BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
FACULTY OF CHEMICAL AND BIOENGINEERING
GYÖRGY OLÁH DOCTORAL SCHOOL

IDENTIFICATION OF INTRINSICALLY UNSTRUCTURED PROTEINS FROM VARIOUS CELLULAR EXTRACTS AND LARGE SYSTEMATIC ERRORS IN THEIR QUANTITATIVE EXAMINATIONS

THE MAIN POINTS OF THE Ph.D. THESIS

Edit Szőllősi-Szitás

Témavezető: Péter Tompa D.Sc.

Konzulens: András Salgó D.Sc.

HUNGARIAN ACADEMY OF SCIENCES
Biological Research Center, Institute of Enzymology

2011.

1. INTRODUCTION

In the last 15-20 years many functionally active proteins and protein domains were described as unstructured proteins by NMR studies and other physicochemical examinations (circular dichroism spectroscopy, Gelfiltration chromatography, differencial scanning microcalorimery, SAXS etc.). Although these proteins are similar to the denaturated state of globular proteins, they play an important role in **regulation of cell cycle** ($p21^{Cip1}$/$p27^{Kip1}$ [5]), in **gene expression** (FlgM [6]), in **inhibition** (calpastatin [7]), in **signal transduction** [8] and assist the folding to RNA and protein as **chaperones** (Ribosomal S12 protein [9]).

The long standing dogma ("the lock and key" structure-function model of Fisher [10]) that tied protein function to a well defined trhree-dimensional structure has been increasingly challenged lately by the recognition that for many proteins / protein domains the native, functional state is intrinsically unstructured/disordered. That is why it is absolutely necessary that the model be expanded and reconsidered. The problem is that the number of the structurally well characterized disordered proteins / protein regions are limited because so far they have been identified only by the causal observation of the structural anomaly of proteins studied for their functional interest. However the protein disorder predictor algorithms suggest that there are much more IUPs in the various proteoms. For example the 20% of Jurkat cellular extracts and the 70% of human serum proteins are heat-stable by Kim at al. [11]. (The heat-stability is one of the basic disorder properties.) In the Swiss-Prot database the 11% of all the proteins and the 6-17 % of various proteoms are disorder and the 50-67% of proteins contains an amino acid part that is longer than 50 AA by the PONDR predictor [11].

1.1. AIMS

The aim of my Ph.D. thesis was to work out a simple method, a novel diagonal two-dimensional gelelectrophoresis technique for identification of novel IUPs from various cellular extracts. The method needs to separate the IUPs from globular proteins and each other. The principle of the novel technique is based on two essential disordered properties (heat-stability and resistance of chemical denaturation) combining a native gel (first dimension) with an 8 M urea second step. Besides my aim was to demonstrate the efficacy of the novel method using known and novel IUPs. The examinations of *Eschericia coli*, *Saccharomyces cerevisiae* és *Drosophila melanogaster* cellular extracts show that the technique is suitable for identifying of IUPs from higher order eucaryote organisms and it is applicable for quality comparison of proteins in case of subspecies of more complicated organisms.

In the course of the detailed structure analyses of proteins I observed that the wide-spread used convenient quantitative methods (Bradford process, BCA assay) provide strange, systematic errors. That is why my aim was to examine and demonstrate the unsuitability of these quantitative methods for IUPs.

2. LITERATURE BACKGROUND

Intrinsically unstructured proteins (IUPs)

The intrinsically unstructured proteins are similar to the denaturated states of globular proteins therefore they are frequently characterized as random coil-like proteins [12, 13] in spite of the fact that there are not real random coil structure under strong denaturating circumstances [14]. The disordered state of IUPs is not the consequence of the environmental effects but their important functional characteristic. The precise wording: the intrinsically unstructured proteins are able to form many kinds of structures quickly because of their lack of well-defined 3D structures.

IUPs, as a class differs from globular proteins in that they are enriched charged, hidrophil – so-called disorder promoting amino acids – and are depleted hidrophobe so-called order-promoting amino acids [15].

According to the function the class of IUPs can be devided in two parts. One of the parts contains those IUPs that do not bind to their partner in their function state. These entropic chains generate force or exert resistance against physical forces via global conformational changes. The other part of the IUPs class contains those proteins that take part in the process of molecular recognition that is to say bind permanently or temporarily to their partner molecules. This group is divisible in five subgroup: display sites, chaperones, effectors, assamblers, scavengers.

