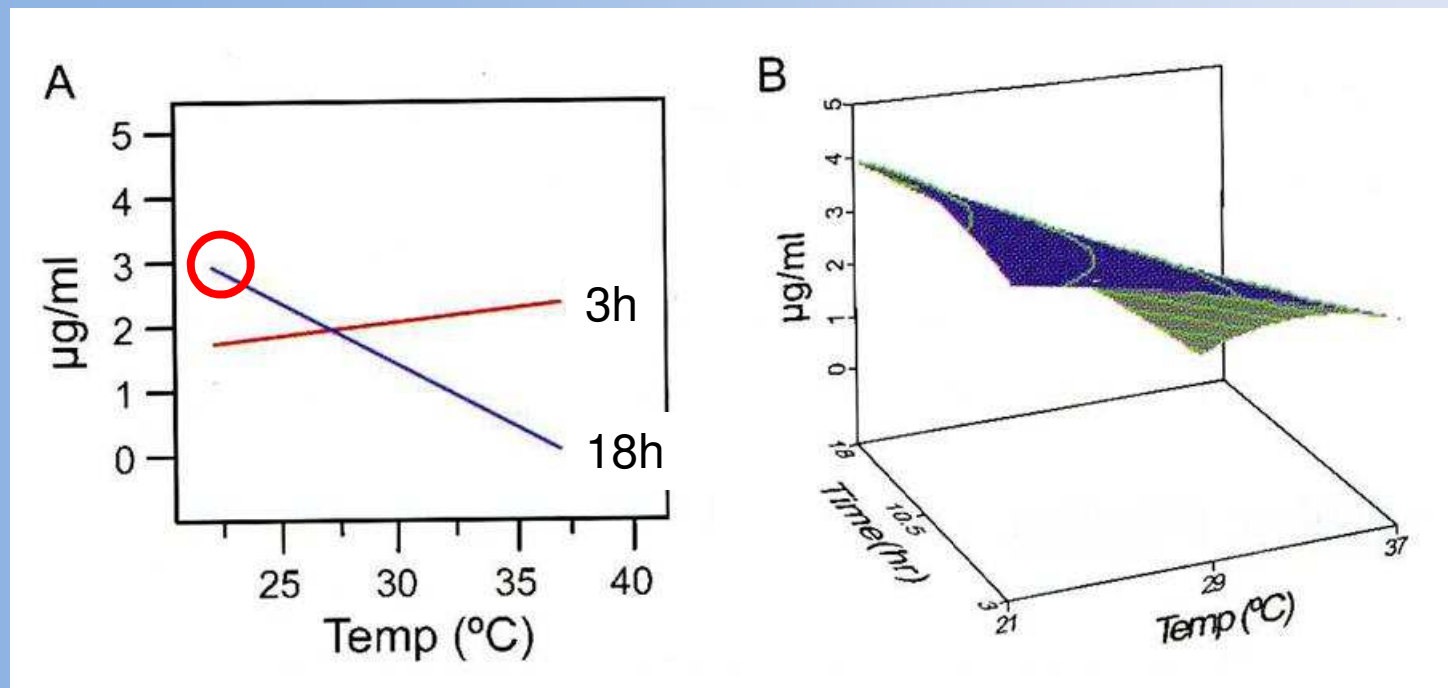


V. Protein expression

5.1. Designing experiments for high-throughput protein expression

Response surface model:

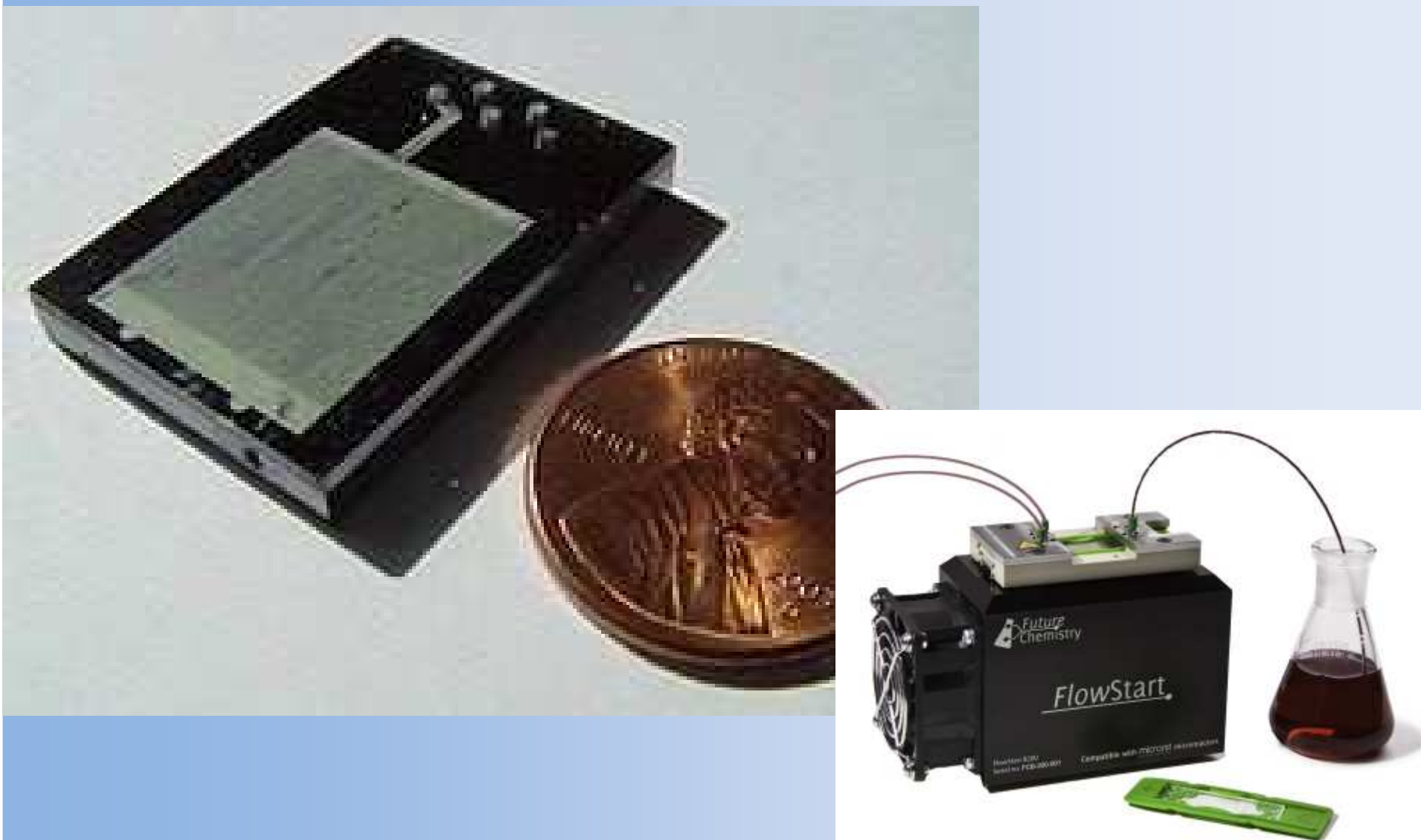
- fine-tunes the conditions
- capable to identify minimum or maximum



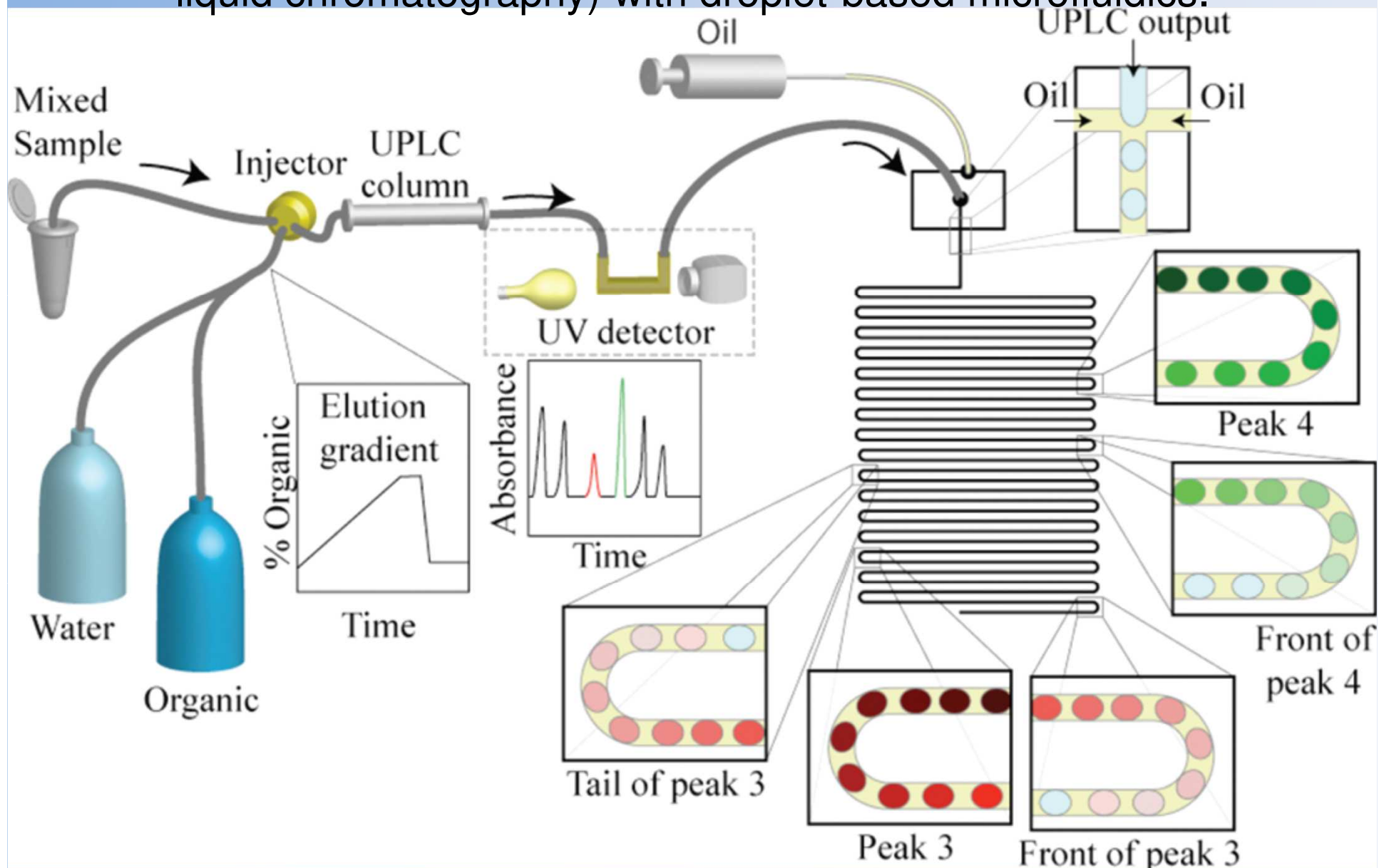
Design of experiment is merely a statistical tool, a means to an end.

It does not guarantee success and cannot replace technical expertise or creativity in experimental work.

Microreactor technologies developed at [LLNL](#) use micromachining techniques to miniaturize the reactor design.



Schematic diagram of integration of nano-UPLC (ultra performance liquid chromatography) with droplet-based microfluidics.



Intuition, experience and good judgment

5.2. Approaches for efficient protein production

- I. Genetic approach x ***protein knowledgebase*** (biochemical approach)
- II. Expression density x ***functional activity***
- III. Expression system x ***medium engineering***
- IV. Troubles with removing tag fusion proteins x ***less convenient purification with classical chromatography***

IV. DNA cloning

4.2. The key questions before DNA cloning

4.2.1. DNA-protein analysis

4.2.1.2. Secondary structure prediction

www.expasy.ch jpred3

```
MASTDSESETRVKSVRTGRKPIGNPEDEQETSKPSDDEF LRGKRVLVDDNFISRKVATG
--EEEEEE----EEEEEEEEEE-----EEEEEE--HHHHHHHHH
KLKMGVSEVEQCDSGKEALRLVTEGLTQREEQGSVDKLPFDYIFMDCQMPMDGYEATR
HHHH----EEEE--HHHHHHHHHH-----EEEEEE----HHHHH
EIRKVEKSYGVRTPIIAVSGHDPGSEEARETIQAGMDAFLDKSLNQLANVIREIESKRH
HHHH-----EEEEEE----HHHHHHHHHH----E---HHHHHHHHHHHHH---
```

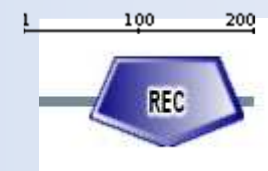
4.2.1.3. Domains detected by SMART

www.expasy.ch SMART

```
KRVLVDDNFISRKVATGKLKMGVSEVEQCDSGKEALRLVTEGLTQREEQGSVDKLP
FDYIFMDCQMPMDGYEATREIRKVEKSYGVRTPIIAVSGHDPGSEEARETIQAGMDA
FLDKSLNQLANVI
```

Confidently predicted domains, repeats, motifs and features:

Name	Begin	End	E-value
REC	43	171	1.19e ⁻²⁶



V. Protein expression

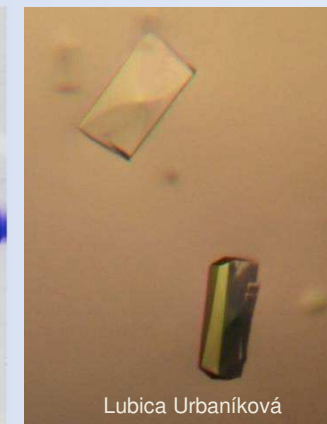
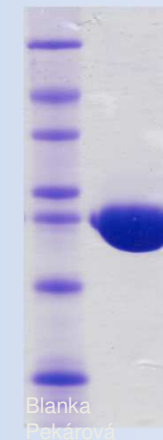
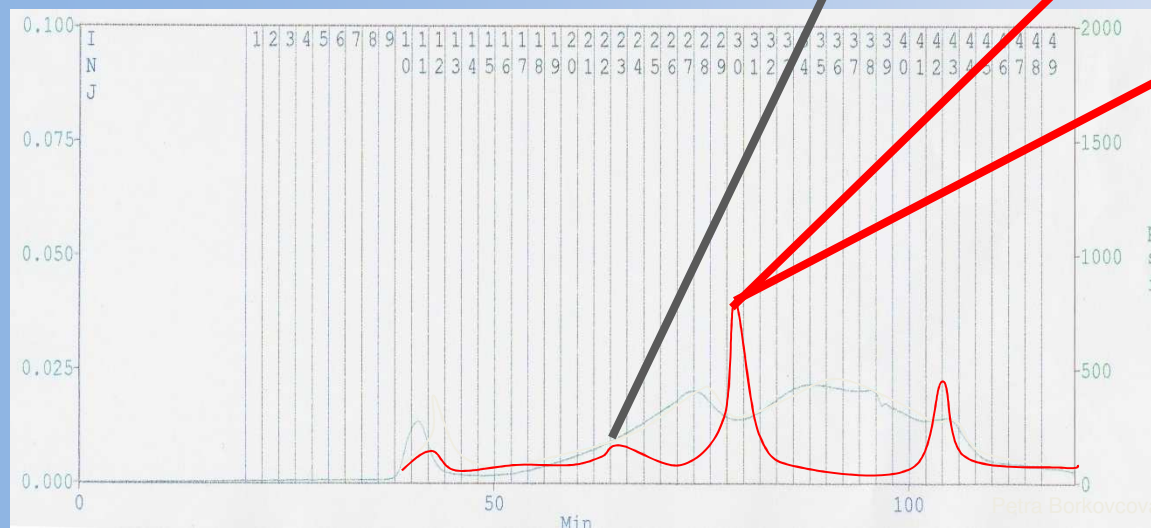
5.2. Approaches for efficient protein production

5.2.1. Genetic approach x *protein knowledgebase*

Histidine kinase from
A. thaliana - CKI 1



Three constructs of
receiver domain.
pET 28



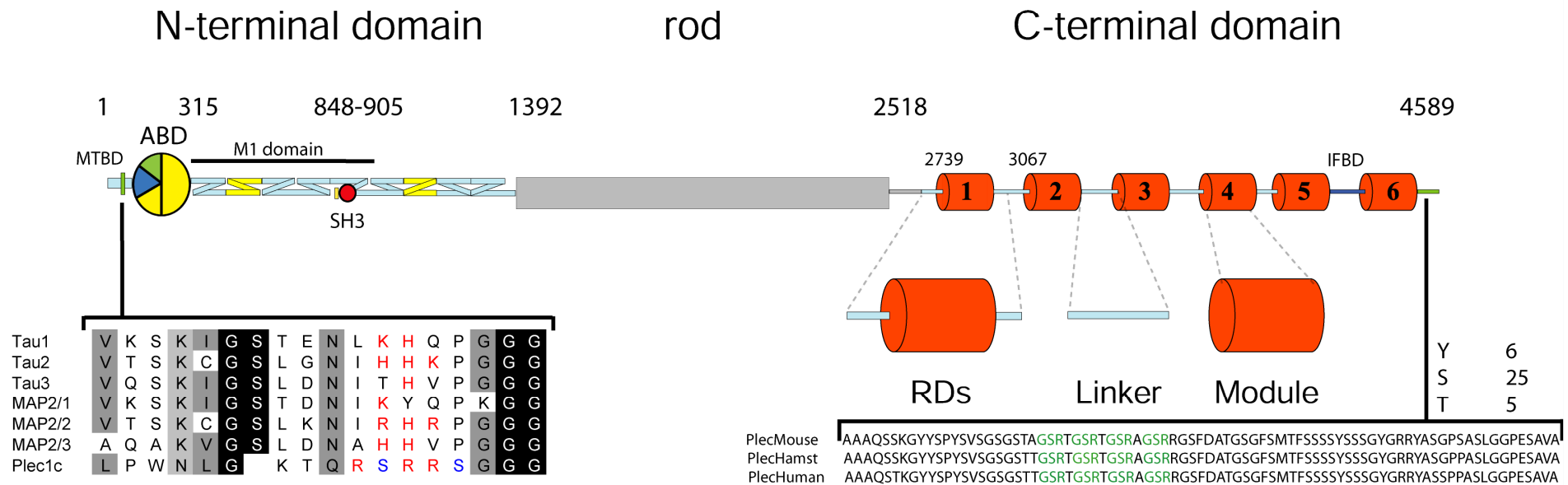
10

V. Protein expression

5.2. Approaches for efficient protein production

5.2.1. Genetic approach x *protein knowledgebase*

Plectin



Kd of Plectin (Ex 1-24) for actin 320 nM (Ex 2-8) 25uM
Kd of Plectin (R5) for vimentin (IF) 100 nM
Kd of Plectin (Ex 2-8) for integrin beta 4 170 nM

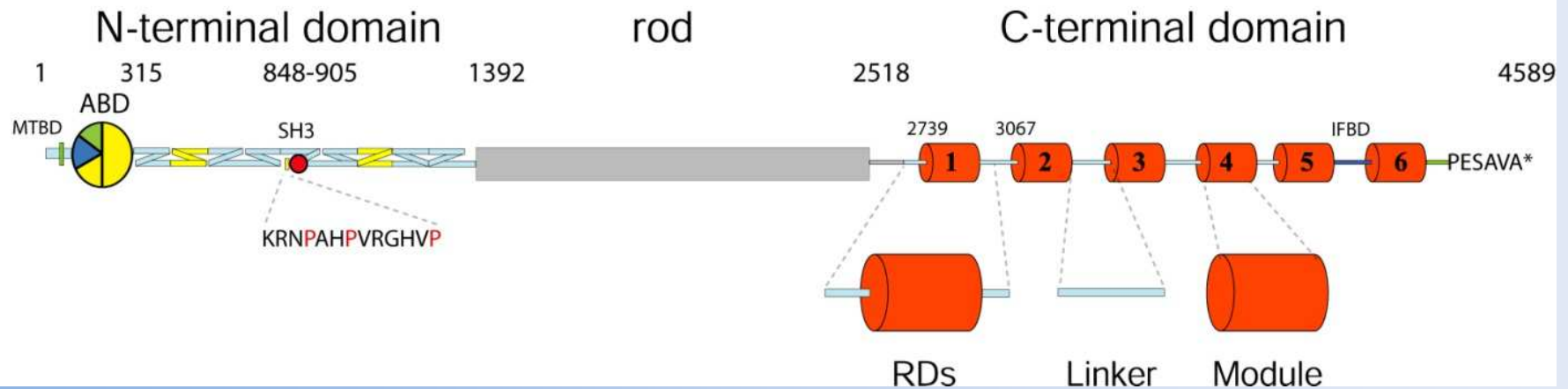
Kd for microtubules in case of MAP2 1-3 uM

V. Protein expression

5.2. Approaches for efficient protein production

5.2.1. Genetic approach x *protein knowledgebase*

Plectin



Spectrin	←	978
Actinin DM		834
Itk Tyr kinase		167
Envoplakin M		404
Actinin CE		847
Periplakin M		391
Kakapo DM		793
Kakapo AC		941
Desmoplakin M		536
MACF		858
Dystonin M		877
Plectin M		931

SH3 domain

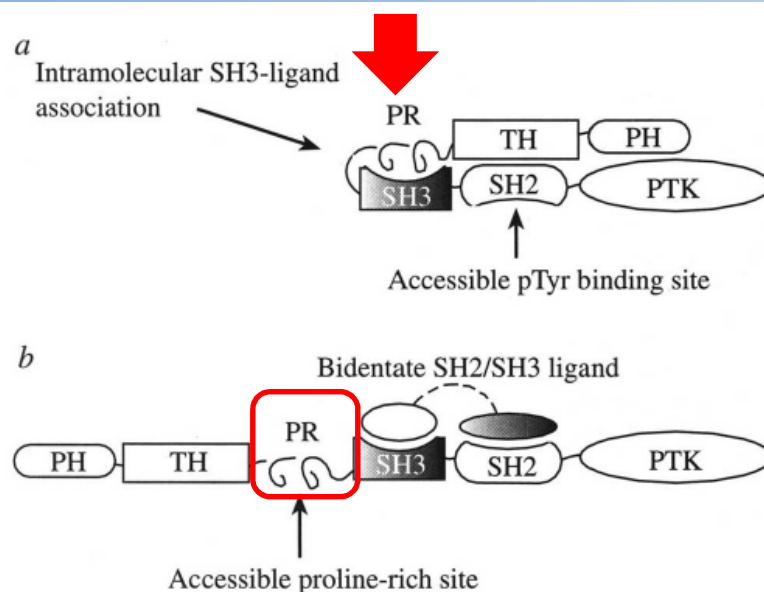
W	I	A	L	Y	D	F	E	A	R	S	R	R	E	V	S	M	K	K	N	D	V	L	T	L	L	S	S	I	N	K	D	W	R	K	V	E	A	D	D	H	Q	F	V	P	A	V	Y	V	R	K	L	A	1037
W	R	S	L	F	P	F	E	G	Q	G	M	R	M	D	K	G	E	V	M	L	L	K	S	K	T	N	D	D	W	W	C	V	R	K	D	N	G	V	E	G	F	V	P	A	N	Y	V	R	E	V	E	890	
W	I	A	L	Y	D	Y	Q	T	N	D	P	Q	E	L	A	L	R	C	D	E	E	Y	L	L	D	S	S	E	I	H	W	R	V	Q	D	K	N	G	H	E	G	Y	A	P	S	S	Y	L	V	E	K	S	240
W	D	S	I	C	D	W	D	S	G	E	V	Q	L	R	G	E	R	C	T	L	K	D	N	A	D	P	Y	T	W	L	V	Q	G	P	G	G	E	T	K	S	A	P	A	A	C	L	W	I	P	A	973		
W	T	A	L	C	D	Y	S	D	E	N	V	T	I	K	A	G	D	D	V	L	L	D	N	S	D	L	I	K	W	T	I	R	D	I	S	G	A	E	G	Q	V	P	S	V	V	F	R	I	914				
W	E	A	L	C	D	F	E	G	E	Q	G	L	I	S	R	G	Y	S	Y	T	L	Q	K	N	N	G	E	S	W	E	L	M	D	S	A	G	N	K	L	I	A	P	A	V	C	F	V	I	456				
W	Q	A	I	C	A	Y	K	Q	Q	Q	L	Q	I	E	K	G	E	T	V	T	L	L	D	N	S	C	R	V	K	W	R	V	R	T	A	K	G	Q	E	C	P	I	P	C	A	C	L	L	862				
W	Q	S	I	C	S	Y	K	Q	C	N	I	S	L	D	K	N	E	T	C	T	L	L	D	T	S	C	R	V	K	W	R	V	K	T	S	K	G	V	E	C	S	V	H	C	V	C	L	L	1008				
L	R	A	L	C	D	Y	K	Q	D	Q	K	I	V	H	K	G	D	E	C	I	L	K	D	N	N	E	R	S	K	W	Y	V	T	C	P	G	G	V	D	M	L	V	P	S	V	G	L	I	604				
W	K	A	I	C	D	Y	R	Q	I	E	I	T	I	C	K	N	D	E	C	V	L	E	D	N	S	Q	R	T	K	W	K	V	I	S	P	T	G	N	E	A	M	V	P	S	V	C	F	L	I	926			
I	K	A	I	C	D	Y	R	Q	I	E	I	T	I	Y	K	D	D	E	C	V	L	A	N	N	S	H	R	A	K	W	K	V	I	S	P	T	G	N	E	A	V	V	P	S	V	C	F	T	V	944			
L	I	A	V	C	D	Y	K	Q	V	E	V	T	V	H	K	G	D	Q	C	Q	L	V	G	P	A	Q	P	S	H	W	K	V	L	S	C	S	S	E	A	A	V	P	S	V	C	F	L	V	1000				

V. Protein expression

5.2. Approaches for efficient protein production

5.2.1. Genetic approach x *protein knowledgebase*

	PRD domain	SH3 domain	PRD																
Spectrin	978	AR	VIALYDFEARSRREVS	MKKNDVLTLLSSINKDW	WRVEADDHQG	FVPAVYVRKLA	PDE	L	1037										
Actinin DM	834	PH	VKSLFPFE	GQG	MKMDKGEVMLL	LKSKTND	WCVRKDNGVEGFVP	PAN	YVREVE	890									
Itk Tyr kinase	167	KKPLPPT	PEDNRRSFQRE	EETL	VIALYDYQTND	PQRLALRCD	EYYLLD	SS	SIHW	WRVQDKNGHEGYAP	SSYLVEKS	PNN	L	240					
Envoplakin M	404	PLPQRRN	PSKQPLH		VDSICDWD	SGEVQLLRGERCT	LKDNADPYT	WLVQGP	CGETKS	SAPAA	CLWIPA	PDP	E	973					
Actinin CE	847	PLWQRGER	IPH	PIK	VTALCDYS	DENVTIKAGDDVYL	LDNSDLIK	WTIRDISG	AEGQVP		SWVFRI	PPT	D	914					
Periplakin M	391	LKYRRET	PLK	PIP	VEALCDFE	GEQCLISRGYSYTL	LQKNN	GESWEL	MSAGNKLIAP		AVCFVI	PPT	D	456					
Kakapo DM	793	PLNKRRQ	PVNRQGP		VQAICAYK	QQGQLQIEKGETVTL	LDNSGRVK	WRVRTAKG	QEGPIP		GACLLL	PPP	D	862					
Kakapo AG	941	PLKQRKQ	PVNRQCT		VQSI	CSYK	QCNISLDKNETCTLL	LDTSGRVK	WRVKT	SKGVEG	SVH	GWCLLL	PPP	D	1008				
Desmoplakin M	536	LKPRNPDYRS	NKPII		LRALCDYK	QDQKI	VHKGDE	CEILKDN	NERSK	WYVTG	PCGVDMLVP	SWGLII	PPP	N	604				
MACF	858	LKPRNPDHVL	KSTLS		VKAICDYR	QI	EITICKN	DECVL	EDNS	QRTK	WVISP	TGNE	AMVP	SWCFLI	PPP	N	926		
Dystonin M	877	LKPRNPDNPL	KTSIF		IKAICDYR	QI	EITIKYK	DECVL	ANN	SHRAK	WVISP	TGNE	AVVP	SWCFTV	PPP	N	944		
Plectin M	931	LKPRNPAHP	VRGHV		LI	AVCDYK	QV	EVTW	HKGD	QCQLV	CPAQP	SHW	KWLS	GSS	SEAAVP	SWCFLV	PPP	N	1000



a) Model of observed intramolecular interaction showing the observed interaction between the Itk proline-rich region and SH3 domain.

