

A Network-based Analysis of Polyanion-binding Proteins Utilizing Yeast Protein Arrays^{*}

Nazila Salamat-Miller[‡], Jianwen Fang[§], Christopher W. Seidel[¶], Aaron M. Smalter[‡], Yassen Assenov^{||}, Mario Albrecht^{||}, and C. Russell Middaugh^{‡**}

The high affinity of certain cellular polyanions for many proteins (polyanion-binding proteins (PABPs)) has been demonstrated previously. It has been hypothesized that such polyanions may be involved in protein structure stabilization, stimulation of folding through chaperone-like activity, and intra- and extracellular protein transport as well as intracellular organization. The purpose of the proteomics studies reported here was to seek evidence for the idea that the nonspecific but high affinity interactions of PABPs with polyanions have a functional role in intracellular processes. Utilizing yeast protein arrays and five biotinylated cellular polyanion probes (actin, tubulin, heparin, heparan sulfate, and DNA), we identified proteins that interact with these probes and analyzed their structural and amino acid sequence requirements as well as their predicted functions in the yeast proteome. We also provide evidence for the existence of a network-like system for PABPs and their potential roles as critical hubs in intracellular behavior. This investigation takes a first step toward achieving a better understanding of the nature of polyanion-protein interactions within cells and introduces an alternative way of thinking about intracellular organization. *Molecular & Cellular Proteomics* 5: 2263–2278, 2006.

A living cell consists of a number of only partially defined compartments in constant communication. Both the interior and exterior of a cell and its compartments possess regions of high negative charge density. Exterior negatively charged entities such as phospholipids, proteoglycans, and polysialic acids as well as intracellular components including inositol phosphates, nucleotides, actin and tubulin microfilaments, DNA, RNA, and ribosomes render cells an ultimately highly crowded and extremely polyanionic environment. Considering the nature of this environment, it seems reasonable to

postulate that numerous nonspecific interactions between cellular polyanions and proteins can take place. Cellular polyanions are of important functional significance to cells because at a minimum they (a) perform regulatory functions (e.g. binding of growth factors to proteoglycans) (1, 2), (b) transfer genetic information (RNA/DNA), (c) play a multitude of roles in cell structure and dynamics (actin and tubulin), (d) potentially serve as chaperones for protein folding (3–5), (e) stabilize proteins (6–8), and (f) perhaps facilitate non-classical transport of proteins into and out of the cell (9–11). In a recent exploration of the nonspecificity of polyanion-protein interactions, two-dimensional gels of proteins from cellular extracts in the presence and absence of matrix-bound polyanions were compared (12). It was demonstrated that hundreds to thousands of COS-7 proteins interacted with polyanions under the experimental conditions used. The polyanion with the greatest extent of interaction was heparin, binding 944 of 1,751 proteins resolved. It was observed that no direct relationship apparently exists between the overall net charge of the identified proteins and their binding potential to polyanions. This presumably reflects the presence of highly localized regions of positive charge in PABPs¹ and negates any requirement for an overall positive charge for protein-polyanion interaction (12). It was also observed that certain proteins in COS-7 cell extracts bound to actin, tubulin, and DNA with little obvious preference for any specific polyanion. This indiscriminate behavior of PABPs suggests that the observed interactions were not highly selective (specific) despite their high affinity. Examples of such nonspecific interactions of this sort have been observed previously. For instance, heparin-induced antibodies, which are characteristic of heparin-induced thrombocytopenia, were initially thought to be specific for heparin. Nevertheless an array of other linear polyanions such as heparan sulfate (HS), dextran sulfate, polyvinyl sulfate,

From the [‡]Department of Pharmaceutical Chemistry, University of Kansas, Lawrence, Kansas 66047, [§]Bioinformatics Core Facility, University of Kansas, Lawrence, Kansas 66045, [¶]Stowers Institute for Medical Research, Kansas City, Missouri 64110, and ^{||}Max Planck Institute for Informatics, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany

Received, June 28, 2006, and in revised form, August 24, 2006

Published, MCP Papers in Press, September 18, 2006, DOI 10.1074/mcp.M600240-MCP200

¹ The abbreviations used are: PABP, polyanion-binding protein; HS, heparan sulfate; DS, dextran sulfate; SF, Supplemental Fig(s); ST, Supplemental Table(s); SGD, *S. cerevisiae* Genome Database; MEME, Multiple Em for Motif Elicitation; GO, gene ontology; PCE, protein continuum electrostatics; BiNGO, Biological Network Gene Ontology tool; HB-EGF, heparin-binding epidermal growth factor-like growth factor; FGF, fibroblast growth factor; PDB, Protein Data Bank; PEO, polyethylene oxide; EDC, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide biotin-LC-hydrazide.

polyvinyl phosphate, and polyvinyl sulfonate were able to induce a biological effect similar to heparin (13). Many similar examples of such nonspecific interactions can be found (14–22).