The lack of well-defined 3D structure gives rise to several considerable benefits in protein functionality. One of them is *the reversible but specific interaction* with their target molecules based on local structure elements of the IUPs. In the course of the interaction, the partner induces local folding in the IUP therefore the conformal entropy of protein decreases and the specificity and the strength of binding separate from each other (high specificity couples with small affinity). Additional benefit of structural disorder is *the increased speed of interaction* thanks to the high/large distance of interaction and the initial aspecific binding of the IUPs. This property of the IUPs can play an important role in regulation. Furthermore, some *IUPs can specifically bind several different partners* in a process termed binding promiscuity or one-to-many signaling and it has been suggested that the IUP in these cases may specifically adopt different structurures. This property is advantageous in case of assembling of multiprotein complexes and the spatial coordination of various partner molecules. In the latter case an IUP enables to attend different functions that can increase the complexity of protein-protein interactions without the increase of gene numbers.

The disordered feature was experimentally verified in case of about 500 proteins [16]. It was demonstrated that the disordered feature of the majority of these proteins is essential from the point of view their function. It is well known the transactivator domain of p53, that plays a role in regulation of cell division and forms interactions during apoptosis, is completely unstructured. The long central part of the BRCA1 protein, that interacts with DNA, p53, c-myc, Rad50, is mostly disordered. (The BRCA1 protein takes part in the development of breast cancer.) The coordinates of the C-terminal domain of the biggest subunit of RNA-polimerase II complex are absent from the X-ray structure of the protein so this domain is completely unstructured. However it is well known that this region is responsible for the synchronization of the many-sided development of mRNA maturation, if it is removed it will be incompatible with the life (it causes letal mutation).

3. THE MOST IMPORTANT MATERIALS AND METHODS

3. 1. The most important materials

Buffers of SDS-PAGE

Compressing gel buffer:	60.6 g TRIS; 4.0 g SDS; volume 1 l pH= 6.8,
Separating gel buffer:	181.7 g TRIS; 4.0 g SDS; volume 1 l; pH= 8.8,
Acrylamide stock solution:	100 g acrylamide; 2.64 g N, N'-methylene-bis-acrylamide; volume: 330 ml; filtering process
Running buffer (10× cc.):	60.6 g TRIS; 288.2 g glycine; 20.0 g SDS; volume: 2 l,

Buffers of native PAGE:
These only differ in SDS content from the buffers of SDS-PAGE, the other components and the quantities are the same.

Buffers of urea PAGE: The solutions of native PAGE + 8 M urea

3. 2. The novel diagonal two-dimensional gelelectrophoresis

Two characteristic properties (the heat-stability and the resistance to denaturating agents eg. 8 M urea) of the IUPs provide the principle of the novel gelelectrophoresis technique [1]. The separation of the IUPs from globular proteins in a cellular extract could be solved by the combination of a native gelelectrophoresis of heat-treated proteins followed by a second, denaturing gel contaning 8 M urea. The heat-treatment thus results in a good initial separation from globular proteins, most of which aggregate and precipitate that can be removed from cellular extracts by ultracentrifugalize so the detection of the IUPs will not be disturbed by these globular proteins.

In the native gel (first dimension), the IUPs and the rare heat-stable globular proteins will then be separated according to their charge/mass ratios.

After the primary separation the protein zone is cut off from the whole gel and it is put into the compressing gel buffer with 8 M urea for 45 min. The pH changing is absolutely necessary in order to obtaine the compressing effect. After 45 min. the first dimension gel is to be combined with an 8 M urea second step is rationalized by the usual structural indifference of IUPs to chemical denaturation by urea. As the urea is uncharged and the IUPs are just as „denatured" in 8 M urea as under native conditions, they are expected to run the same distance in the second dimension and end up along the diagonal. The heat-resistant globular proteins, on the other hand, will unfold in urea, slow down in the second gel, and accumulate above the diagonal. Due to this effective separation, the IUPs can be sent to subsequent MS identification.