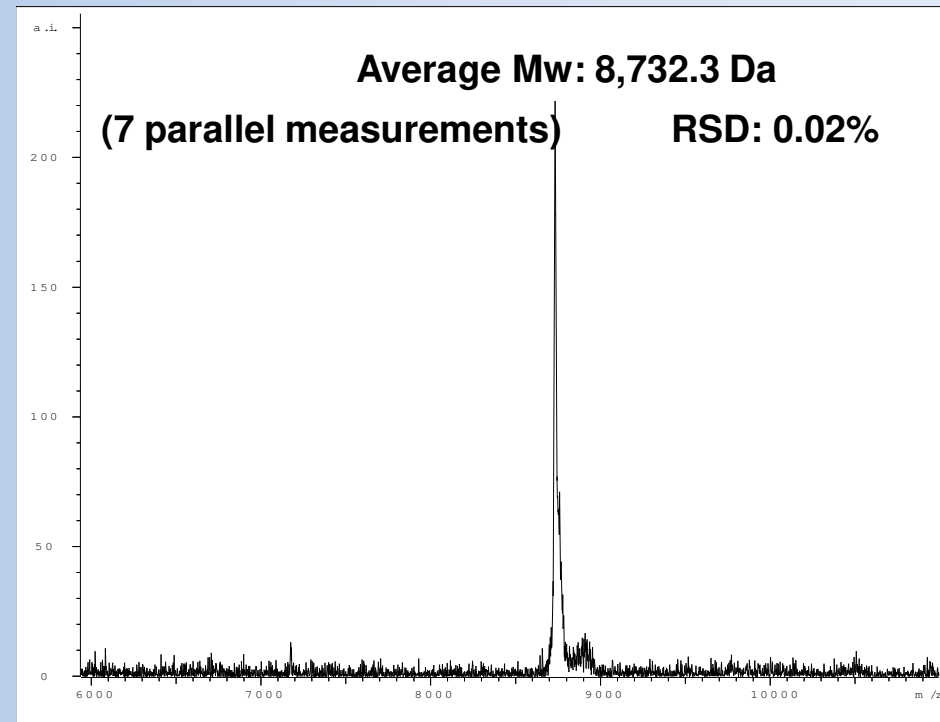
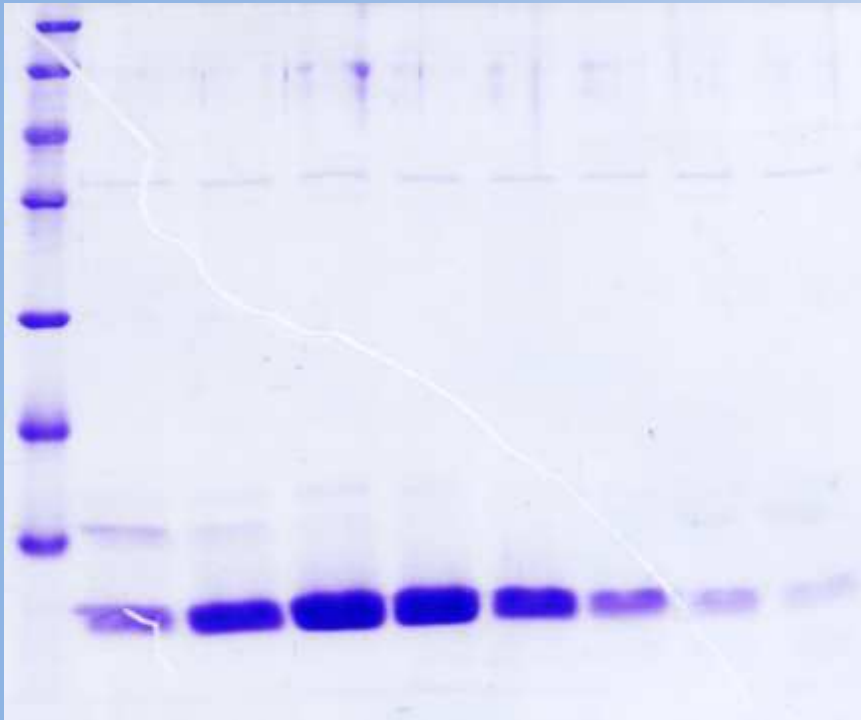
b) Model of the opening of the intramolecular complex by interaction with bidentate ligand for the Itk SH3 and SH2 domains. (Andreotti et al., Nature 1996)

V. Protein expression

5.2. Approaches for efficient protein production

5.2.1. Genetic approach x *protein knowledgebase*

SH3 domain of plectin with surrounding proline rich regions
(Sarc homology domain soluble in citrate buffer of pH 3.5)



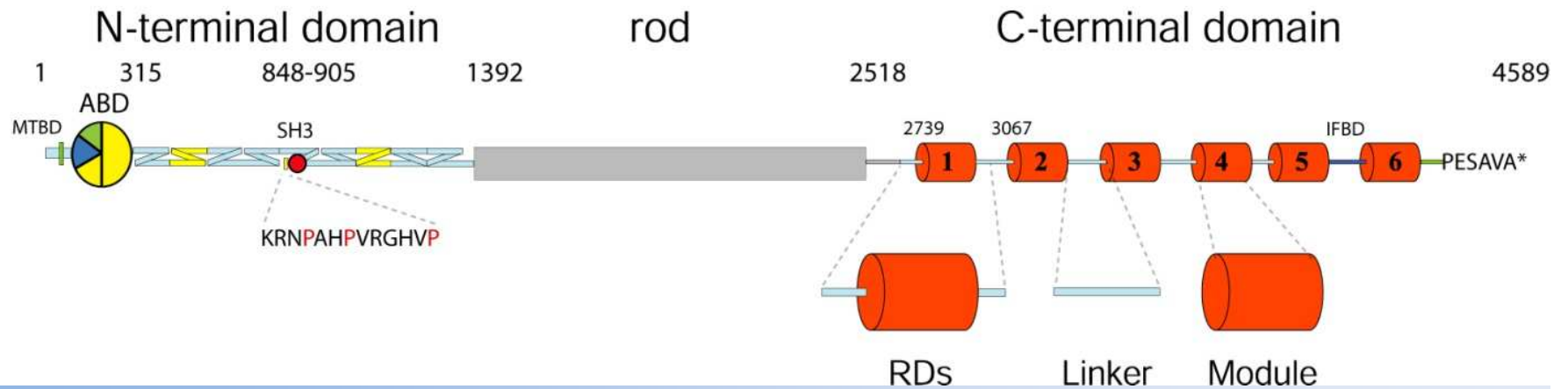
Theoretical pI/Mw (average) for the protein sequence

Theoretical pI/Mw: 7.78 / **8,726.11**

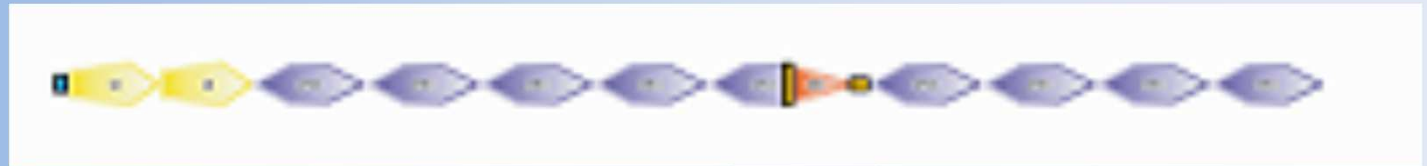
mefKAIIVQLKPRNPAHPVRGHHVPLIAVCDYKQVEVTVHKGD
QCQLVGPAQPSHWKVLSGSSSEAAVPSVCFLVPPPQNQEf

V. Protein expression

Plectin



Exon 1d-30



Exon 20-21



Exon 16-24



Exon 16-21

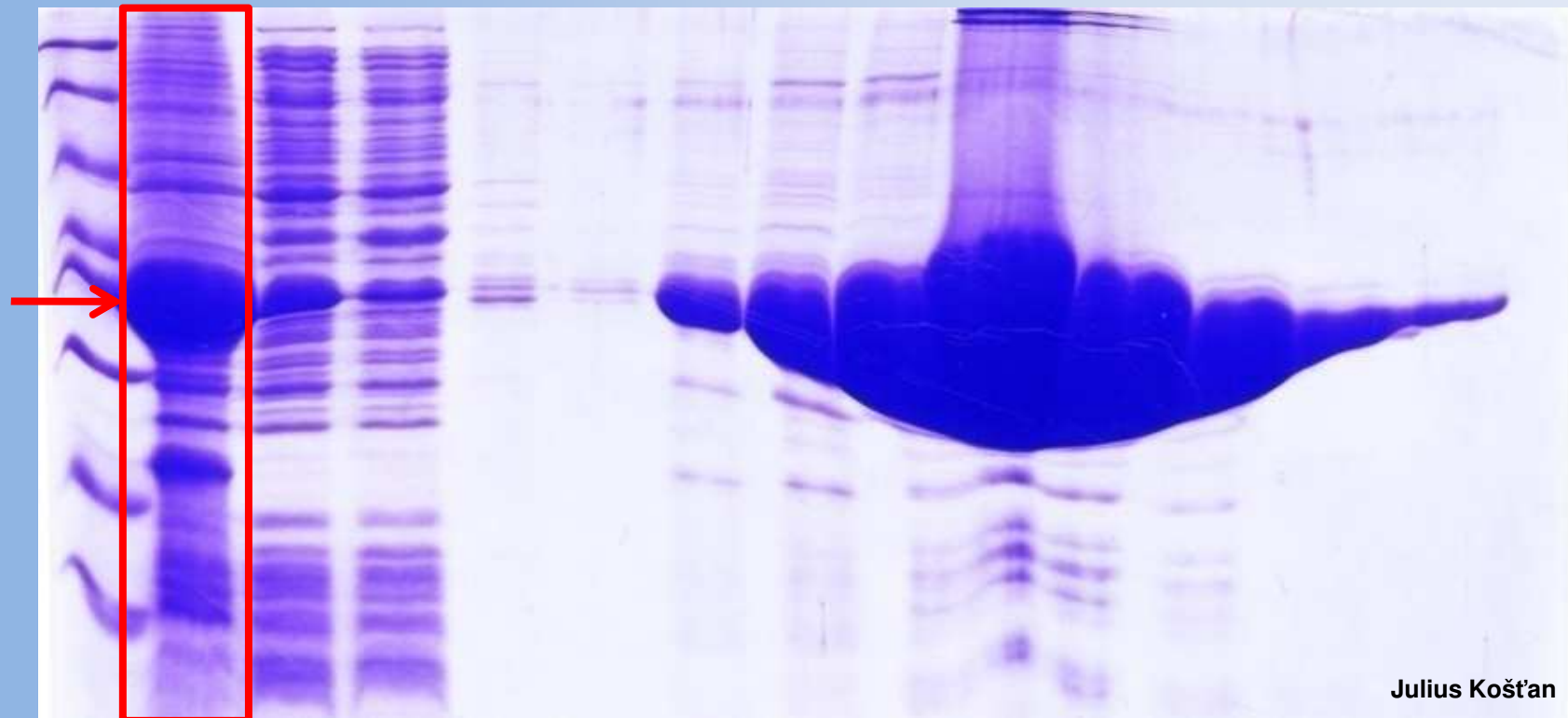


V. Protein expression

5.2. Approaches for efficient protein production

5.2.2. Expression density x *functional activity*

Expression and purification of plectin's ABD (Actin Binding Domain) in three isoforms.



V. Protein expression

5.2. Approaches for efficient protein production

5.2.2. Expression density x **functional activity**

Maize recombinant β -glucosidase produced in *E. coli*.

<i>Cultivation condition</i>	<i>Yield (mg)</i>	<i>Specific activity (nkat/mg) / (total activity nkat)</i>
LB medium	380	1.9 (966 nkat)
<i>TB medium - pH 6</i>	230	3.8 (874 nkat)
<i>TB medium - pH 7</i>	230	4.2! (966 nkat)
<i>TB medium - pH 8</i>	410	2.8 (1,148 nkat)
<i>Additive of cellobiose (LB medium)</i>	400	2.7! (1,080 nkat)

Radka Fohlerov

Result: TB medium (pH 7.0) supplemented by cellobiose shows 3.1 x higher β -glucosidase specific activity than in common LB medium.

V. Protein expression

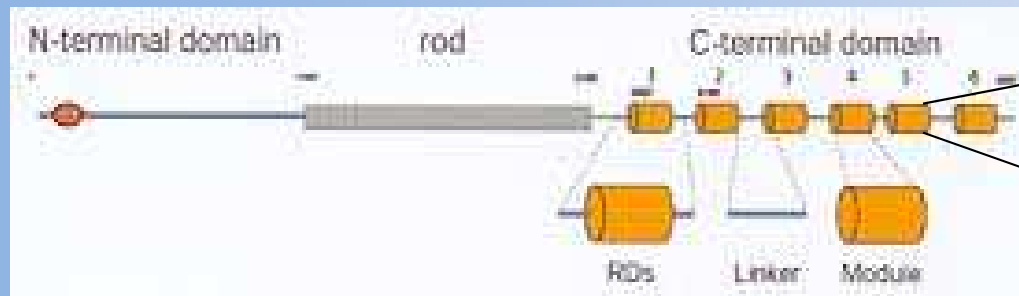
5.2. Approaches for efficient protein production

5.2.2. Expression density x *functional activity*

The cytolinker protein:
plectin



Plectin is one of the main linker proteins for the cytoskeleton.