The potentially wide ranging functions that polyanionic molecules could perform, their nonspecific *versus* specific binding to cellular proteins, and the highly crowded environment of cells suggest a number of fundamental questions. It seems quite possible that the crowded polyanionic environment of cells serves as some kind of functional surface/network where many (but by no means all) cellular events such as protein folding, protein-protein interactions, regulation, metabolism, and trafficking occur. Thus, we ask here whether it is plausible that cellular polyanionic surfaces might provide a matrix to direct the organization and function of a network of PABPs. We hypothesize that the cell can be viewed as a dense, functional network of polyanionic surfaces that facilitate, direct, and/or regulate certain cellular events through a gradient of relatively nonspecific interactions with PABPs. To investigate this hypothesis, we chose five model cellular polyanions to probe protein arrays containing 4,087 unique yeast proteins (23) and subsequently analyzed their potential interactions and functional significance.

EXPERIMENTAL PROCEDURES

Lyophilized G-actin from bovine muscle (molecular mass, 43 kDa), heparin and HS from porcine intestinal mucosa (molecular masses of 18 and 14 kDa, respectively), the metachromatic dye azure A, calf thymus double-stranded DNA, and dextran sulfate (average molecular mass, 5 kDa) were purchased from Sigma. Bovine brain tubulin (molecular mass, 100 kDa) was donated by Dr. Richard Himes of the division of Biological Sciences at the University of Kansas. EZ-link biotin-LC-hydrazide, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), MES, and EZ-link psoralen-PEO₃-biotin were purchased from Pierce. Protein array kits, including buffers and controls, were obtained from Invitrogen.

Biotinylation of Polyanions

Actin and Tubulin—Actin was dialyzed against 10 mM PBS overnight to avoid a cross-reaction between the Tris salt contained in its lyophilized form and the biotinylation reagent. Tubulin was dialyzed against two exchanges of 10 mM PBS for 3 h. Actin and tubulin (~2.0 mg/ml) were biotinylated according to the manufacturer's instructions using a biotin-XX-sulfosuccinimidyl ester (5 nmol/ μ l) included in the Invitrogen protein array kit. Excess biotin was removed using a gel filtration resin provided in the kit. The biotinylation efficiency was assessed by performing Tris-glycine SDS-PAGE and a Western blot of the biotinylated polyanions and the provided reference proteins according to standard protocols. Detection and visualization of samples were performed using a streptavidin-alkaline phosphatase conjugate and a chemiluminescent substrate, respectively.

Heparin and HS—Compounds were dissolved in 0.1 M MES buffer to a final concentration of ~4.0 mg/ml. The biotinylation reaction was conducted by addition of 25 μ l of biotin hydrazide in dry DMSO and 12.5 μ l of freshly prepared EDC with final concentrations of 50 and 5.0 mM, respectively, to the MES/heparin or MES/HS solution. Solutions were stirred overnight at room temperature. The carbodiimide chemistry coupled biotin-LC-hydrazide to heparin and HS through uronic acids (carboxylate groups) (24). Excess biotin was removed

using either dialysis cassettes (molecular weight cutoff 3,500) or polyacrylamide desalting columns (Pierce). Biotinylation efficiency was assessed by performing a dot blot on nitrocellulose membranes (Schleicher & Schuell). Approximately 2 μ l of each biotinylated sample was placed on the membrane, and the blot was air-dried. Biotinylated tubulin or BSA (both at a molar ratio of 9:1) and non-biotinylated polysaccharide or buffer were used as positive and negative controls, respectively. This procedure was repeated three times followed by blocking and washing of the membrane according to the Western blot protocols. Detection and visualization of samples were performed as above.

Calf Thymus DNA—A final concentration of 200 μ M EZ-link psoralen-PEO₃-biotin in doubly deionized H₂O was added to a 1 mg/ml solution of DNA in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.4) after which DNA was boiled and immediately cooled in dry ice/ethanol bath. The reaction centrifuge tube was placed on ice and irradiated with a long wavelength UV source (Mini Ralight lamp, UVP Inc., San Gabriel, CA) for 40 min. The biotinylated solutions of DNA were stored at –20 °C. Detection and visualization of biotinylated DNA was performed through a dot blot assay as described above.

Quantification of Polyanions

Actin and Tubulin—Final concentrations were determined by UV spectroscopy at 290 nm for actin and 280 nm for tubulin using extinction coefficients of 2.66×10^4 (25) and $1.15 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$,² respectively, with a cell pathlength of 1.0 cm.

Heparin and HS—Polysaccharide concentrations were determined using the metachromatic dye azure A (26). A stock solution of azure A was prepared in 0.2% (v/v) aqueous formic acid, pH 3.5, at a final concentration of $5.04 \times 10^{-5} \text{ M}$. One milliliter of the dye stock solution was titrated with 1–10 μ l of 1.0 mg/ml polysaccharide solution prepared in 10 mM PBS and vortexed (27). The absorbance of each sample was measured at 620 nm where the extent of disappearance of azure A was followed as a function of polysaccharide concentration. In each case, the non-biotinylated heparin and HS were used as standards for the biotinylated samples. UV absorbance confirmed that azure A was bound to both biotinylated and non-biotinylated forms of heparin and HS. These experiments were conducted after observing that biotinylated HS failed to bind to another metachromatic dye, 1,9-dimethylmethylene blue. This is in contrast to the non-biotinylated form, which shows binding to this dye.