3. 3. Basic methods for the structure analysing of IUPs

The application and the execution of the basic methods of the structure analysing of IUPs [17] are performed in the articles below:

- Csizmok V. et al. (2006). „ A novel two-dimensional electrophoresis technique for the identification of inrinsically unstructured proteins." Mol. Cell Proteomics 5(2): 265-273
- Szőllősi, E. et al. 2008. „Intrinsic structural disorder of Df31, a *Drosophila* protein of chromatin decondensation and remodeling activities." J. Proteome Res. 7: 2291-2299

4. RESULTS

My Ph.D. thesis covers two fields. One of them is the identification of IUPs from cellular extracts, while the other field focuses on the examinations concerning the inaccuracy of wide-spread used convenient quantitative methods. In the following part I would like to summarize the achieved results.

4. 1. 2DE examination of *E. coli* and *S. cerevisiae* extracts

After the novel 2D technique's parameters were optimalized the principle of method was controlled by studied and well characterized IUPs [1]. Henceforth *E. coli* and *S. cerevisiae* cellular extracts were analysed by 2D electrophoresis. The separated proteins were sent for identifying by MALDI-TOF MS [1]. We generated the PONDR pattern and the CH plot for all control and newly identified proteins.

Besides clear cut cases, there were several contradictory proteins that would be difficult to correctly classify by any single method. To clarify this ambiguity, we selected those cases where the domination of disorder is dubious. These proteins and some clear cut cases were cloned and expressed, and their disorder was assessed by various independent means, such as their position on the gel (on/above diagonal), PONDR pattern, location on the CH plot, coil content by CD, and apparent M_W and R_s by GF chromatography (**Table 1.**).

	Protein	2D position	PONDR (%)	CH plot	$coil_{CD}$ (%)	$M_{w,app}/M_w$
Controls	**ERD10**	ON	64.23	8.68	82.8	5.0[a]
	NACP		37,10	-1,67	81,3	4,5
	CSD1		81,56	2,95	78,8	4,5
	MAP2c		74,95	0,71	88,6	5,0
	b-kazein		41,15	-6,13	74,5	4,0
	IPMDH	ABOVE	59,42	-7,37	28	1,8[a,b]
	ovalbumin		27,72	-6,92	32,0	1,0
E. coli	**dnaKs**	ON	39.07	3.67	26.8	2.7
	yhgI		41,36	1,37	30,0	2,0
	ORF1		69,14	7,98	80,8	4,0[a]
	GlyH		53,49	-0,87	26,4	2,0
	ORF2	ABOVE	15,73	-0,81	4,5	4,7[c]
S. cerev.	**sRib**	ON	20.69	0.32	41.6	4.9
	TFIIA		63,29	6,26	76,1	4,4[a]
	tropoM		49,25	12,46	12,6	2,0
	Ubi6		51,70	18,13	30,9	2,0
	CaM	ABOVE	52,38	0,54	31,7	2,0

Table I. Comparison of various measures of disorder for selected proteins. [a] For some proteins, GF chromatography was repeated in 1 M NaCl to ensure that the high apparent M_w was not the result of the aggregation.

[b] In the case of IPMDH the value of $M_{w,app}/ M_w$ =1.8 reflects the formation of a dimer resistant to high salt concentrations.

[c] The high values of $M_{w,app} / M_w$ of ORF2 are probably the result of the coliled-coil structure of the protein as demonstrated by CD spectroscopy.

In all, the 2D technique correctly identified the global folded/unfolded character of proteins, most closely agreed with the structur analysing technique (GF and CD spectroscopy) and identified both extended (U or PMG type) and more compact (MG type) disordered proteins correctly.

4.2. 2DE examination of larva and imago *Drosophila melanogaster* extracts

The 2D/MS technique was applied to the examination of the proteome of larva and imago *Drosophila melanogaster*, predicted to contain as much as 17% fully disordered proteins. Several potential IDPs have been identified. The separated proteins were sent for identifying by MALDI-TOF MS. The amino acid sequences of the identified proteins were examined by PONDR and IUPred predictor and CH plot [2]. Only the functionally and structurally interest proteins (CG9894-PA, Df31 and Tropomyosin) were further characterized. The simple, indirect diagnostic tests (**heat-stability** and **proteolytic sensitivity**) did not deny the disorder characters of the examined *D. melanogaster* proteins. BSA was used as control globular protein.