Plectin R5

V. Protein expression

5.2. Approaches for efficient protein production

5.2.2. Expression density x **functional activity**

Converted pET 15b + IF binding domain of plectin

- R5 d. plectin (pH 7.9)
- R5 d. plectin (pH 7.9, urea, dialysis)
- R5 d. plectin (pH 7.9, urea, refolding HR)
- R5 d. plectin (pH 11, purification pH 9.0)

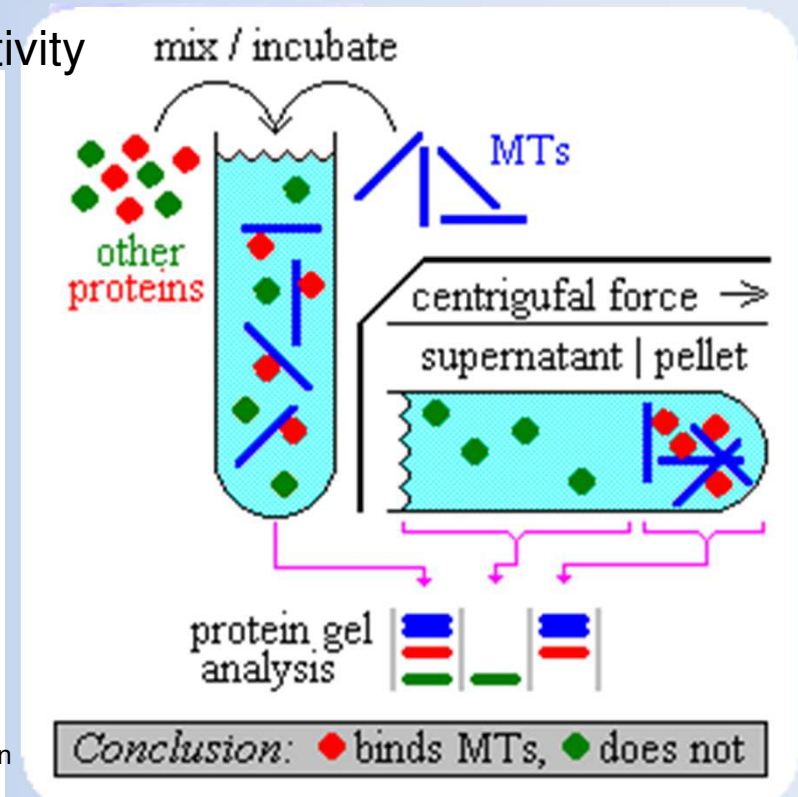
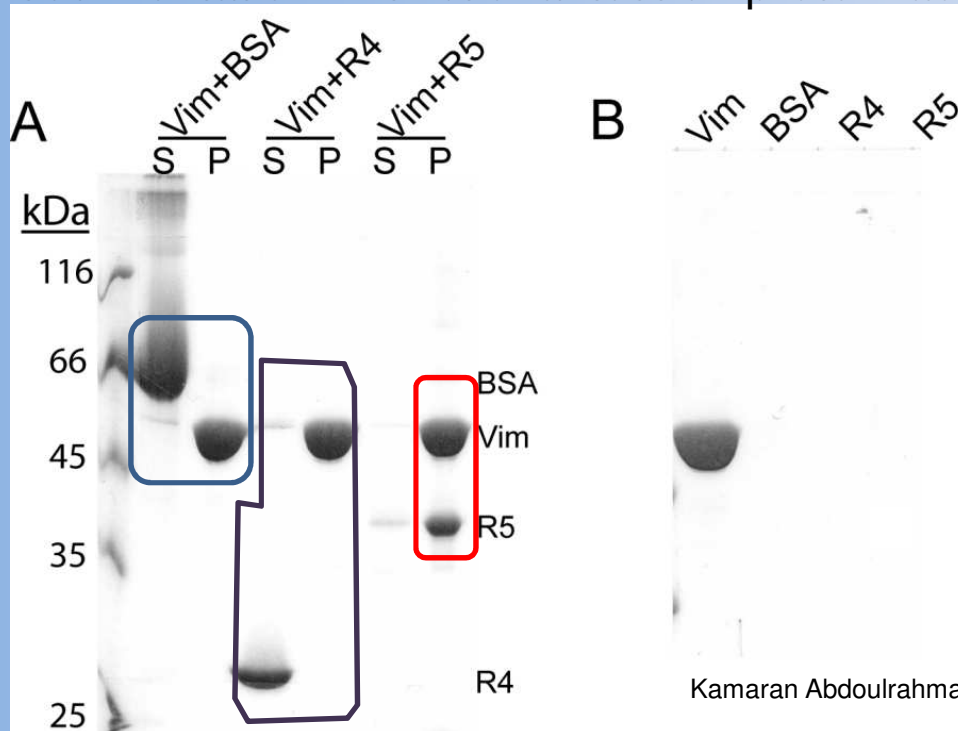
Insoluble form

Func. act.(45%)

Func. act.(60%)

Func. act. (≥95%)

Co-sedimentation – functional test on protein activity



V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

Examples of *E. coli* expression systems and web pages for further information

Vector system	Promoter/ induction method	Special host strains required	Protein tag	Source (website)
Pinpoint	<i>tac</i> /IPTG or T7 IPTG	Yes	Biotin binding domain	www.promega.com
* pET	T7 IPTG	Yes	His ₆ , T7 gene	http://www.merckbiosciences.co.uk
* pGEX	<i>tac</i> /IPTG	No	GST	www.amershambiosciences.com
pBAD	<i>araBAD</i>	Yes	His ₆ , GFP	www.invitrogen.com
pLEX	<i>P_L</i> /trp	Yes		
* pRSET	T7 IPTG	Yes	His ₆ , T7	
pPROTet	<i>P_{Ltet}</i> /anhydrotetracycline	No	His ₆	www.clontech.com
pTYB	T7 IPTG	Yes	Chitin binding domain	www.neb.com
* pMAL	<i>tac</i> /IPTG	Yes	Maltose binding domain	
pQE	T5/IPTG	Yes/TOPP	His ₆	www.qiagen.com
pCAL	T7/IPTG	Yes	Calmodulin binding peptide	www.stratagene.com
pFLAG	<i>tac</i> /IPTG	Yes		www.sigmaldrich.com

V. Protein expression

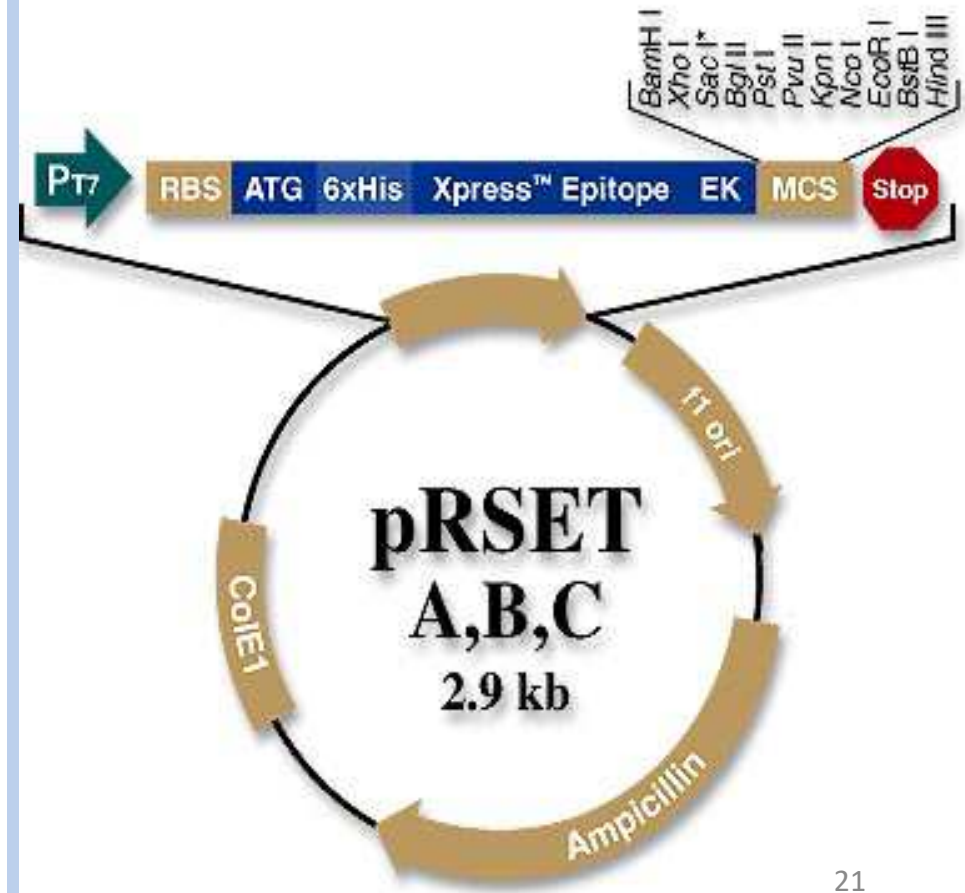
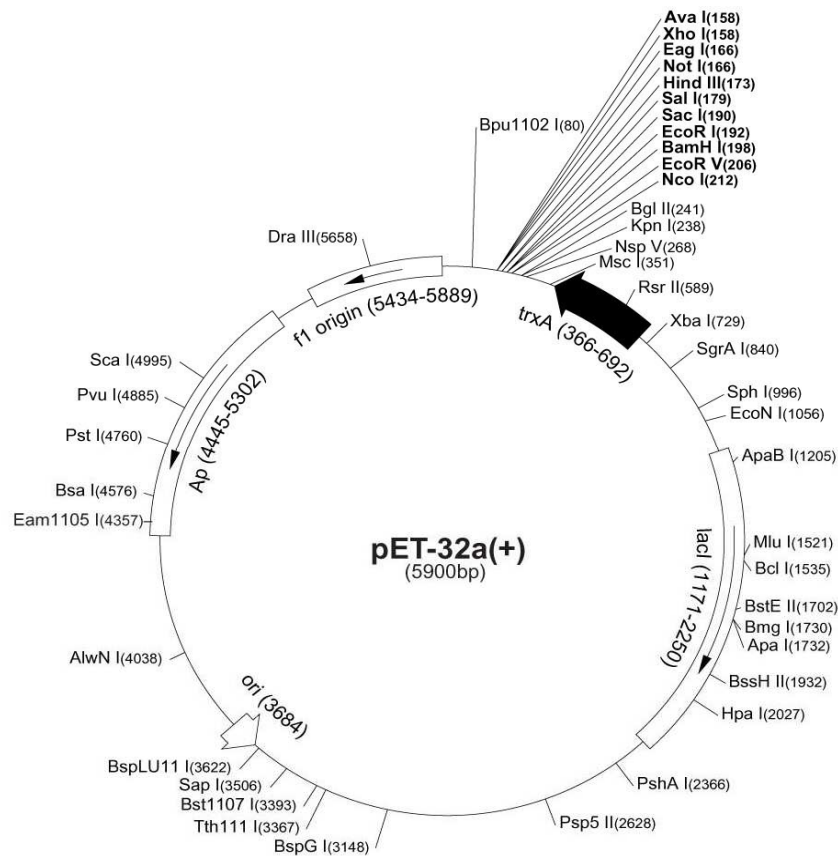
5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

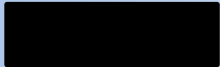
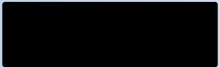
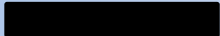
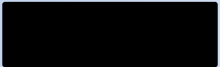

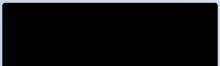
pET32a::AHP1
pET32a::AHP5

AHP – phosphotransfer protein in
cytokinin signalling pathway of *A. thaliana*

pRSETB::AHP1
pRSETB::AHP5



Bands of protein of interest on SDS-PAGE

Disintegration <i>E. coli</i> after induction	Normal buffer 	With urea 	100% soluble
Disintegration <i>E. coli</i> after induction	Normal buffer 	With urea 	50% soluble
Disintegration <i>E. coli</i> after induction	Normal buffer 	With urea 	insoluble

- **Disintegrate** *E. coli* in native buffer and divide into two same parts.
- **Denature** second part of the crude extract by chaotropic compounds (urea).
- Sediment both extracts and load on **SDS-PAGE**
- **Scan** the gel after staining and subsequent de-staining.
- Evaluate differences between signals from protein denaturated by chaotropic compounds and protein signal from native buffer.