Calf Thymus DNA—DNA was quantified using an optical density of 1.0 for a 50 μ g/ml solution at 260 nm.

Protein Array Experiments

The yeast protein microarrays contain 4,087 purified proteins from *Saccharomyces cerevisiae* ORFs that are expressed as N-terminal GST-His₆ fusion proteins, purified, and spotted in duplicate on a nitrocellulose-coated glass slide (23). These arrays were probed with the biotinylated polyanions in the current study. Due to the high background associated with biotinylated heparin and HS with this type of coating, modified surface chemistry arrays from Invitrogen were utilized. According to the manufacturer's protocol, the arrays were probed with 120 μ l of a 25–50 μ g/ml concentration of each biotinylated polyanion and exposed to streptavidin-Alexa Fluor® 647 followed by subsequent washings, drying, and scanning at 635 nm to acquire the fluorescence image. Protein arrays were scanned using a GenePix 4000B microarray scanner (Molecular Devices Co., Sunnyvale, CA) at 635 nm with 10- μ m pixel size. The results were analyzed using GenePix® Pro 5.0 software. Each image was analyzed using the appropriate grid available from Invitrogen. To investigate the nature of

² T. Mitchison, personal website, protocols.

any polyanion-protein interactions, each array, already probed with the biotinylated probes, was incubated with three different concentrations of NaCl solution (0.05, 0.15, and 0.75 M) in the probing buffer for 30 min, and the number of positive protein signals were reanalyzed after each experiment using the appropriate grid. For all scanned images at 0–0.75 M NaCl, the program “Prospector” (Invitrogen) was utilized to identify significant polyanion-protein interactions. Interactions that were considered statistically significant demonstrated a mean signal greater than median signal of all protein spots on the array plus one times the standard deviation (+1S.D.). This +1S.D. threshold has a Z-score that is greater than 1. For such threshold values, interacting proteins with low, medium, and high confidence levels are included in the dataset. Upon increase in the confidence level of interactions, +2S.D. and +3S.D. are introduced. We decided to include the greatest number of proteins involved in any type of binding to polyanions (specific or nonspecific) and, therefore, chose to study PABPs at the level of +1S.D. in most cases unless indicated otherwise. All results and data analyses were obtained from protein array experiments that included 0.15 M NaCl (corresponding to physiological ionic strength) unless indicated otherwise.

Competition Experiments with Dextran Sulfate (DS)

An experiment with a high concentration of DS (6 mM) was conducted that differed from the NaCl experiments in two ways. 1) Each array was exposed to DS for 24 h in contrast to the 30-min exposure time for the NaCl experiments, and 2) each experiment was conducted at room temperature as opposed to 4 °C for the NaCl experiments. Following the incubation of protein arrays with DS, they were washed, dried, and scanned, and the resulting images were analyzed using GenePix 5.0.

Searching for Positively Charged Amino Acids in the Sequence of PABPs Using MATLAB

An in-house MATLAB program (version 7.0.4; The MathWorks, Inc., Natick, MA) was used to search for the presence of the basic amino acids Lys, Arg, and His in PABPs (test set) and to calculate their amount (percentage) in each protein. Additionally this program was used to analyze percentage of basic amino acids in the population of all yeast proteins that were spotted on the array (population set). The translated protein sequence for each ORF was obtained from the *S. cerevisiae* Genome Database (SGD) batch download tool (db.yeast-genome.org/cgi-bin/batchDownload). The results from the test and population sets were compared utilizing a one-sample *t* test. This test calculates a two-tailed *p* value between the means of one particular basic amino acid percentage for proteins that interacted with each of the five different probes used and the same amino acids in the population set (QuickCalcs, GraphPad Software, Inc.). Furthermore in an effort to determine whether the enrichment of positive residues was more pronounced locally, we used Perl script to calculate the number of positive pitches in PABPs and other yeast proteins. The positive pitches are defined as a sequence run containing three or more positive residues of any five continuous residues.

pI Calculations

The code for the program for theoretical calculations of *pI* values for yeast PABPs was obtained from the code author.³ This program was used to generate and compare the theoretical *pI* values for the test set as well as the yeast protein population on the array. The results of the test and population sets were compared utilizing a one-sample *t* test.

³ C. Putnam, personal communication.

Searching for a Sequence Signature in PABPs Using MEME and InterPro

Sequence Signature Analysis—A sequence signature is defined as a highly conserved region, a sequence pattern that is found repeatedly in a group of related sequences. By this definition, a sequence signature could be a protein family, functional domain, functional site, or any conserved region of unknown function. Thus, the actual physical manifestation of a signature can vary greatly in size (28). In this study, sequence signatures were derived from MEME motifs. Signature analysis was performed for the sequences of the yeast PABPs binding to one or more polyanion probes using MEME (version 3.0.10) (29). MEME implements an unsupervised learning algorithm for discovering gapless signatures in a group of related protein sequences. In this study, we chose the maximum number of signatures to be 50 and minimum and maximum width of the signatures to range from 5 to 300 amino acids, respectively. The *E*-value cutoff was set to 0.1 to report only statistically significant results. In addition, the type of distribution was set to zero or one occurrence per sequence, and the minimum number of sites was set to three. Therefore, only signatures that were shared by three or more PABPs were discovered by the MEME search.

