

Towards a matrix mechanics framework for dynamic protein network

Sanjoy K. Bhattacharya

Received: 18 November 2009 / Revised: 18 December 2009 / Accepted: 29 December 2009 / Published online: 9 January 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Protein–protein interaction networks are currently visualized by software generated interaction webs based upon static experimental data. Current state is limited to static, mostly non-compartmental network and non time resolved protein interactions. A satisfactory mathematical foundation for particle interactions within a viscous liquid state (situation within the cytoplasm) does not exist nor do current computer programs enable building dynamic interaction networks for time resolved interactions. Building mathematical foundation for intracellular protein interactions can be achieved in two increments (a) trigger and capture the dynamic molecular changes for a select subset of proteins using several model systems and high throughput time resolved proteomics and, (b) use this information to build the mathematical foundation and computational algorithm for a compartmentalized and dynamic protein interaction network. Such a foundation is expected to provide benefit in at least two spheres: (a) understanding physiology enabling explanation of phenomenon such as incomplete penetrance in genetic disorders and (b) enabling several fold increase in biopharmaceutical production using impure starting materials.

Keywords Protein–protein interaction · Dynamic network · Channeling · Matrix mechanics · Mathematical foundation · Time-resolved proteomics

Description

Static and single node protein networks

Protein–protein interaction information can be derived from numerous individual (such as immunoprecipitation) or high throughput (such as yeast two-hybrid, phage display and mass spectrometry) experiments. The networks of these interactions are currently analyzed and created by computer programs based on manually curated static experimental data. Protein interactions within a living cell are dynamic and occur with strikingly high specificity. A static network enables only a very limited imputed glimpse of a possible dynamic molecular scenario. Current approaches are not focused on capturing time resolved interactions of proteins, especially in a high throughput manner. A second problem is the lack of satisfactory mathematical foundations for particle interactions within viscous liquids (a situation that proteins within the cytoplasm experience). Current approaches also do not take into account the compartmentalization of proteins.

If we compare the current protein interaction maps they are highway maps for airplanes that is, a three dimensional positions collapsed on the highway as if the z -axis does not exist. Secondly to navigate only an equivalent of highway with standing but not moving vehicles are shown.

In order to develop a mathematical molecular interaction foundation (*not model*), a large experimental data set is necessary. Once this foundation is in place, computer programs capable of constructing dynamic networks for time resolved interactions from the experimental data can be used to simulate models and help predict sequential events accommodating additional incremental static or dynamic interaction data by extrapolation. A *foundation*

S. K. Bhattacharya (✉)
Bascom Palmer Eye Institute, University of Miami,
1638 NW 10th Avenue, Suite 706A, Miami, FL 33136, USA
e-mail: sbhattacharya@med.miami.edu

will include the assumption of equivalent of a highway and also moving players.

Mathematical foundation

The basis on which the mathematical statements are true or the concepts based on which the mathematical statements are build can be construed as mathematical foundation. In contrast a model is a subset of mathematical statements that may enable simulation of a limited situation. Matrix mechanics/quantum mechanics for particles in an isolated gaseous state was developed in 1925. This mathematical foundation was formulated to describe microscopic observables and can be pertinent in describing protein–protein interactions. In matrix mechanics the physical properties of particles have been interpreted as matrices that evolve in time. The foundation was formulated based on discrete transition states and transition probabilities. In the viscous state time resolved protein movement is expected to show discrete movement and interactions and a similar formulation may enable description of such states.

However, for interacting particles in a viscous fluid, such a developed foundation does not yet exist. We propose here necessity to develop a mathematically based foundation for intracellular protein interactions in two steps (a) trigger and capture the dynamic molecular change for a select subset of proteins using high throughput time resolved proteomic mass spectrometry with capture of compartmentalization information to the maximum extent possible and, (b) use this information to build a mathematical foundation and computational algorithm for a dynamic compartmentalized protein interaction network. Such an approach may benefit from several model systems to capture protein–protein interaction for select proteins that are part of cytoskeletal channeling, short and long lasting signaling. Five or more different model systems encompassing multicellular organisms is expected to provide sufficient information to build a time resolved channeling events. To capture variability amid diversity information derived from land and aquatic animal, plant and animals capable of aviation may be included in such analysis. The inherent non-linearity in biological systems will necessitate inclusion of a number of independent systems to build the foundation.