V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

5.2.3.1. Temperature

Expression system

AHP1

<i>Plasmid</i>	pET32a+		pRSETB	
<i>Temperature (°C) growth/induction</i>	Soluble form (%)	Insoluble form (%)	Soluble form (%)	Insoluble form (%)
22°C/22°C	62%	38%	71%	29%
37°C/22°C	0%	100%	82%	18%
37°C/28°C	0%	100%	8%	92%

AHP5

<i>Plasmid</i>	pET32a+		pRSETB	
<i>Temperature (°C) growth/induction</i>	Soluble form (%)	Insoluble form (%)	Soluble form (%)	Insoluble form (%)
22°C/22°C	78%	22%	76%	24%
37°C/22°C	67%	33%	81%	19%
37°C/28°C	61%	39%	81%	19%

V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

5.2.3.1. Temperature

Production of soluble AHP proteins using *E. coli* expression vector pRSET at different temperatures (%)

Temperature (°C) growth/induction	AHP1	AHP2	AHP3	AHP4	AHP5	AHP6
37°C/28°C	8%	85%	100%	0%	76%	0%
37°C/22°C	82%	73%	100%	0%	81%	51%
22°C/22°C	71%	78%	100%	30%	81%	73%

V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

5.2.3.2. Medium pH

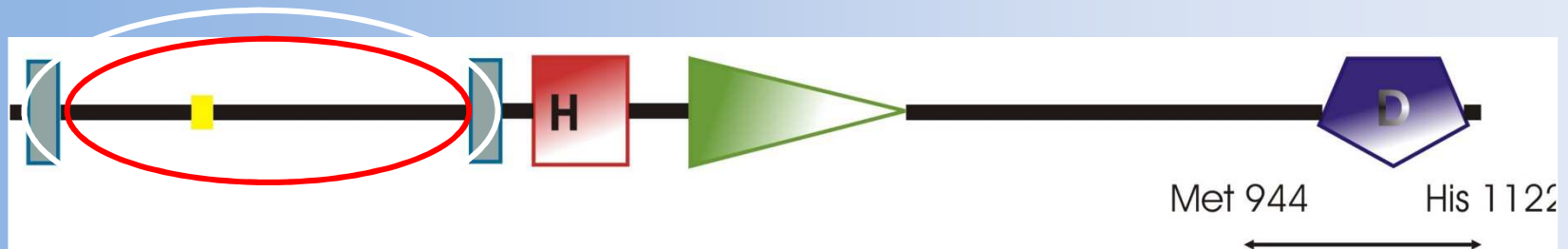
Production receiver domain of plant histidine kinase AHK4 in <i>E. coli</i> by <i>pET161DEST</i>			
pH	6.0	7.0	8.0
Soluble fraction	35%	89%	100%

- **Disintegrate** *E. coli* in native buffer and divide into two same parts.
- **Denature** second part of the crude extract by chaotropic compounds (urea).
- Sediment both extracts and load on **SDS-PAGE**
- **Scan** the gel after staining and subsequent de-staining.
- Evaluate differences between signals from protein denaturated by chaotropic compounds and protein signal from native buffer.

V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*



pDEST17::CKI1ex1 – 371 AA, Mw = 42 kDa

pDEST17::CKI1ex2 – 419 AA, Mw = 47 kDa

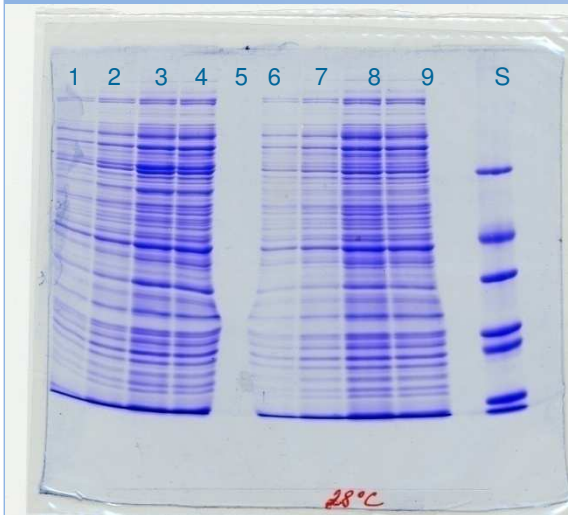
V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

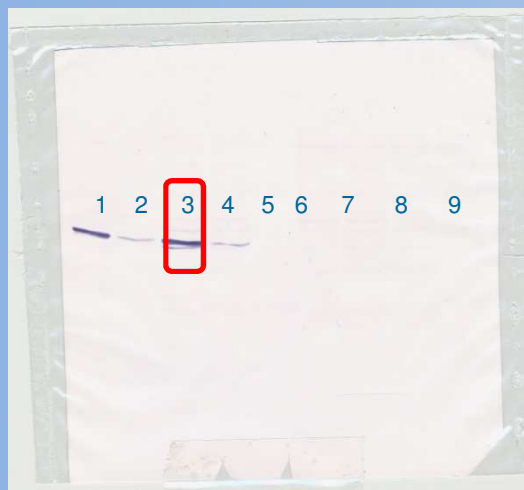
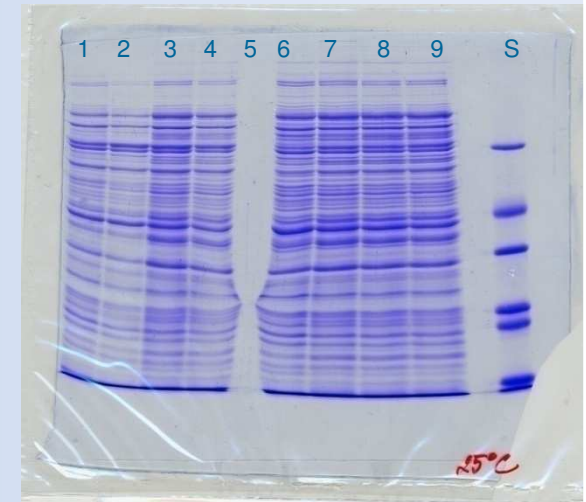
Growth temperature
37°C, expression 28°C

Growth and expression 25°C



- 1: BL21 — 1 h after induction = OD 0.5
- 2: 4 h after induction = OD 0.5
- 3: 1 h after induction = OD 2
- 4: 4 h after induction = OD 2
- 5: ←
- 6: C43 — 1 h after induction = OD 0.5
- 7: 4 h after induction = OD 0.5
- 8: 1 h after induction = OD 2
- 9: 4 h after induction = OD 2

S: 14–66 kDa



- 1: BL 21 before induction
- 2: 1 h after induction = OD 0.5
- 3: 3 h after induction = OD 0.5
- 4: 2 h after induction = OD 2
- 5: -
- 6: C43 before induction
- 7: 1 h after induction = OD 0.5
- 8: 3 h after induction = OD 0.5
- 9: 2 h after induction = OD 2

S: 14–66 kDa



Petra Borkovcová

V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

Expression *E. coli* BL21 (DE3)_{RIL}

TB medium: pH 6.0, pH 7.0, [redacted]

growth [redacted] → induction [redacted]

growth 25°C → induction 22°C, 13.5 h

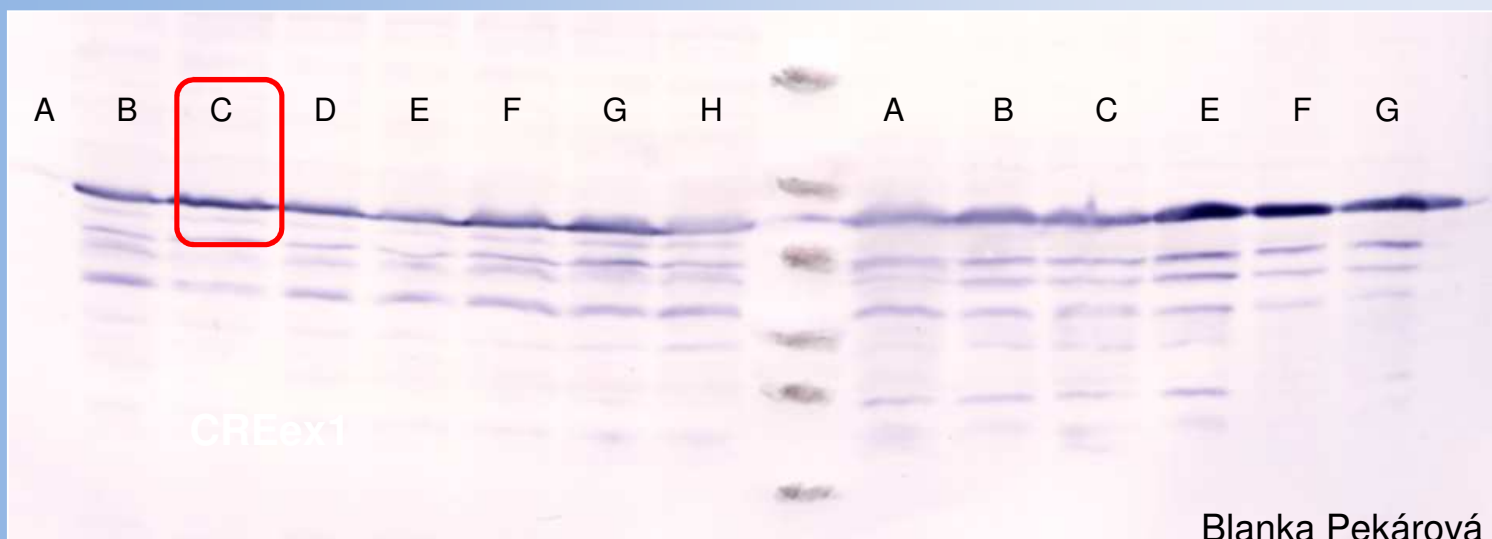
Lysis buffers:

- A. Citrate b., pH 3.6, Triton X-100
- B. Tris b., pH 7.9, Triton X-100
- C. [redacted]
- D. Tris b., pH 7.9, Triton X-100
- E. Tris b., pH 7.9, CTAB
- F. Tris b., pH 7.9, NONIDET P-40
- G. Tris b., pH 7.9, SDS
- H. Tris b., pH 7.9

CKI1ex1

supernatant

pellet



V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

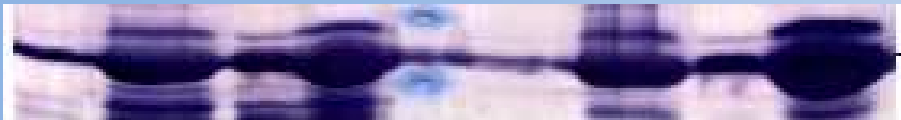
E. coli BL21(DE3) Arctica

Buffer engineering

1) pH

2) detergent

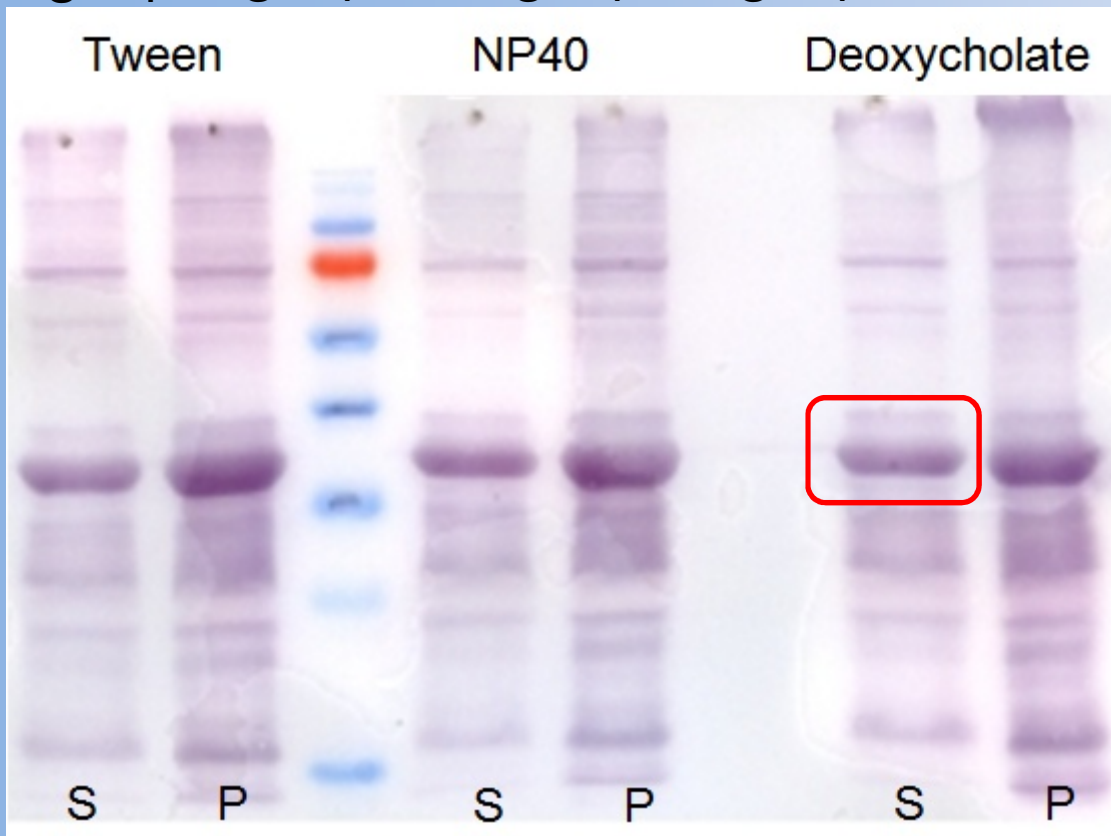
Tris Glycine Tris ↓ Glycine pH10,6



S P S P S P S P

Western blot

detected by poly- His antibodies



E. coli BL21(DE3)
Arctica RP



V. Protein expression

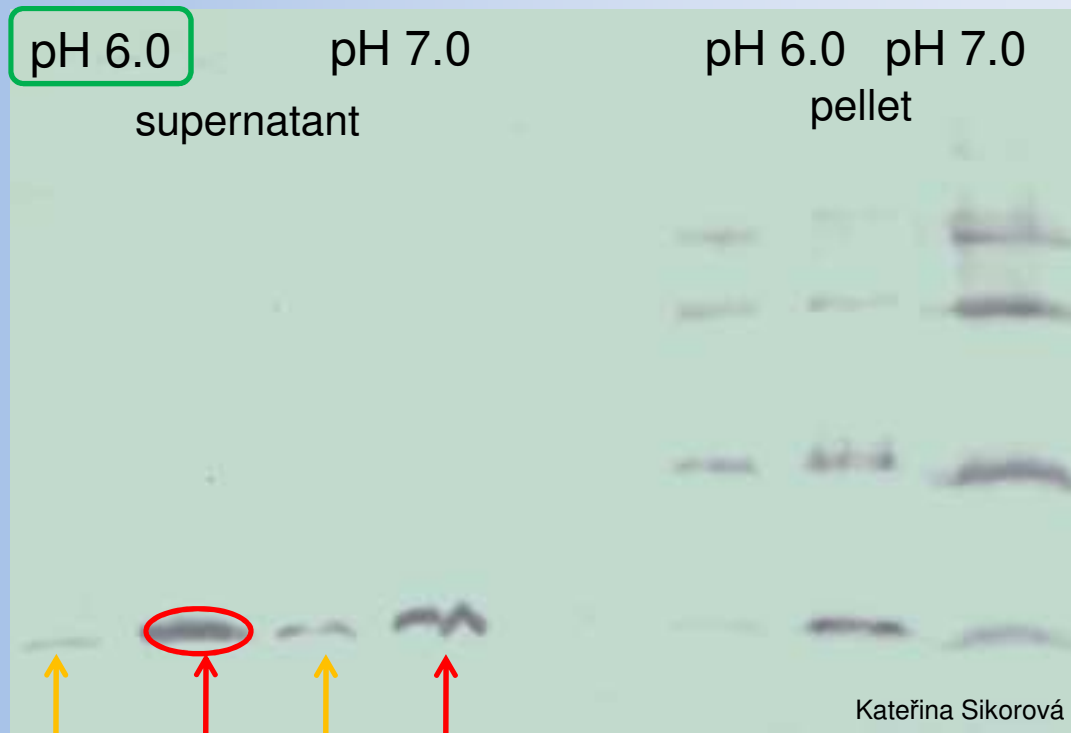
5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