Sequence Signature Characterization—The consensus sequence of each signature model was searched for in all InterPro member databases using an on-line version of InterProScan (www.ebi.ac.uk/InterProScan/). InterPro is an integrated collection of the most commonly used databases of protein families, domains, and functional sites (30). The program InterProScan allows a user to search for sequence signatures in any number of these databases simultaneously.

Additionally in an effort to determine the distribution of known polyanion binding domains (protein domains that bind actin, ATP, DNA, fibronectin, and RNA) in PABPs, we used domain information for *S. cerevisiae* in the Integr8 web portal (www.ebi.ac.uk/integr8/FtpSearch.do?orgProteomeld=40). A list of yeast domains that bind polyanions such as ATP (GO: 0005524), actin (GO: 0003779), DNA (GO: 0003677), and RNA (GO: 0003723) was obtained from the InterPro database (www.ebi.ac.uk/interpro/) (fibronectin did not result in any hits). The rationale for this investigation was to relate any identified sequence signature in PABPs to known domain data.

Investigation of the Three-dimensional Structures of PABPs Using Macroscopic Electrostatic Models and Protein Continuum Electrostatics

Our list of 529 unique PABPs was compared against a list of 5,140 yeast proteins in www.expasy.org/cgi-bin/lists?yeast.txt. We utilized protein continuum electrostatics (PCE), developed by Miteva *et al.* (31), which is a web tool based on macroscopic electrostatics with atomic details to calculate the electrostatic properties of PABPs. Submitting a protein crystal structure to the PCE server enables calculation of surface potentials and generation of electrostatic energies via infinite difference solutions to the Poisson-Boltzmann equation (31). The crystal structures of the matched proteins were submitted to the PCE server (bioserv.rpbs.jussieu.fr/PCE) with the default values of 4.0 and 80.0 for protein internal and solvent dielectric constant, respectively. A value of 0.15 was used for the ionic strength, and a range of −3.0 (red) to +3.0 kcal/mol (blue) was chosen for the electrostatic potential distribution surfaces of the matched proteins. Prior to the submission of structures to the PCE server, the crystal structures were examined to ensure that they contained only the most native-like conformation with the longest amino acid chains and no ligands and/or mutations. Additionally a choice of the highest possible resolution and *R*-factors were preferred when selecting the solved PABPs structures. A set of random non-PABPs was also selected and

subjected to the same treatment to compare their electrostatic properties with those of PABPs.

Relationship between Phosphorylation and PABPs

The ORF list of kinase substrates in yeast was obtained from the supplemental information published by Ptacek *et al.* (32) and was compared with the lists of PABPs as well as the proteins that did not interact with our probes (non-PABPs) using MATLAB (version 7.0.4). Statistical analysis of the results was conducted using Fisher's exact test.⁴

PABPs and Intrinsic Protein Disorder

A sequence analysis of PABPs for the extent of their disorder was conducted utilizing the pipeline interface of DisEMBL (33). Based on artificial neural networks with training on three different datasets, this program predicts disorder based on various definitions. In this study, MATLAB was used to calculate the percentage of "hot loops" for PABPs (for three statistical levels), hubs (we have defined hubs as proteins with more than 15 interacting partners), and non-PABPs from DisEMBL output files, respectively. An unpaired, two-sampled unequal variance *t* test was used to investigate the statistical difference between the means of percentages of hot loops for PABPs and non-PABPs.

Essential Genes, Protein Homologues, and PABPs

A database was created to include the name of all essential genes in yeast (34). These data have demonstrated that ~18% of the genes in yeast are essential. We created a query between the ORFs of essential genes and PABPs to investigate how many of the PABPs were expressed by essential genes. Statistical analysis of the results was conducted using Fisher's exact test.⁴

For a protein homologue study, BLAST 2.2.13 was downloaded from the National Center for Biotechnology Information (NCBI) server and installed in a PC work station. All-against-all BLAST searches were conducted on the yeast proteome to determine the number of homologues of yeast PABPs *versus* the whole yeast proteome. The default Blossum 62 scoring matrix was used in the BLAST searches, and the *E*-value cutoff was set to 0.01 to ensure only statistically significant homologues were identified.

Visualization of Potential PABP Networks

Cytoscape (version 2.2), an open-source program, provided a useful tool with which to visualize PABP interaction networks (35). In an attempt to build a cellular network between our PABPs and other proteins from different databases, we used *Saccharomyces* high quality data (YeastHighQuality.sif) as the sample data in Cytoscape, which is based on the known interactions in the Biomolecular Interaction Network Database (36). This dataset contained 3,025 proteins (nodes) and 6,886 protein-protein interactions (edges). We also generated another database from the published interactions of von Mering *et al.* (37) and considered only the high and medium confidence protein-protein interactions (2,617 proteins and 11,855 interactions). We specified two aims in this section: 1) to observe the extent of interaction between PABPs serving as hubs and other yeast proteins and 2) to investigate the presence of any PABP network(s) that formed a coherent, functionally rich network (see "Searching for Enriched Functional Proteins in PABPs Utilizing Biological Network Gene Ontology tool (BiNGO)").