Examples of systems

The time resolved proteomic analysis may utilize systems such as: (a) the effects of adrenaline on protein–protein interactions in a time resolved manner in neurons (b) the light-driven reaction in photoreceptors of mammalian retina (Calvert et al. 2006) (c) the touch-me-not plant (*Mimosa pudica*) in response to touch (Fromm and Lautner

2007) (d) the production of mucous by the hagfish (*Eptatretus stouti*) in response to stress or touch (Venter et al. 2001), and (e) analysis of protein expression in the brain of zebra finches (*Taeniopygia guttata*) in response to bird songs by a tutor (Mello et al. 1992). Such analysis will enable capturing time resolved data for cytoskeletal molecules as well as for rapid and stable long term cytosolic signaling. An example of rapid signaling process are phosphorylation and acetylation that are carried out by two opposite pair of enzymatic activities (kinase-phosphatase and acetylase-deacetylases) (Selvi and Kundu 2009; Shenolikar 1986). The deimination is an example of stable posttranslational modification catalyzed by peptidyl arginine deiminases and renders modification of protein-bound arginines but not free arginines (Bhattacharya 2009). In contrast to phosphorylation and acetylation, no known reversal enzyme exists for removal of deimination rendering it a relatively long term signaling event. The time resolved proteomic analysis of five systems will enable development of mathematical foundation differentiating structural and signaling components as well as normal cellular operating/metabolic processes. Combining information from multiple systems is thought to be better than any one of them and information derived from five systems will be sufficient to build the foundation (Stadnik et al. 2008).

In terms of developing a mathematical foundation in this direction, disparate tools exists that which in combination will act as guiding elements for analyzing high throughput time resolved data.

However, as stated above neither data nor mathematical foundation or computational programs exist that enables constructing compartmentalized and dynamic time resolved interaction networks.

Compartmentalization and channeling of proteins: cellular infrastructure and effects

Within cells sophisticated compartmentalization often prohibits non-specific interactions and facilitates specific interactions. Facilitated interactions are often stated as channeling. Studies have revealed that for macromolecule probes and for metabolite sized probes the cytoplasmic diffusion coefficients are 10–100 folds and 2–4 times smaller, respectively than the corresponding values in water. This suggests shorter traversed path and presence of channeling for probe molecules in the cytoplasm (Uyeda 1992). In mammalian cells glycolytic enzymes have been shown to exhibit biphasic partitioning between cytoplasmic and cytomatrix-bound states. The higher concentrations of cytomatrix associated enzymes augments flow or channeling of substrates (Ovadi and Srere 1992; Uyeda 1992). Inconsistency between experimentally determined concentrations

of metabolites in situ and kinetic properties of individual enzymes and the metabolic flux in site supports the existence of channeling within the citric acid cycle in mitochondria (Keleti and Welch 1984) and a similar situation exists in chloroplast. The computer models have shown that hypothetical steady-state processes with channeling do (Mendes et al. 1992) or do not (Cornish-Bowden and Cardenas 1993) provide a kinetic advantage. The metabolic channeling is now widely accepted concept. In metazoans, a cell is highly compartmentalized and most input and output of material across the cellular boundary is controlled. Extracellular matrix space also has some control of inflow and outflow. Within the cell the endoplasmic reticulum and other organelles have highly organized and compartmentalized structure and dynamic movement of materials is well organized. Towards developing a mathematical foundation where like matrix mechanics the physical properties of proteins (or protein–protein interactions) interpreted as matrices that evolve in time have to take such compartmentalization into account. The infrastructure or the compartmentalization is well suited for regulation. It is the infrastructure and regulation that enables producing several fold production of the chemicals in response to stimuli (Chakrabarti et al. 2003). Such compartmentalization may also help prevent system failure due to a mutant protein in some individuals to have full interaction with its interactors. Such incomplete interactions due to compartmentalized manner may help explain incomplete penetrance for some mutants (Ashley and Warren 1995; Grigoryev et al. 2004; Murphy 1970; Shapiro et al. 1969).

Availability of mathematical tools

The differential equation to describe complex rate kinetics with boundary conditions as well as simplification functions such as Laplace and Fourier transform has been in existence for a long time. These mathematical tools can be considered as nuts and bolts to describe compartmentalized time resolved situations. Historically mathematical platforms have been developed to simulate chemical reactions (Gillespie 1977) including probabilistic algorithms according to deterministic kinetic rates of chemical reactions (Barbuti et al. 2008). However, such models are either inadequate or number of interacting molecules within cells are simply too large for such most such platforms. Attempts have been recently been made to characterize currents in non-equilibrium statistical mechanics. Non-zero currents in the non-equilibrium state in stochastic dynamics models of natural systems have been noted to characterize these systems (Debnath et al. 2006; Gaveau and Schulman 2009). Recently some platforms have been developed for simulation of large biochemical reaction networks (Sinityn et al. 2009; Slepoy et al. 2008). In parallel, attempts have been