The results of **Far-UV Circular Dichroism (CD) spectroscopy** show a typical disorder spectrum (large negative ellipticity near 200 nm) in case of the native (pH=7.5) Df31 and CG9894-PA protein. These proteins do not contain any secondary-stuctural elements because of the negative shoulder between 210-230 nm is not present in the spectrum. But in the CD spectrum of the Tropomyosin this negative shoulder is present therefore it contains secondary-stuctural elements.

^{1}H NMR measurements confirmed the results of CD spectroscopy. Considering the molecular mass of Df31, the one-dimensional ^{1}H NMR spectrum has unexpectedly sharp resonance lines, and a lack of signal dispersion in both the low-field and high-field regions. Note that neither low-field (eg. NH resonance) nor high-field proton resonance (eg. methyl resonances of Val, Ile or Leu) can be detected as isolated NMR signals [2]. But the NMR signals of Tropomyosin showed robust overlapping that is the properties of globular proteins. It was established that in the Tropomyosin are more stable secondary-structural elements and Df31 is random coil state by NMR signals.

To determine the disorder the most reliable and accepted methods are the CD spectroscopy and the NMR measurements. Additional examinations are carried out only in rare cases. In case of the Df31 protein it was necessary to perform additional studies (GF chromatogrphy, cross-linking, DSC microcalorimetry) because Crevel et al. tried to describe the properties of clearly disordered Df31 protein from the point of view of globular proteins [18, 19].

In the **gel filtration chromatography**, on a Superdex 200 gel-filtration column, the 18 kDa Df31 protein elutes at an apparent M_w of about 67 kDa ($M_{w,app}=3.7\times M_w$). This confirms that the native Df31 chain, indicating a hydrodynamic behavior characteristic of IUPs. The difference between our value and the M_w reported earlier (250 kDa) might come from post-translational modification of Df31 originally isolated from D. melanogaster, as opposed to the recombinant protein used in our study [18].

In accord with the previous paragraph, high apparent M_w may in principle be explained by an oligomeric structure of the protein. The oligomeric structure of the globular proteins is carried out routinely by **cross-linking studies**. In my experiment it was used dimethyl-suberimidate reacts with Lys residues as cross-linker agent in case of a known monomeric IUP, α-synuclein (NACP), in case of the homodimeric glycolytic enzyme, phosphoglycerate-mutase (PGM), as well as in case of Df31. NACP and Df31 showed no sign of cross-linking even under the longest incubation time, 60 min, which indicates a monomeric structure whereas in PGM, a new band at about 2.5 times the original M_w appeared.

Differential scanning microcalorimetry (DSC) curves are characteristic of the lack of a cooperative structural transition in IUPs. To test if Df31 is disordered by this criterion, the temperature dependence of the heat capacity of Df31 and of a control globular protein, lysozyme was recorded. Lysozyme underwent a critical exothermic transition at 72°C, but the calorimetric curve of Df31 was entirely flat, thus, Df31 lacks a folded structure [2].

4.3. Quantitative analyses of IUPs

In the course of my experiment work of proteins, it was dicovered that the routine laboratory quantitative identifying methods of proteins cause large systematic errors in case of IUPs [3]. The inexact determination of protein quantity gives rise to numerous errors so the experimental results are falsified.

To compare the performance of different assays (Bradford process, BCA assay, UV absorbance assay, densitometric examination of SDS-PAGE), six globular proteins [20, 21] and eight IUPs were chosen for testing (**Table 2.**). The Arg content of IUPs was between 0-12%, the aromic amino acid (Tyr, Trp, Phe) content of IUPs was between 0-8.5%. Concentrations of the proteins were assayed by a variety of conventional methods. The results of quantitve analyses are showed in **Table 2**.