5.2.3.3. Buffer for desintegration (pH)

LTP-2 (non-specific lipid transporting protein from wheat)

pH medium



Buffer for disintegration (pH)

Glycine buffer pH 10.6

Phosphate buffer pH 7.2

Screening for solubility

- Optimum Solubility Screen
 - Optimizing protein solubility and stability with salts and buffers prior to crystallization can improve crystallization results
 - Microbatch experiments can identify salt and buffer conditions that increase a protein's solubility

Acta Cryst. (2006) D**62**, 833 - 842. Izaac, C. A. Schall and T. C. Mueser. Assessment of a preliminary solubility screen to improve crystallization trials: uncoupling crystal condition searches.

Acta Cryst. (2004) D**60**, 1674 -1678. B. K. Collins, S. J. Tomanicek, N. Lyamicheva, M. W. Kaiser and T. C. Mueser. A preliminary solubility screen used to improve crystallization trials: crystallization and preliminary X-ray structure determination of *Aeropyrum pernix* flap endonuclease-1

Acta Cryst. (2004) D**60**, 1670 -1673. J. Jancarik, R. Pufan, C. Hong, S.-H. Kim and R. Kim. Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins.

Incomplete factorial: 950 cocktails

35 Salts and 8 Buffers

Salt	Occurrences (# cocktails)		
NH ₄ Br	28	KCl	31
NH ₄ Cl	29	KNO ₃	32
NH ₄ NO ₃	27	KH ₂ PO ₄	30
NH ₄ H ₂ PO ₄	32	KSCN	29
(NH ₄) ₂ HPO ₄	26	RbCl	28
(NH ₄) ₂ SO ₄	30	NaBr	28
Ca(C ₂ H ₃ O ₂) ₂	12	NaCl	28
CaCl ₂	13	Na ₂ MoO ₄	28
LiBr	32	NaNO ₃	30
LiCl	30	NaH ₂ PO ₄	23
Mg(C ₂ H ₃ O ₂) ₂	17	Na ₂ S ₂ O ₃	23
MgCl ₂	26	Zn(C ₂ H ₃ O ₂) ₂	22
MgSO ₄	27	K ₂ HPO ₄	25
MnCl ₂	22	CoSO ₄	24
KC ₂ H ₃ O ₂	30	Li ₂ SO ₄	27
KBr	33	K ₃ PO ₄	28
K ₂ CO ₃	34	NH ₄ SCN	29
		MnSO ₄	14
		Mg(NO ₃) ₂	23
		TOTAL	950



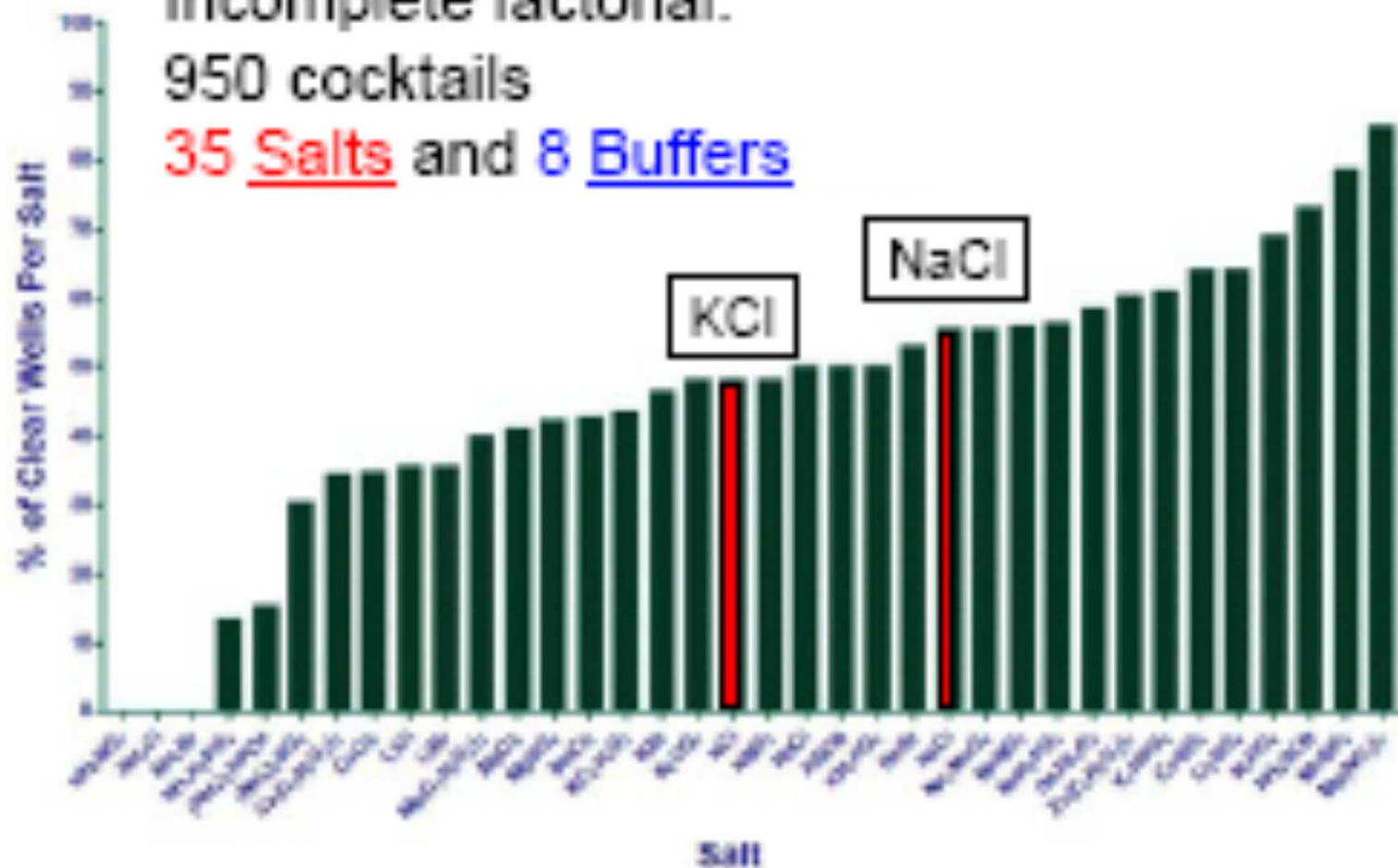
Buffer	pH	Occurrences (# cocktails)
Citrate	4.2	116
Acetate	5	144
MES	6	124
MOPS	7	124
HEPES	7.5	104
Tris	8	134
TAPS	9	109
CAPS	10	95
TOTAL		950

Data to calculate this for every protein
Statistically significant data

Incomplete factorial:

950 cocktails

35 Salts and 8 Buffers



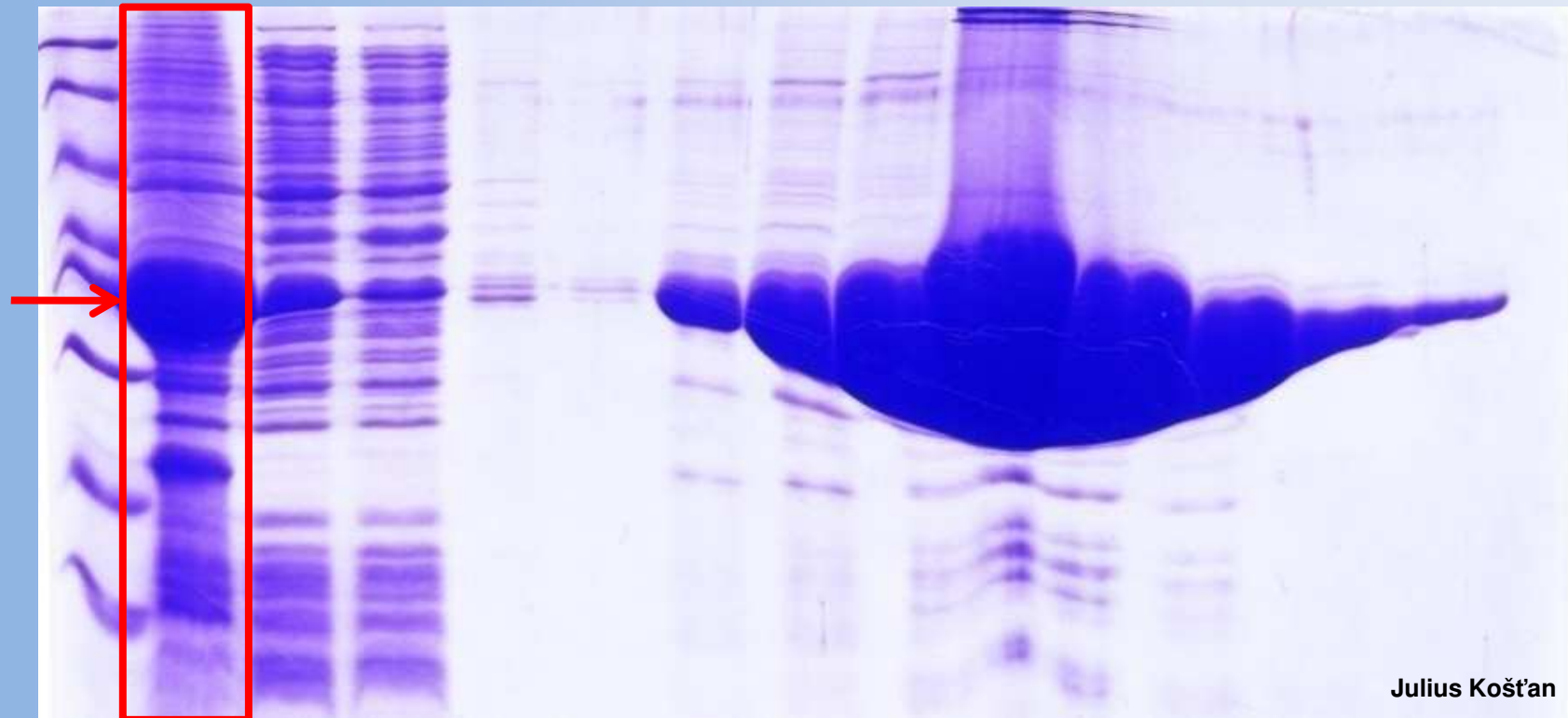
Increasing solubility →

V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

Expression and purification of plectin's ABD (actin binding domain) in three isoforms.



V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

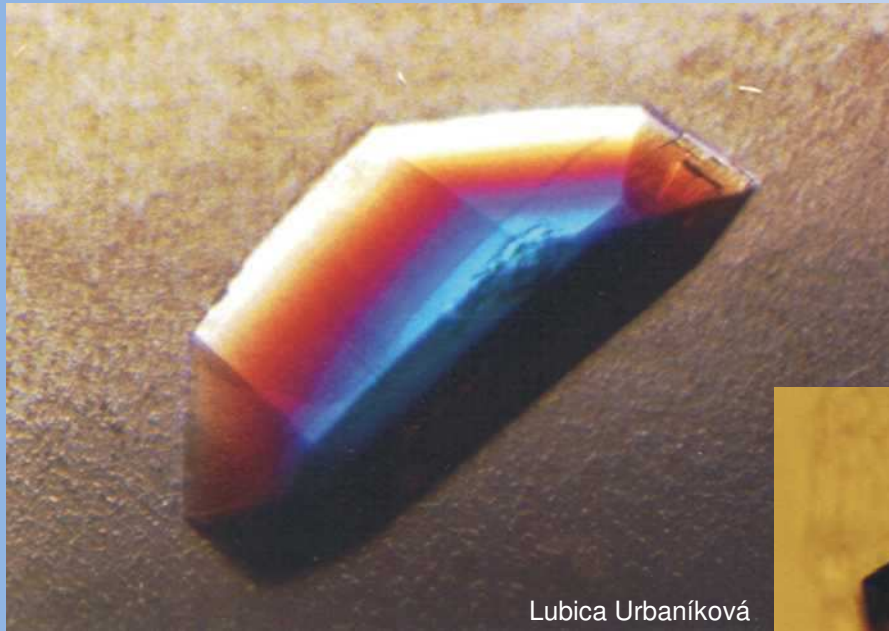
Monoclinic crystals of plectin ABD

Precipitant solution:

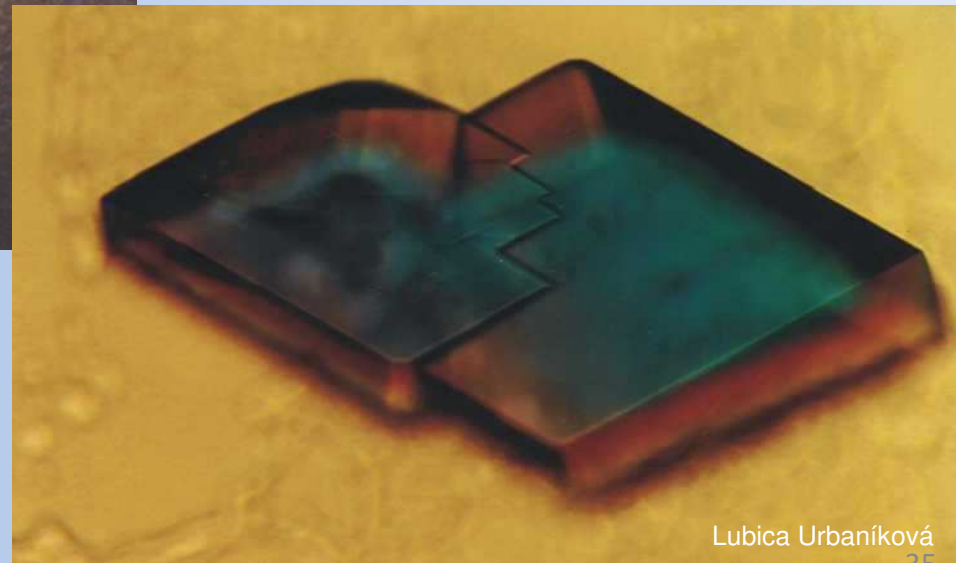
0.1 M TRIS buffer pH 8.5

10% PEG 4000

2% dioxane



Space group $P2_1$
2 molecules in asym. unit



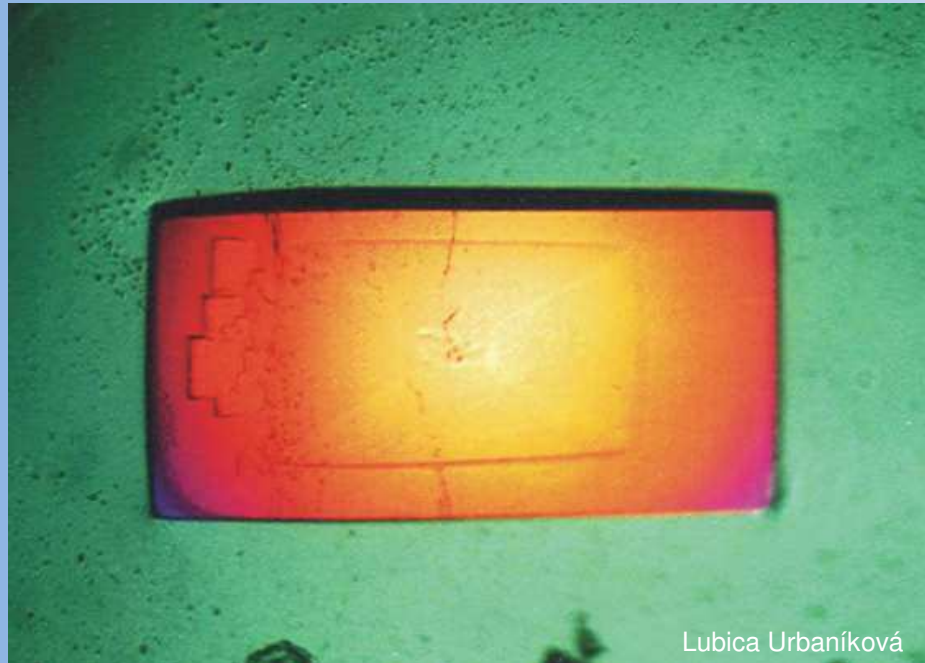
V. Protein expression

5.2. Approaches for efficient protein production

5.2.3. Expression system x *medium engineering*

Orthorhombic crystals of plectin ABD

Precipitant solution:



0.1 M Cacodylate buffer pH 6.5
6–8% PEG 8000
0.2 M Ca acetate
2% dioxane

Space group $P2_12_12_1$
1 molecule in asym. unit
2.0 Å resolution

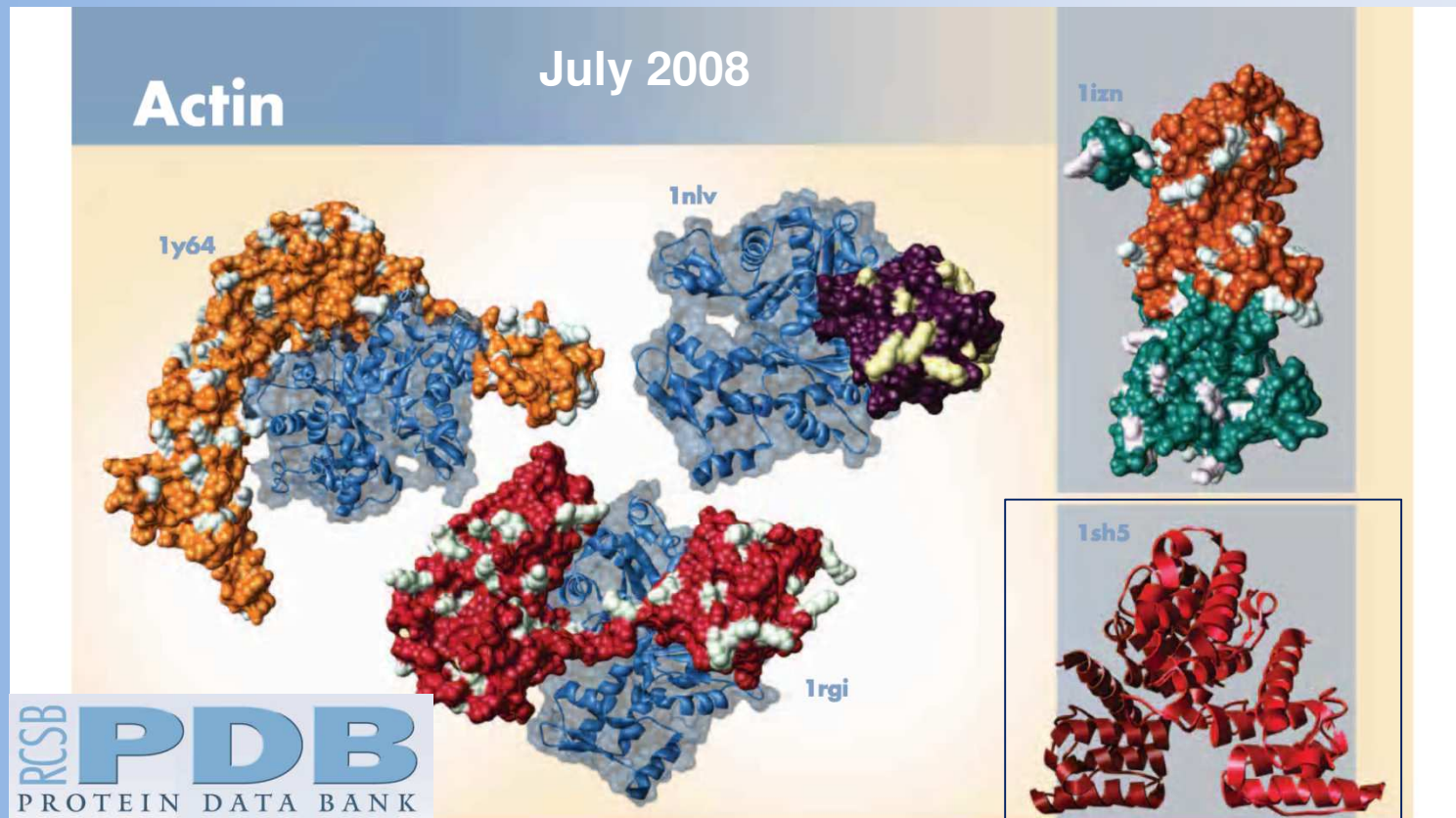
V. Protein expression

1nlv: S.M.Vorobiev, B. Strokopytov, D.G. Drubin, C. Frieden, S. Ono, J. Condeelis, P.A. Rubenstein, S.C. Almo. The structure of non-vertebrate actin: Implications for the ATP hydrolytic mechanism (2003). *Proc.Natl.Acad.Sci. USA* **100:5760-5765.**

1rgi: L.D.Burnick, D. Urosev, E. Irobi, K. Narayan, R.C. Robinson (2004). Structure of the N-terminal half of gelsolin bound to actin: roles in severing, apoptosis and FAF. *EMBO J.* **23:2713-2722.**

1izn: A.Yamashita, K. Maeda, Y. Maeda (2003). Crystal structure of CapZ: structural basis for actin filament barbed end capping. *EMBO J.* **22:1529-1538.**

1sh5: J. Sevcik, L. Urbanikova, J. Kostan, L. Janda, G. Wiche (2004). Actin-binding domain of mouse plectin: crystal structure and binding to vimentin. *Eur.J.Biochem.* **271:873-1884.**



The cytoskeleton is an intracellular maze of filaments that supports and shapes the cell. The most plentiful type of filament is composed of actin, shown here in blue. The cytoskeleton, however, is not a static structure, since it must respond to the changing needs of the cell.

The proteins shown here help to reshape the cytoskeleton by assembling or disassembling actin filaments as necessary. A molecule of ATP, which is

bound inside each actin molecule, is important in this process. When it is hydrolyzed to ADP, the filament becomes unstable and falls apart.

Gelsolin breaks down actin filaments by assisting the hydrolysis of ATP and blocking the sites of interaction with other actin proteins. Two different fragments of gelsolin are shown in 1nlv and 1rgi bound to actin.

The protein CapZ forms a cap on the actin filaments shown in 1izn, which

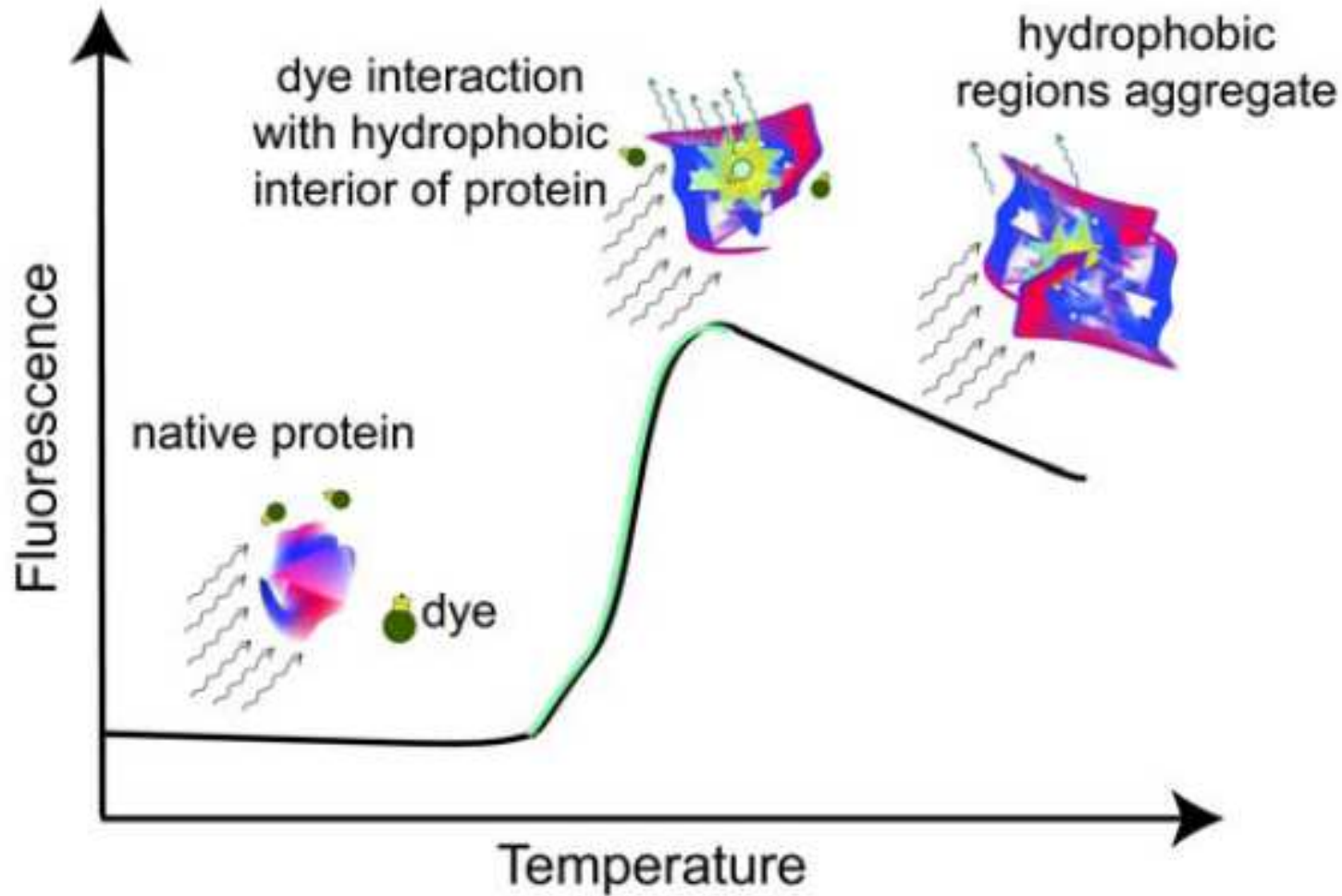
limits assembly.

The protein formin assists the assembly of actin by aligning two actin proteins in the proper orientation which starts the process of filament growth. One domain of formin is shown bound to actin in 1y64.

Plectin links neighboring actin filaments into higher order structures. The actin-binding domain is shown in 1sh5.