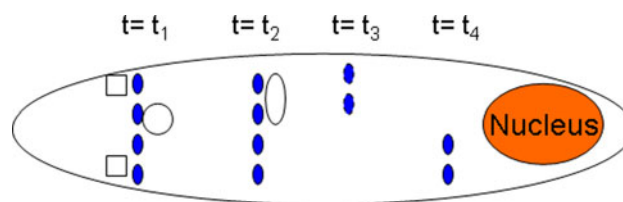


Fig. 1 Schematic diagram depicting the time dependent interaction of molecules within the boundary of a cell. Different molecules are shown by different shapes *circle, square* etc., *filled circle* (representing one type: molecule A) shift in time resolved interaction is illustrated, $t = t_1 \dots t_i$ represent different time points, where t_i is a given time point. The cell nucleus has been shown to have a boundary

made to determine the gross statistical properties of complex networks. Biological protein–protein networks have been found to be hierarchical scale free power networks (Barabasi 2009; Kitano 2007; Ravasz et al. 2002). However, beyond gross statistical properties, such analysis has not provided great deal of additional functionally meaningful insight. In recent years, a path integral approach has been used to develop variational approximations for the calculation of survival probability for rate processes with dynamical disorder. Jensen’s inequality has been used to derive both upper and lower bounds for survival probability (Harland and Sun 2007). An integration of these approaches is a starting point for the development of mathematical foundation for dynamic particle interaction in non-isolated and viscous environment (Chiam et al. 2006; Costa et al. 2009; Takahashi et al. 2005).

A simplistic time resolved protein–protein interaction within the cytosol has been depicted in Fig. 1. In this schematic diagram, the protein molecules A (represented by blue ellipsoids) are captured at different locations interacting with different proteins (hollow square or circles) or are degraded as in time $t = t_3$.

Ancillary techniques and devices for time resolved proteomics

Several techniques that accelerate fractionation (Patel et al. 2008), and methods to capture the time resolved capture of cell/tissue regions are increasingly becomes available. The devices such an ultra thin compressor for fast freezing and a laser fast protein cross-linker are under development. Power of one dimensional proteomic analysis for capture and identification of proteins from disparate systems are now well realized (Benjamin et al. 2008; Menegay et al. 2008; Sloley et al. 2007a, b). One dimensional proteomic analyses are quick and simple fractionation procedures that utilize separation in only one dimension (in contrast to two dimensional fractionation) such as 1D gel or one chromatographic column followed by mass spectrometric identification. One dimensional fractionation and analysis

is a compromise with reduced resolution in fractionation versus procedural loss (Zhou et al. 2005). Recently tandem affinity purification and mass spectrometry along with other methods and bioinformatic analysis has revealed the possibility of higher levels of proteome organization even in simplest of microorganism (Kuhner et al. 2009). These include assembly of proteins into larger, multiprotein complex entities and suggest sequential steps in biological processes, with sharing of components implicating multifunctionality of proteins. Such assembly and interactions are dynamic and evidence suggest wide-ranging temporal waves of protein–protein interactions with cells (Cohen et al. 2008; Li and Zeng 2009).

Benefits

Ability to have detailed dynamic network of interaction will expand our understanding and enhance our translational goals. A gene mutation usually shows an effect in an individual but does not in others is termed as *incomplete penetrance*. All individuals within an isogenic population are expected to have parallel and normal development. However, gene mutations affect the outcome of individual organisms differently. How the genetic control allows the penetrance of multiple fate for individuals harboring same mutation remain unclear. Some elegant studies utilizing sporulation mutants of *Bacillus subtilis* and *Clostridium oceanicum* have shown that initial expression of discrete morphological traits occurs at low penetrance of mutations. The stabilization of effect of mutation occurs by gradual adjustment of genetic parameters and often determines different fate for individual organisms harboring the same mutation (Eldar et al. 2009). Time resolved interaction data will demonstrate case studies as to what is different between two cellular systems that results in failure in one but not in the other systems. A generalized understanding will help extrapolate such data and enable prediction depending upon time resolved interactions and presence of non-specific surroundings. Such interaction details will also provide insight into anomalies in protein function, moonlighting functions of proteins and in addition shed light into the redundancies in the system. Time resolved data will enable capture of small molecule interactions and thus help unravel their interactions to bring out specific and non-specific side effects. Such time resolved data will leave gaps and will help bridge the gaps by algorithms which will predict obvious scope for interaction in some interactions and help identify new interactions through predictions. Lastly, cytoplasm is just not a bag of juice and in the environment of cytoplasm the interactions enable many fold higher production of some chemicals in response to a single stimuli (for example, epinephrine) amidst the presence of very many non-specific entities. Achieving such

production amidst “impurities” is a chemical or biochemical engineers dream (Chakrabarti et al. 2003). Learning dynamic interactions of proteins and small molecules will enable mimicking such production of biochemicals in reactors.