	Bradford assay	**BCA assay**	**% SDS-PAGE**	**205 nm**	**280nm**
IUPs					
λ-N	33±1	31±12	42±11	12±2	13±2
Caskin(603-804)	4±1	3±8	7±1	26±1	13±2
CSDI	5±1	72±8	11±1	78±1	311±10
FlgM	25±4	50±22	10±3,5	36±3	
GlyH	40±4	86±10	32±7	64±1	183±2
MAP2C	14±0	14±1	16±1	10±1	37±5
NACP	18±0	44±4	25±5	38±4	100±4
Titin(5805-6005)	28±5	39±5	31±5	39±1	52±6
Globular proteins					
ADH	69±3	85±9	101±5	111±2	94±9
BSA	97±2	104±13	108±10	102±2	80±2
Cytochrome *C*	144±15	122±24	104±9	146±1	361±5
Myoglobine	130±5	102±18	139±34	130±6	238±1
Ovalbumin	66±3	137±20	129±8	161±2	146±5
Trypsinogen	66±18	129±25	69±10	117±6	80±7

Table 2. Protein concentrations determined by colorimetric and UV absorbance assays [18]. For comparison of the results, the values are given as percentages of the absolute values determined by gravimetry. Data represen means ±SD of 8 to 10 separate experiments (colorimetric assays) or 3 to 5 separate experiment (UV absorbance). BSA was used as a protein standard for all of the assays.

The values determined show large negative systematic errors, sometimes exceeding an order of magnitude, which would propagate into deduced parameters. To make a comparison (mainly for globular proteins) between all three of the colorimetric assays it was declared that the Bradford assay is the least accurate. In contact with the pI_{acidic} and the pI_{basic} proteins it was determined that all colorimetric methods provided approximately equivalent values in cases of the pI_{basic} proteins, but in cases of pI_{acidic} IUPs the results were substantially better in the course of BCA assay [3].

In conclusion, my studies show that traditional and convenient assays for determining protein quantity are generally not applicable to IUPs.

5. THESES

1. A novel diagonal two-dimensional gelelectrophoresis technique was *made/worked out* for the identification of a large number of intrinsically unstructured proteins from various organisms. The technique provides a new mode for the separation IUPs from globular proteins and other IUPs [1].

2. Novel IUPs were identificated from *E. coli* and *S. cerevisiae* extracts. The efficiency of the technique was certified by the structure analyses of these proteins [1].

3. The novel methode is highly reproducible, is easy to perform, and is readily adaptable to high throughput studies. Its comparison with other experimental and bioinformatic techniques showed that it provides dependable assessment of global structural disorder even in contradictory cases [1].

4. In course of the 2D examination of larva and imago *D. melanogaster* extracts it was observed that there are much more disordered proteins in *D. melanogaster* extracts [2] than in *E. coli* extract or in *S. cerevisiae* extract. This observation amplifies the fact of the disorder is growing with the complexity of organisms [11, 22].

5. My presented results provided unambiguous proofs about the coiled-coil Tropomyosin of *D. melanogaster* is a member of the α-helical protein family and about the CG9894-PA and the Df31 protein are unequivocally the members of disordered protein family [2].

6. It was demonstrated that the routine laboratory quantitative identifying methods of proteins cause large systematic errors in case of IUPs. The results of the quantitative examinations of well studied IUPs and globular proteins showed that the inexact determination of protein quantity gives rise to numerous errors so the experimental results are falsified [3].

7. The only way about this quantity problem is the direct determination of protein quantity by measuring mass, nitrogen content by the micro-Kjeldahl method, or UV determination of posthydrolysis derivated amino acids [3].

6. APPLICABILITIES

6. 1. Verification of disordered character by novel 2DE technique

Although its resolving power does not match that of the conventional 2D technique, it enables specific applications in the rapidly advancing field of IUPs in two different but related directions.

Its first practical application is the rapid characterization of a single protein in terms of its IUP status [1]. As demonstrated (**Figure 1.**), a simple run of the 2D format can tell with high certainty if the protein is structured or unstructured. Thus, this simple approach of running two known IUPs with an uncharacterized protein obtained in a cloning/purification experiment can establish the order/disorder status of protein available in a minute quantity only. The well charaterized IUPs run to, or very near, the diagonal of second gel, the heat-denaturized globular proteins do not appear on the gel, whereas the heat-resistant globular proteins remain way above the diagonal.

Given that the technique can provide information on a protein of very small quantity and limited purity, it will be a useful complement to other techniques (CD and NMR spectoscopy) that are more demanding on protein quantity and quality.

Actually with a good antibody at hand, it may even be performed on cell/tissue extracts without prior purification.