Prior to accomplishing the first goal, a query was created to obtain

the number of interacting partners between each ORF of PABPs in the von Mering *et al.* (37) dataset. Due to a lack of hub definition, we defined high and low connection hubs as proteins with more than 40 and 15 interacting partners, respectively, in this portion of our study.

Searching for Enriched Functional Proteins in PABPs Utilizing Biological Network Gene Ontology Tool (BiNGO)

BiNGO, an implemented plug-in for Cytoscape, was used to determine the statistical overrepresentation of gene ontology (GO) categories in the yeast PABP network (38). This tool permits mapping of the statistically predominant functional categories of genes or proteins based on the GO hierarchy (38). BiNGO uses a hypergeometric test (as opposed to a binomial one) in which sampling occurs without replacement to search for the statistically predominant functional categories in terms of *p* values and GO diagrams. Such diagrams contain color-coded nodes of different sizes based on *p* values that create an easy-to-follow visualization. This program was used to identify a significant functional category for the yeast PABPs as well as their first neighbors. A similar sized set of random proteins was also generated and tested using BiNGO to ensure the significance of the results generated from the experimental set.

Investigating PABP Networks

The yeast interactome was compiled by combining the yeast protein-protein interaction networks from the databases BioGrid (www.thebiogrid.org), Comprehensive Yeast Genome Database (mips.gsf.de/proj/yeast/CYGD/interaction), and IntAct (www.ebi.ac.uk/intact/index.jsp). Only the experimentally detected binary physical interactions were selected from each database. Regarding a specific topological property of subnetworks formed by PABP datasets, we investigated the number of protein-protein interactions involving at least one PABP (39).

We constructed 50,000 random subnetworks by arbitrarily picking the same number of proteins as in the respective PABP dataset and including their interaction partners. We used the resulting histogram of the number of interactions contained in each random subnetwork to estimate a *p* value quantifying the significance of the subnetwork for the respective PABP dataset. In addition, a 95% binomial confidence interval for the *p* value was calculated. To give a concrete example of our evaluation procedure, there were 108 PABPs in the dataset detected at the statistical level of +3S.D., and 103 of them were present in the compiled yeast interactome. Thus, we generated 50,000 different random subnetworks of 103 proteins and their direct interaction partners. We then counted the protein-protein interactions in each subnetwork to drive a histogram.

RESULTS AND DISCUSSION

Biotinylation of Polyanions—Successful biotinylation of polyanion probes was confirmed through Western and dot blots. Different panels in Supplemental Fig. 1 (SF 1) demonstrate the Western blots of biotinylated actin and tubulin (at a molar ratio of 9:1) as well as the biotinylated standards used in these experiments. Dot blots for heparin, HS, and DNA also verified their positive biotinylation.

Quantification of Heparin and HS—Azure A, in the absence of heparin or HS, exhibited a peak maximum at ~630 nm (SF 2). Detailed results of quantification of heparin and HS are reported in the supplemental material.

Protein Array Experiments—To identify PABPs in yeast, arrays containing 4,087 yeast proteins were used. Examples

⁴ Ø. Langsrud, personal website (Fisher's exact test).

TABLE I
Yeast proteins that interact with the five polyanion probes

The information provided in the table is extracted from the SGD at db.yeastgenome.org/cgi-bin/GO/goEvidence.pl (see also ST 1). IEP, inferred from expression pattern; ND, no biological data available; TAS, traceable author statement; IDA, inferred from direct assay; IGI, inferred from genetic interaction; IMP, inferred from mutant phenotype; IPI, inferred from physical interaction; ISS, inferred from sequence or structural similarity.