In conclusion two broad problems have been identified here: (a) a lack of time resolved and compartmental high throughput protein–protein (or biomacromolecular) interaction data and (b) a lack of a mathematical foundation for understanding and modeling particle interactions in a viscous state with surrounding non-specific interactions. To develop this foundation, data acquisition is necessary. Matrix mechanics, better known as quantum mechanics, was developed upon the observation of discrete transition spectra for the hydrogen atom. Lack of availability of equivalent data for particle interaction in viscous space is the greatest impediment to developing a mathematically based biological molecular interaction foundation. The development of mathematical foundation enabling compartmentalized and time resolved biomolecular interaction data will help translational goals in the following ways: (a) explain the phenomenon of *incomplete penetrance* of gene mutations, (b) the efficacy and side effects of modulators such as small molecules on intended protein targets, and (c) extrapolation of new interactions or lack of identified molecules in a time resolved shifting network and (d) expand our understanding of enhancing production of biochemicals in presence of surrounding impurities.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Ashley CT Jr, Warren ST (1995) Trinucleotide repeat expansion and human disease. *Annu Rev Genet* 29:703–728
- Barabasi AL (2009) Scale-free networks: a decade and beyond. *Science* 325:412–413
- Barbuti R, Maggiolo-Schettini A, Milazzo P, Tiberi P, Troina A (2008) Stochastic calculus of looping sequences for the modeling and simulation of cellular pathways. *Trans Comput Syst Biol* IX:86–113
- Benjamin A, Kashem M, Cohen C, Caldwell Busby JA, Salgado-Commissariat D, Helekar SA, Bhattacharya SK (2008) Proteomics of the nucleus ovoidalis and field L brain regions of zebra finch. *J Proteome Res* 7:2121–2132
- Bhattacharya SK (2009) Retinal deimination in aging and disease. *IUBMB Life* 61:504–509
- Calvert PD, Strissel KJ, Schiesser WE, Pugh EN Jr, Arshavsky VY (2006) Light-driven translocation of signaling proteins in vertebrate photoreceptors. *Trends Cell Biol* 16:560–568
- Chakrabarti S, Bhattacharya S, Bhattacharya SK (2003) Biochemical engineering: cues from cells. *Trends Biotechnol* 21:204–209
- Chiam KH, Tan CM, Bhargava V, Rajagopal G (2006) Hybrid simulations of stochastic reaction-diffusion processes for