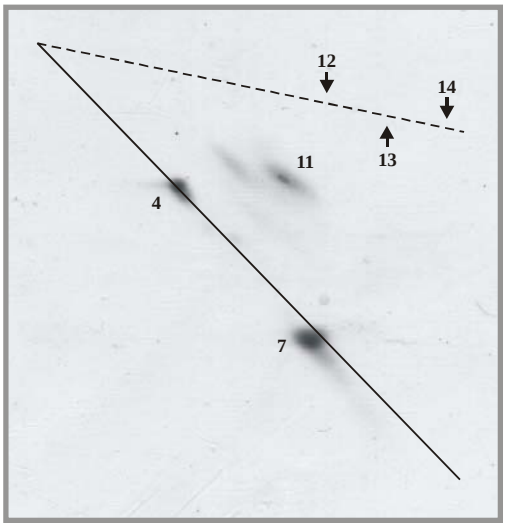

Fig. 1. The native / 8 M urea 2D electrophoresis of IUPs and globular proteins [1]
1 μg each of four globular proteins (11, fetuin; 12, IPMDH; 13, BSA; and 14, ovalbumin) and two IUPs (4, ERD10; and 7, CSD1) mixed in a total volume of 30 μl were boiled for 10 min. The precipitated protein was pelleted by centrifugation for 10 min at 10,000 × *g*, and the supernatant was run on a small format (8 × 6 cm) 7.5% native / 8 M urea 2D gel. The two IUPs define the diagonal. Three of the globular proteins (BSA, ovalbumin, and IPMDH despite its thermophilic origin) precipitate upon heat treatment and are absent from the second gel: their expected positions are marked. Serum fetuin resists heat treatment but runs above the diagonal in the second dimension.

6. 2. Identification of novel IUPs from various cellular extracts

The technique applicable for identification of novel IUPs from various cellular extracts (*Escherichia coli, Saccharomyces. cerevisiae, Drosophila melanogaster)* by combining the novel diagonal 2D gelelctrophoresis technique with MALDI-TOF MS. Beyond that the method open up new vistas for the qualiry characterization of various developmental state of species.

PUBLICATIONS ON THE SUBJECT OF THE Ph.D. THESIS

1. Csizmók, V., Szőllősi, E., Friedrich, P., Tompa, P. (2006). „A novel two-dimensional electrophoresis technique for the identification of intrinsically unstructured proteins." Mol. Cell Proteomics 5(2): 265-273 (I.F.: 8,8; I.: 8)

2. Szőllősi, E., Bokor, M., Bodor, A, Perczel, A., Klement, E., Medzihradszky K. F., Tompa, K., Tompa, P. (2008). „Intrinsic structural disorder of Df31, a *Drosophila* protein of chromatin decondensation and remodelling activities." J. Proteome Res. 7: 2291-2299 (I.F.: 5,13; I.: 1)

3. Szőllősi, E., Házy, E., Szász, Cs., Tompa, P. (2007). Large systematic errors compromise quantitation of intrinsically unstructured proteins." Anal. Biochem. 360: 321-323 (I.F.: 3,287; I.: 2)

4. Szőllősi-Szitás E., Kalmár, L., Tompa, P. (2011). „Investigation of IUPs by high throughput proteomic and bioinformatic methods" Biokémia. In press.

ORAL AND POSTER PRESENTATION ON THE SUBJECT OF THE Ph.D. THESIS ON THE NATIONAL CONFERENCES

1. Veronika Csizmók, Edit Szőllősi, Péter Friedrich, Péter Tompa: Identification of intrinsically unstructured proteins by a novel 2D electrophoresis technique
   The IX. Conference organized by Molecular Biology Section of The Hungarian Ciochemistry Association (Poster presentation), Sopron, May 10-13, 2004.

2. Edit Szőllősi: Identification of intrinsically unstructured proteins by a novel 2D electrophoresis technique. (Oral presentation)
   3nd Conference of Ph.D. sudents at faculty of chemical engineering (BME), Budapest, February 7, 2006.

3. Edit Szőllősi: a novel 2D electrophoresis technique for larval and imago *Drosophila melanogaster* cellular extracts. (Oral presentation)
   4th Conference of Ph.D. sudents at faculty of chemical engineering (BME), Budapest, February 7, 2007.

PRESENTATION ON THE SUBJECT OF THE Ph.D. THESIS ON THE INTERNATIONAL CONFERENCES

1. Veronika Csizmók, Edit Szőllősi, Péter Friedrich, Péter Tompa: Identification of intrinsically unstructured proteins by a novel 2D electrophoresis technique
   EMBO/HHMI Central European Scientist Meeting, Budapest, February, 2005.