ORF	Name	Location	Biological process	Molecular function
YBR031W	Ribosomal protein L4A (L2A) (rp2) (YL2)	Cytosolic large ribosomal subunit (sensu Eukaryotae) (TAS)	Protein biosynthesis (TAS)	Structural constituent of ribosome (TAS)
YBR118W	Translational elongation factor EF-1 α	Eukaryotic translation elongation factor 1 complex (63), ribosome (TAS)	Translational elongation (TAS)	Translation elongation factor activity (TAS)
YCR016W	Hypothetical protein	Nucleolus (IDA), nucleus (IDA)	Unknown	Unknown
YDL051W	La homologous protein	Nucleolus (IDA), nucleoplasm (IDA), nucleus (IDA)	tRNA processing (IMP)	RNA binding (IDA)
YDL078C	Malate dehydrogenase	Peroxisomal matrix (TAS), peroxisome (IDA)	NADH regeneration (TAS), fatty acid β -oxidation (TAS), . . .	L-Malate dehydrogenase activity (IMP, TAS)
YGR026W	Unknown	Unknown	Unknown	Unknown
YHL033C	Ribosomal protein L8A (rp6) (YL5) (L4A)	Cytosolic large ribosomal subunit (sensu Eukaryotae) (TAS)	Protein biosynthesis (TAS)	Structural constituent of ribosome (TAS)
YHR117W	Protein translocase 71-kDa component of the outer membrane of mitochondria	Mitochondrial outer membrane (IDA)	Unknown	Protein transporter activity (IPI, ISS)
YKR009C	Multifunctional β -oxidation protein in fatty acid oxidation	Peroxisomal matrix (TAS)	Fatty acid β -oxidation (TAS)	3-Hydroxyacyl-CoA dehydrogenase activity (TAS), enoyl-CoA hydratase activity (TAS)
YLR003C	Hypothetical protein	Nucleus (IDA)	Regulation of DNA replication (IGI)	Unknown
YLR150W	Purine motif triplex-binding protein	Cytoplasm (IDA), cytosolic ribosome (sensu Eukaryotae) (IDA), nuclear telomere cap complex (IGI)	Antiapoptosis (IDA), telomere maintenance (IMP)	DNA binding (IDA), telomeric DNA binding (TAS)
YMR153W	Karyopherin docking complex component of the nuclear pore complex nuclear pore complex subunit	Nuclear pore	mRNA binding, protein export from nucleus, . . .	Structural molecule activity
YOR056C	Essential nuclear protein involved in proteasome maturation and synthesis of 40 S ribosomal subunits; required for cleavage of the 20 S pre-rRNA to generate the mature 18 S rRNA	Nucleus (IDA), preribosome (IDA, IEP, IPI)	Processing of 20 S pre-rRNA, proteasome assembly, ribosomal small subunit biogenesis	RNA (IDA, ISS) and protein binding (IDA, IMP)
YOR309C	Hypothetical protein	Cellular component unknown (ND)	Unknown	Molecular function unknown (ND)
YPR086W	Transcription factor TFIIIB homolog	Nucleoplasm	Transcription initiation from RNA polymerase II promoter	General RNA polymerase II transcription factor activity
YPR139C	Vacuolar protein sorting	Cytoplasm	Protein-vacuolar targeting (IMP)	Unknown

of images created from different protein arrays are shown in the supplemental material (SF 3). The 16 yeast proteins that interacted with all five polyanion probes at the physiological NaCl concentration of 0.15 M and +1S.D. are shown in Table I (more information is in Supplemental Table 1 (ST 1)). Table II

includes the total number of PABPs that were bound to polyanions at the statistical level of +1S.D. (see "Protein Array Experiments" under "Experimental Procedures"). Using this approach, we identified 135, 60, 264, 348, and 86 yeast proteins interacting with actin, tubulin, heparin, HS, and DNA,

TABLE II
Yeast polyanion-binding proteins

Shown is the number of proteins that significantly interacted with model polyanions using yeast protein arrays at 0.15 M NaCl solution and +1S.D. level.

Polyanion	Number of PABPs on yeast array
Actin	135
Tubulin	60
Heparin	264
HS	348
DNA	86

TABLE III
The effect of NaCl on heparin-binding proteins

Shown is the number of heparin-binding yeast proteins at three different NaCl concentrations and accompanying statistical significance. Heparin concentration was 50 μ g/ml.

	NaCl (M)			
	0	0.05	0.15	0.75
Total proteins with positive interaction	4247	4282	4276	4278
Total proteins with significant interaction				
1S.D.	254	261	264	258
2S.D.	107	105	107	104
3S.D.	60	58	59	60

respectively. Furthermore significant overlaps were observed in which some PABPs interacted with more than one polyanion. For example, 216, 88, 46, and 16 proteins interacted with two, three, or four or more probes, respectively. A comprehensive list of the standard ORFs of proteins that interacted with the individual polyanion probes is available in the supplemental tables (ST 2).

To compare the extent and number of proteins that interacted with each polyanionic probe, NaCl incubation experiments were conducted. The number of proteins that interacted with heparin when each array was probed with this polyanion at zero and incrementally increasing concentrations of NaCl is shown in Table III. Similar tables for the other polyanion probes are presented in the supplemental information (ST 3–6). For example, at 0.15 M NaCl, 59 heparin-binding proteins interacted with this probe with high confidence level (+3S.D.) compared with 264 proteins at the level of +1S.D. Subsequently a comparable number of proteins appeared to interact with heparin when this array was exposed to 0 and 0.75 M NaCl. These results were not anticipated because it was thought that, with increasing concentrations of NaCl, a smaller number of heparin-binding proteins should have been observed. Similar results were obtained for experiments with actin, tubulin, HS, and DNA probes (ST 3–6), although a small reduction in binding was seen at high salt concentrations in the case of tubulin, HS, and DNA.