- modeling intracellular signaling pathways. *Phys Rev E Stat Nonlin Soft Matter Phys* 74:051910
- Cohen AA, Geva-Zatorsky N, Eden E, Frenkel-Morgenstern M, Issaeva I, Sigal A, Milo R, Cohen-Saidon C, Liron Y, Kam Z, Cohen L, Danon T, Perzov N, Alon U (2008) Dynamic proteomics of individual cancer cells in response to a drug. *Science* 322:1511–1516
- Cornish-Bowden A, Cardenas ML (1993) Channelling can affect concentrations of metabolic intermediates at constant net flux: artefact or reality? *Eur J Biochem* 213:87–92
- Costa MN, Radhakrishnan K, Wilson BS, Vlachos DG, Edwards JS (2009) Coupled stochastic spatial and non-spatial simulations of ErbB1 signaling pathways demonstrate the importance of spatial organization in signal transduction. *PLoS One* 4:e6316
- Debnath A, Chakrabarti R, Sebastian KL (2006) Rate processes with dynamical disorder: a direct variational approach. *J Chem Phys* 124:204111
- Eldar A, Chary VK, Xenopoulos P, Fontes ME, Loson OC, Dworkin J, Piggot PJ, Elowitz MB (2009) Partial penetrance facilitates developmental evolution in bacteria. *Nature* 460:510–514
- Fromm J, Lautner S (2007) Electrical signals and their physiological significance in plants. *Plant Cell Environ* 30:249–257
- Gaveau B, Schulman LS (2009) Currents in nonequilibrium statistical mechanics. *Phys Rev E Stat Nonlin Soft Matter Phys* 79:021112
- Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361
- Grigoryev DN, Finigan JH, Hassoun P, Garcia JG (2004) Science review: searching for gene candidates in acute lung injury. *Crit Care* 8:440–447
- Harland B, Sun SX (2007) Path ensembles and path sampling in nonequilibrium stochastic systems. *J Chem Phys* 127:104103
- Keleti T, Welch GR (1984) The evolution of enzyme kinetic power. *Biochem J* 223:299–303
- Kitano H (2007) Biological robustness in complex host-pathogen systems. *Prog Drug Res* 64:241–263
- Kuhner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, Rode M, Yamada T, Maier T, Bader S, Beltran-Alvarez P, Castano-Diez D, Chen WH, Devos D, Guell M, Norambuena T, Racke I, Rybin V, Schmidt A, Yus E, Aebersold R, Herrmann R, Bottcher B, Frangakis AS, Russell RB, Serrano L, Bork P, Gavin AC (2009) Proteome organization in a genome-reduced bacterium. *Science* 326:1235–1240
- Li RX, Zeng R (2009) Dynamic proteomics for investigating the response of individual cancer cells under drug action. *Expert Rev Proteomics* 6:19–21
- Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. *Proc Natl Acad Sci USA* 89:6818–6822
- Mendes P, Kell DB, Westerhoff HV (1992) Channelling can decrease pool size. *Eur J Biochem* 204:257–266
- Menegay M, Lee D, Tabbara KF, Cafaro TA, Urrets-Zavalía JA, Serra HM, Bhattacharya SK (2008) Proteomic analysis of climatic keratopathy droplets. *Invest Ophthalmol Vis Sci* 49:2829–2837
- Murphy EA (1970) Incomplete penetrance and the ascertainment of recessive characters in man I. Introduction and distribution theory. *J Chronic Dis* 23:9–27
- Ovadi J, Srere PA (1992) Channel your energies. *Trends Biochem Sci* 17:445–447
- Patel N, Solanki E, Picciani R, Cavett V, Caldwell-Busby JA, Bhattacharya SK (2008) Strategies to recover proteins from ocular tissues for proteomics. *Proteomics* 8:1055–1070
- Ravasiz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL (2002) Hierarchical organization of modularity in metabolic networks. *Science* 297:1551–1555
- Selvi RB, Kundu TK (2009) Reversible acetylation of chromatin: implication in regulation of gene expression, disease and therapeutics. *Biotechnol J* 4:375–390
- Shapiro LR, Raab EL, Leopold IH, Hirschhorn K (1969) Hereditary optic atrophy. An autosomal dominant with incomplete penetrance. *Arch Ophthalmol* 81:359–362
- Shenolikar S (1986) Control of cell function by reversible protein phosphorylation. *J Cyclic Nucleotide Protein Phosphor Res* 11:531–541
- Sinitzyn NA, Hengartner N, Nemenman I (2009) Adiabatic coarse-graining and simulations of stochastic biochemical networks. *Proc Natl Acad Sci USA* 106:10546–10551
- Slepoy A, Thompson AP, Plimpton SJ (2008) A constant-time kinetic Monte Carlo algorithm for simulation of large biochemical reaction networks. *J Chem Phys* 128:205101
- Sloley S, Smith S, Algeciras M, Cavett V, Busby JA, London S, Clayton DF, Bhattacharya SK (2007a) Proteomic analyses of songbird (Zebra finch; *Taeniopygia guttata*) retina. *J Proteome Res* 6:1093–1100
- Sloley S, Smith S, Gandhi S, Busby JA, London S, Luksch H, Clayton DF, Bhattacharya SK (2007b) Proteomic analyses of Zebra finch optic tectum and comparative histochemistry. *J Proteome Res* 6:2341–2350
- Stadnik J, Ganzha M, Paprzycki M (2008) Are Many Heads Better Than One-On Combining Information from Multiple Internet Sources. In: Kacprzyk J (ed) *Studies in Computational Intelligence*, vol 162. Springer, Berlin, pp 177–186
- Takahashi K, Arjunan SN, Tomita M (2005) Space in systems biology of signaling pathways—towards intracellular molecular crowding in silico. *FEBS Lett* 579:1783–1788
- Uyeda K (1992) Interactions of glycolytic enzymes with cellular membranes. *Curr Top Cell Regul* 33:31–46
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J,

- Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X (2001) The sequence of the human genome. *Science* 291:1304–1351
- Zhou S, Bailey MJ, Dunn MJ, Preedy VR, Emery PW (2005) A quantitative investigation into the losses of proteins at different stages of a two-dimensional gel electrophoresis procedure. *Proteomics* 5:2739–2747