2. Veronika Csizmók, Edit Szőllősi, Péter Friedrich, Péter Tompa: Identification of intrinsically unstructured proteins by a novel 2D electrophoresis technique
   Gordon Research Conferences, ProteinsUSA, Plymouth, June, 2005.

REFERENCES

1. Csizmók, V., Szollosi, E., Friedrich, P., Tompa, P. , *A novel two-dimensional electrophoresis technique for the identification of inrinsically unstructured proteins.* Mol. Cell Proteomics, 2006. **5**(2): 265-273.
2. Szőllősi, E. et al., *Intrinsic structural disorder of Df31, a Drosophila protein of chromatin decondensation and remodeling activities.* J. Proteome Res., 2008. **7**: 2291-2299.
3. Szőllősi, E., Házy, E., Szász, Cs., Tompa, P. *Large systematic errors compromise quantitation of intrinsically unstructured proteins.* Anal. Biochem., 2007. **360**: 321-323.
4. Szőllősi-Szitás E., Kalmár, L., Tompa, P. (2011). „*Investigation of IUPs by high throughput proteomic and bioinformatic methods*" Biokémia. In press.
5. Kriwacki, R.W., et al., *Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity.* Proc Natl Acad Sci. 1996. **93**(21): 11504-9.
6. Daughdrill, G.W., et al., *The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28.* Nat Struct Biol, 1997. **4**(4):285-91.
7. Konno, T., et al., *A circular dichroism study of preferential hydration and alcohol effects on a denatured protein, pig calpastatin domain I.* Biochim. Biophys. Acta, 1997. **1342**(1): 73-82.
8. Clodi, E., Semrad, K., and Schroeder, R. , *Assaying RNA chaperone activity in vivo using a novel RNA folding trap.* EMBO J., 1999. **18**:3776-3782.
9. Triezenberg, S.J., *Structure and function of transcriptional activation domains.* Curr. Opin. Genet. Dev., 1995. **5**(2): 190-6.
10. Fiescher, E., *Einfluss der configuration auf die wirkung den enzyme.* Berichte der Deutsche Chemie Gesichte, 1894. **27**: 2985-2993.
11. Dunker, A.K., et al., *Intrinsic protein disorder in complete genomes.* Genome Inform Ser Workshop Genome Inform, 2000. **11**: 161-171.
12. Schweers, O., et al., *Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure.* J Biol Chem, 1994. **269**(39): 24290-7.
13. Gast K, D.H. et al., *Prothymosin alpha: a biologically active protein with random coil conformation.* Biochemistry, 1995 **34**(40): 3211-8.
14. Shortle, D., *The denatured state (the other half of the folding equation) and its role in protein stability.* Faseb J, 1996. **10**(1): p. 27-34.
15. Dunker, A.K., et al., *Intrinsically disordered protein.* J Mol Graph Model, 2001. **19**(1): 26-59.
16. Sickmeier, M., Hamilton, J.A., LeGall, T., *Disperot: The database of disorder proteins.* Nucleic Acids Res, 2007. **35**: D786-D793.
17. Receveur-Brechot, V., Bourhis, J. M., Uversky, V. N., Canard, B., Longhi, S., *Assessing protein disorder and induced folding.* Proteins, 2005. **62**: 24-25.
18. Crevel, G., Cotterill, S., *Df31, a sperm decondensation factor from Drosophila melanogaster: purification and characterization.* Embo J, 1995. **14**: 1711-1717.
19. Crevel, G., Huikeshoven, H., Cotterill, S., *Df31 is a novel nuclear protein involved in chromatin structure in Drosophila melanogaster.* J. Cell. Sci, 2001. **114**: 37-47.
20. Bradford, M.M. *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal. Biochem. 1976. **72:** 248-54.
21. Congdon, R.W., Muth, G. W., Splittgerber, A. G., *The binding interaction of coomassie blue with proteins.* Anal Biochem, 1993. **213**: p. 407-413.
22. Dunker, A.K., et al., *Intrinsic disorder and protein function.* Biochemistry, 2002. **41**(21): 6573-82.

NOTE