Competition Experiments with DS—Because the incubation of protein arrays with increasing concentrations of NaCl did not demonstrate a major reduction in the number of PABPs, another experiment was conducted using DS. This observed

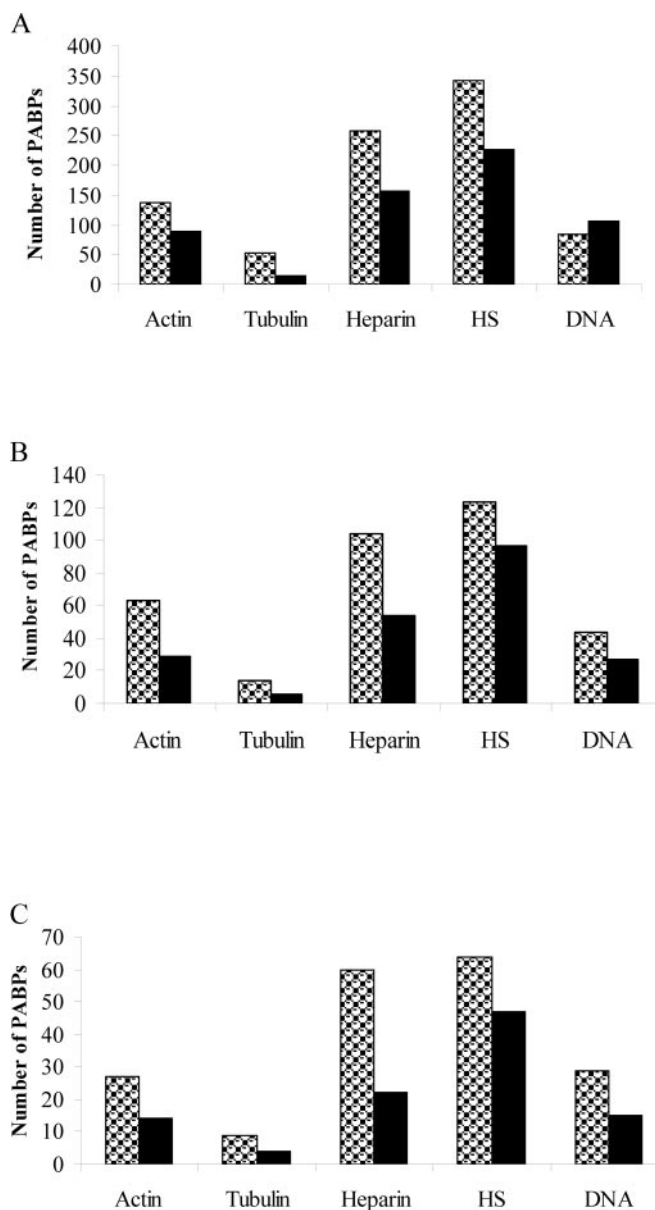


FIG. 1. **Effect of DS on polyanion-binding proteins.** The reduced number of polyanion-binding proteins after a 24-h incubation of the yeast protein arrays with DS is shown. Patterned and solid bars represent the number of PABPs before and after treatment of the array with DS, respectively, with the statistical level of +1S.D. (A), +2S.D. (B), and +3S.D. (C).

lack of reduction in the number of PABPs could have been due to inadequate incubation times with NaCl or the lack of polyanionic character. Thus, we speculated that the higher density of negative charges on DS compared with heparin or HS might permit it to displace the polyanions already bound to the proteins on the array. In fact, the results of DS displacement experiments at three different levels of statistical significance demonstrated an ~50% decrease in the number of yeast proteins that were bound to actin, tubulin, heparin, HS, and DNA (Fig. 1). It was unexpected, however, that the num-

TABLE IV
PABPs are enriched in basic amino acids

Shown are the average percentage of Lys, Arg, and His from the identified PABPs (test set) and the overall population of proteins from the total yeast proteins on the array.

Polyanion	Mean percentage of				Mean pI	Mean no. positive patches
	Lys (experimental)	Arg (experimental)	His (experimental)	All positive residues		
Actin	8.74	5.18	2.14	16.06	7.86	4.76
Tubulin	9.36	5.29	1.85	16.50	8.19	4.61
Heparin	8.54	4.73	2.13	15.40	7.96	4.69
HS	8.57	4.72	2.05	15.34	7.74	4.73
DNA	10.10	5.47	1.89	17.46	8.06	6.43
Average				16.20		5.04
Population	7.60	4.70	2.18	14.48	7.40	3.33
Difference				1.72 (11.9%)		1.71 (51.4%)
p value	0.0081 ^a	0.067	0.050 ^a	0.0137 ^a	0.002 ^a	0.0079 ^a

^a Indicates a significant difference ($p < 0.05$).

ber of DNA-binding proteins increased after incubation with DS (Fig. 1A). This apparent contradiction is probably due to the large standard error of this particular experiment because in the subsequent panels (B and C) this result was not observed. Although the nature of polyanion-protein binding is generally considered to be primarily electrostatic (40, 41), these experiments support the idea that polyanion-protein interactions also involve other non-coulombic types of intermolecular (e.g. hydrophobic and hydrogen binding) interactions (40, 41).