ThermoFluor[®] assay

CD - spectroscopy



ThermoFluor[®] assay

- No data in the literature to support the prediction of crystallization conditions from T_m values
- Literature reports ThermoFluor[®] can identify ligands that stabilize macromolecules to improve crystallization outcomes

Ericsson U.B., *et al.* Thermofluor-based high-throughput stability optimization of proteins for structural studies. (2006) *Analytical Biochemistry* **357**(2) 289-298.

Vedadi M, *et al.* Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. (2006) *PNAS*, **103**(43):15835-40.

V. Protein expression

5.2. Approaches for efficient protein production

5.2.5. Maximizing target protein recovery

- **Protein expression**
 - The knowledge applied in constructing expression vectors;
 - Changes in cultivation conditions:
 - protein induction at different temperature
 - media pH
 - various additives;
- **Protein purification**
 - choice of buffer for disintegrating cells (all buffers should be degassed,
 - The availability of a test for protein activity as a supplement to the solubility test;
 - The use of expression vectors without affinity fusion proteins, which one should not be afraid to do.

V. Protein expression

5.3. Expression system

5.3.1. *E. coli* expression system

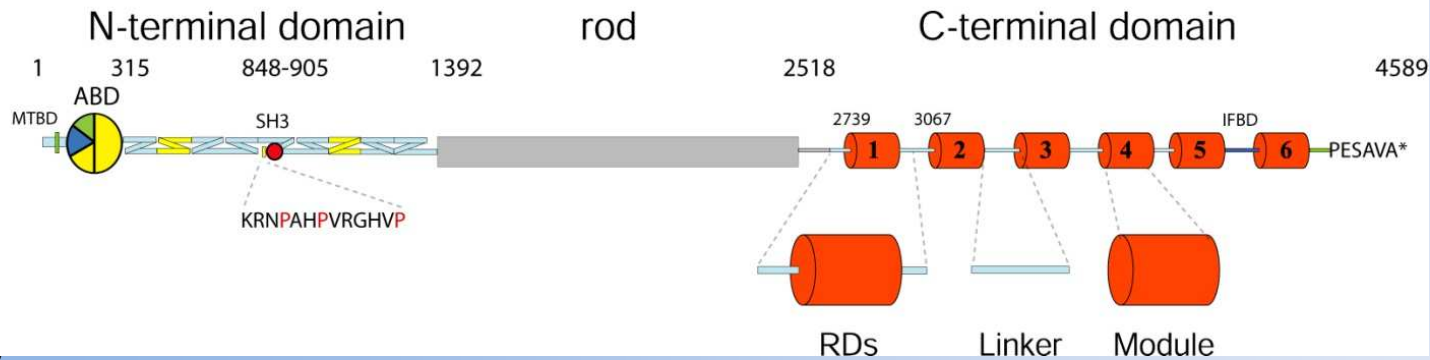
Advantages and Disadvantages of *E. coli*

- Ease of gene manipulation
 - Availability of reagents
 - Easy of producing quantities of protein
 - Speed
 - Low cost
 - Adaptability of the system
-
- Formation of insoluble inclusion bodies
 - Size of the protein
 - Post-translational modification

V. Protein expression

- Size of the protein
- Formation of insoluble inclusion bodies

Plectin



1c-33
MEFHMSGEDS ... TLRRMEEEEF
:
pI/Mw: 6.37 / **160,602.22**

MW 160,602.22 Da

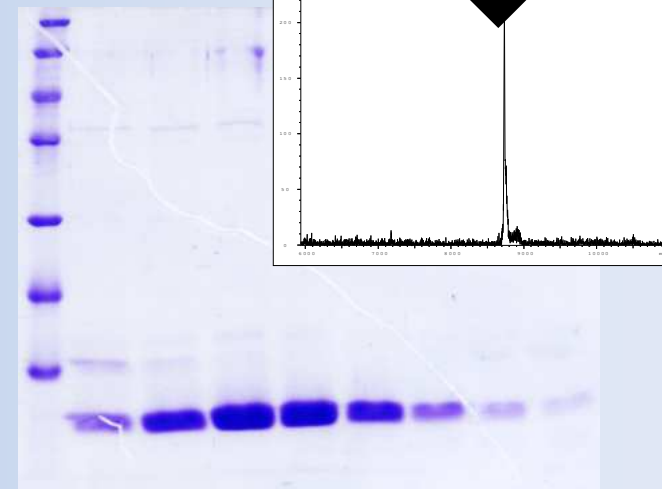
1c-24
MEFHMSGEDS ...
CISELKDIEF :
pI/Mw: 6.38 / **119,799.86**

MW 119,799.86 Da



N-terminal plakin domain of mouse

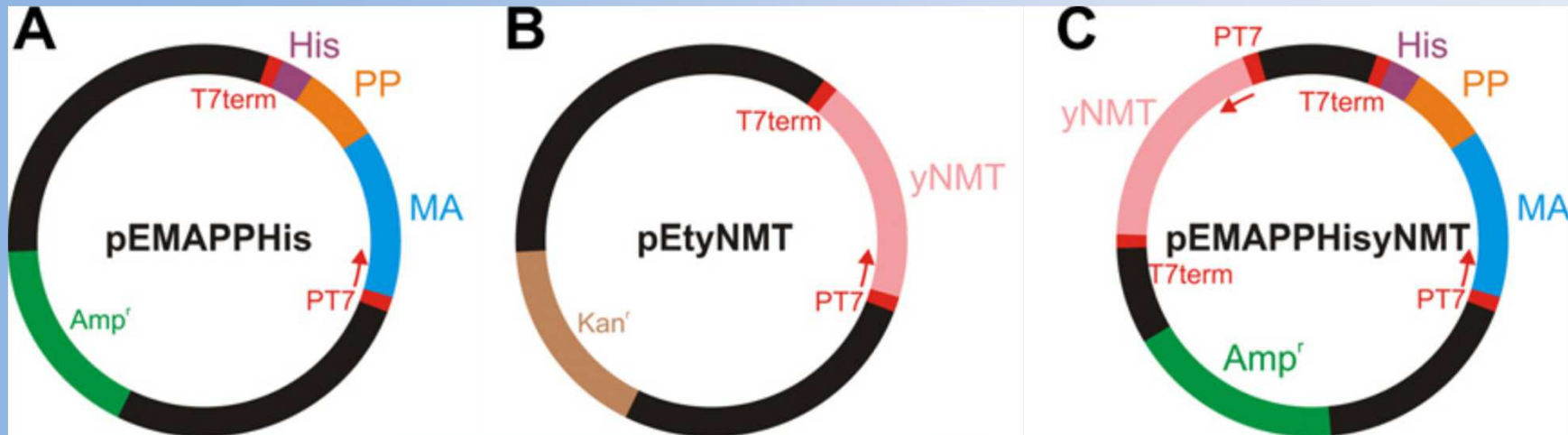
Mw: 8,732.3 Da



SH3 domain of mouse plectin

(A) The sequence encoding the matrix protein and downstream 18 amino acids of phosphoprotein (PP) was introduced to the pET22b vector (Novagen). The resulting vector pEMAPPHis was verified by restriction cleavage and by sequencing.

(B) The vector pETyNMT for the production of yeast N-myristoyltransferase was constructed by inserting the gene encoding the enzyme to the pET29b vector



(C) The plasmid carrying both MAPPHis and NMT genes constructed on the basis of the pET19b and pET11b vectors used in the single-plasmid system.

(Prchal et al., 2011)

V. Protein expression

5.3. Expression system

5.3.2. Baculovirus protein expression system

- HT-bacmid propagation
- HT-suspension-based insect cell transfection
- Methods of recombinant viral titer determination
GFP co-expression, titration assay using
Alamarblue, Cedex cell counter
- HT-miniaturized deep-well block insect cell
expression
- Transient insect cell expression

V. Protein expression

5.3. Expression system

5.3.2. Baculovirus protein expression system

Overview on commercially available baculovirus expression systems

Baculovirus expression kits and vendors	Compatible transfer vectors	Methodology for cloning foreign gene into transfer vector	Transfer of foreign gene into Baculovirus genome	Selection/ Recombination efficiency
BacPAK™ (Clontech)	Based on homologous recombination at polyhedrin locus	Ligase dependent	Homologous recombination in insect cells	≥90%
Bac-to-Bac™ (Invitrogen)	Based on site-specific transposition	Ligase-dependent, Gateway™ adapted	Site-specific transposition in bacterial cells	Selection of recombinants by blue-white selection on agar plates
BaculoDirect™ (Invitrogen)	Based on site-specific recombination	Gateway™ adapted	Site-specific recombination in Eppendorf tube	Antibiotic selection of transfectants in insect cells
flashBac™ (OET/NextGen Sciences)	Based on homologous recombination at polyhedrin locus	Ligase dependent	Homologous recombination in insect cells	100%
BacVector™ 1000, 2000, 3000 (EMD/Novagen)	Based on homologous recombination at polyhedrin locus	Ligase dependent	Homologous recombination in insect cells	≥ 95%
BaculoGold™ (Cloentech)	Based on homologous recombination at polyhedrin locus	Ligase dependent	Homologous recombination in insect cells	≥ 95%
DiamondBac™ (Sigma-Aldrich)	Based on homologous recombination at polyhedrin locus	Ligase dependent	Homologous recombination in insect cells	≥ 95%

V. Protein expression

5.3. Expression system

5.3.3. Cell-free protein expression system

Simple open system which influences:

- Protein folding
- Disulfide bond formation
- Incorporation of unnatural amino acids
- Protein stability
- Expression of toxic proteins

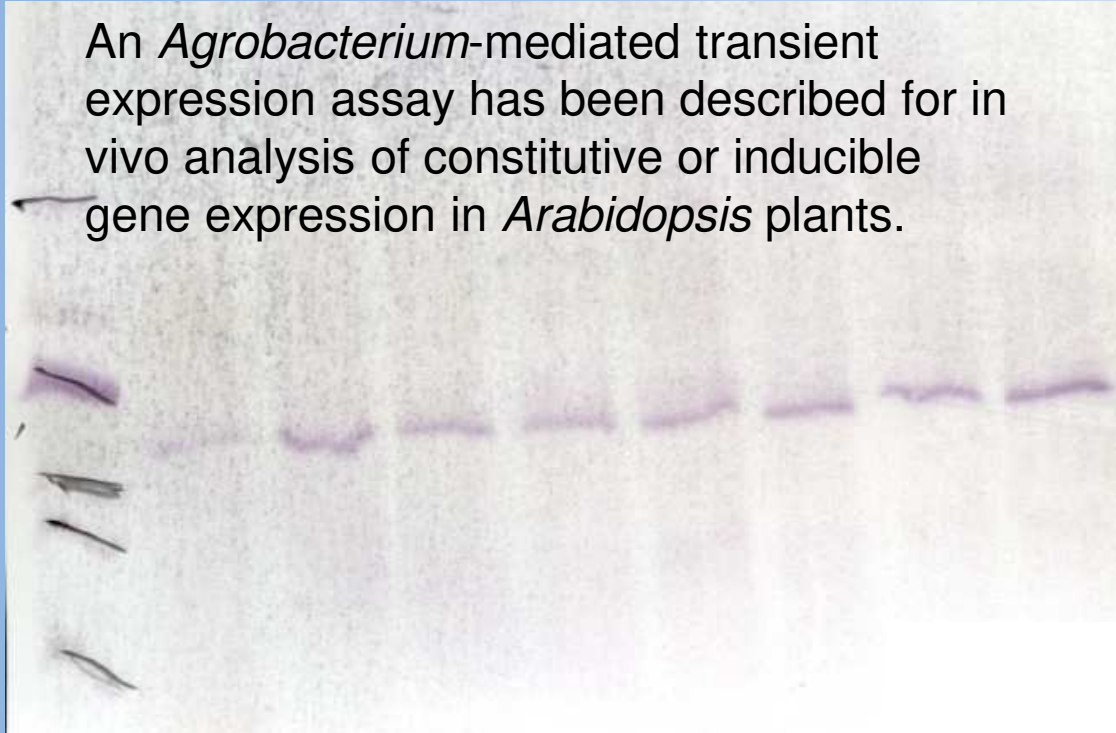
Use the machinery of *E. coli* S30

V. Protein expression

5.3. Expression system

5.3.4. Transient protein expression in tobacco leaves

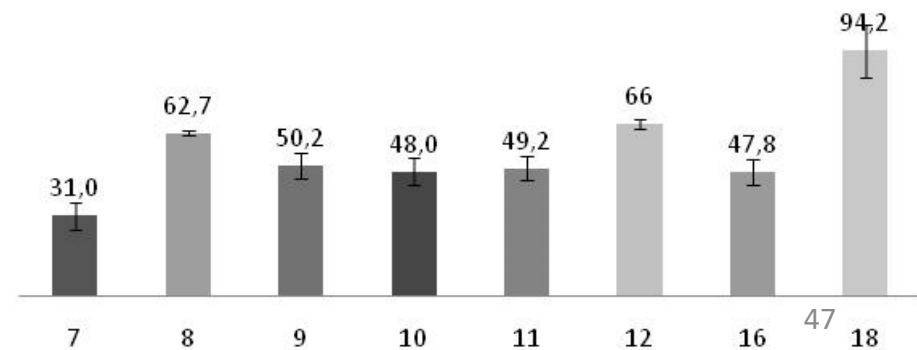
An *Agrobacterium*-mediated transient expression assay has been described for in vivo analysis of constitutive or inducible gene expression in *Arabidopsis* plants.



- Plant number: ca 30
- Weight of tobacco leaves: 7–10 g
- Number of tobacco leaves: 12–15
- Total: ca 3.5 kg~12–15 g protein~120–150 mg scFv

scFv x DHZR in Tobacco

■ 7 ■ 8 ■ 9 ■ 10 ■ 11 ■ 12 ■ 16 ■ 18



Before expression

- **Host strain** (E. coli, Yeast, mammalian cell, plant)
- **Expression vector** (plasmid)
- **Gene construct:**
 - ❖ domain and secondary structure prediction
 - ❖ codon usage
 - ❖ stability
 - ❖ toxicity
- **Buffer options for disintegration**