Searching for Positively Charged Amino Acids in the Sequence of PABPs—It is already well established that proteins that interact with polyanions contain localized regions with high positive charge density (12, 40–42). To confirm that PABPs are enriched for positively charged amino acids relative to other proteins, the percentages of Lys, Arg, and His for the test and population sets were calculated. The results are shown in Table IV. These values were generated by calculating the percentage of the basic amino acids for each ORF of all yeast proteins that interacted with polyanions (PABPs) and the population of proteins present on the arrays, respectively, and subsequently averaging these values over the entire range of test and total proteins. The results from a one-sample *t* test demonstrated a statistically higher percentage of Lys, a marginal increase for His, and a lower amount of Arg in the PABPs (Table IV). Overall there is an 11.9% increase of positively charged residues compared with the general population set (this percent difference was generated using a similarly sized 100 random sets, and the average result was 0.034%). The enrichment of positive residues seems to be more pronounced locally (*i.e.* within small sequence runs) because there is an ~51% increase in the number of positive patches in PABPs (Table IV). These positively charged regions are probably on the surfaces of proteins because it is well established that charged residues are surface-localized (see below). Our data in Table IV are consistent with the presence of localized positive patches on PABPs. It is quite likely that other factors such as the specific three-dimensional geometry

of a positively charged site strongly influence the charge-charge interactions between PABPs and polyanions. For example, it has been argued that Arg is more effective in binding to polyanions than Lys (40) due to an optimal geometry. In our analysis, however, PABPs generally contained more Lys and His than Arg.

pI Calculations—If PABPs are presumably enriched in basic amino acids, their pI should demonstrate this fact. Thus, we investigated the pI values for PABPs as well as the for all proteins on the array. The results, averaged over the total number of each set, demonstrated a mean value that was slightly but significantly ($p < 0.05$) higher than that of all yeast proteins (Table IV). Although it was observed previously that the values of pI did not correlate with the extent of polyanion binding (12), the inclusion of the entire yeast proteome in this analysis demonstrated a slightly more basic nature for the PABPs. This is also consistent with the high percentages of Lys in these proteins.

Searching for a Sequence Signature in PABPs Using MEME and InterPro—Because PABPs were enriched in basic amino acids, a sequence signature study was performed to determine whether they shared any common motifs. A MEME analysis (29) found 23 signatures within the population of PABPs that matched the existing entries in the InterPro database. Most of these are domains present in RNA/DNA-binding proteins, kinases, and chaperone domains. These categories of proteins were also observed in Cytoscape and BiNGO analysis of the PABP clusters along with other functional categories (see below). Although the existence of possible sequence signatures for heparin-binding proteins has been hypothesized previously (43), MEME did not identify any common unique signatures for the more general class of polyanion-binding proteins. Cardin and Weintraub (43) proposed sequences of *XBBBXXBX* and *XB-BXB* (with *B* and *X* as basic and hydrophobic residues, respectively) as consensus sequences from studying 12 known heparin binding regions in four proteins. Hileman *et al.* (40) also suggested a *TXXBXXTBXXXTBB* motif (where *T*

TABLE V
PABPs and the known polyanion binding domains

The result of MEME and domain searches demonstrate some of the PABPs with known ATP, DNA, and RNA binding domains.

	DNA binding (423)		RNA binding (117)		ATP binding (369)		Actin binding (19)	
	PABP	All yeast	PABP	All yeast	PABP	All yeast	PABP	All yeast
Percentage of proteins having the domains	3.5	3.0	1.7	1.0	8.5	10.5	0	2.7

defines a turn and *B* and *X* have the previous definitions) using x-ray and NMR structural data for heparin-binding growth factors.

As stated above, although MEME did not identify any novel sequence signatures, it identified 23 signatures that have been associated with binding sites for polyanions such as ATP, DNA, and RNA. Table V reports the fraction of PABPs that share known polyanion binding domains. It can be seen that 3.5, 1.7, and 8.5% of PABPs share a common DNA, RNA, or ATP binding domain, respectively. It was expected that we might also detect a common actin binding domain. We suspect that the lack of such a finding is due to the fact that the presence of known domains is often a function of particular polyanionic sequences as well as the experimental conditions used in those investigations. In the current study, we focused on the presence of more nonspecific interactions, which may explain some of the discrepancy. The fact that MEME analysis found no novel polyanion binding sequence signatures enforces the idea of nonspecific interactions between positive patches on PABPs surfaces and polyanions rather than highly stereoscopically specific binding sites. Another example of such nonspecific interactions was evident when we compared the list of our DNA-binding proteins to that generated by Hall *et al.* (44). We found that most of the DNA-binding proteins that were identified by Hall *et al.* (44) bound to a least one of our other probes, a clear indication of the nonspecific nature of such interactions.

Investigation of the Three-dimensional Structures of PABPs Using Macroscopic Electrostatic Models and Protein Continuum Electrostatics—Based on the above analysis, it is clear that PABPs contain localized positively charged regions of basic amino acids. To visually inspect this feature, a search for matching yeast PABPs to the ExPASy list was conducted. This resulted in 24 proteins with solved three-dimensional structures. PCE was used to depict the surface potential of these proteins. The crystal structures of the yeast PABPs, depicted in Fig. 2, consistently demonstrated the presence of one or more positively charged patches (colored in *blue*). Although the crystal structure of some proteins (such as Protein Data Bank (PDB) code 1HOW) contain less extended positively charged areas, all of them were experimentally found to bind to one or more of the polyanion probes. Details of structure selection as well as the electrostatic analysis of random protein structures can be found in the supplemental material (ST 7). We also applied this method to the available crystal structures of human polyanion-binding proteins, and

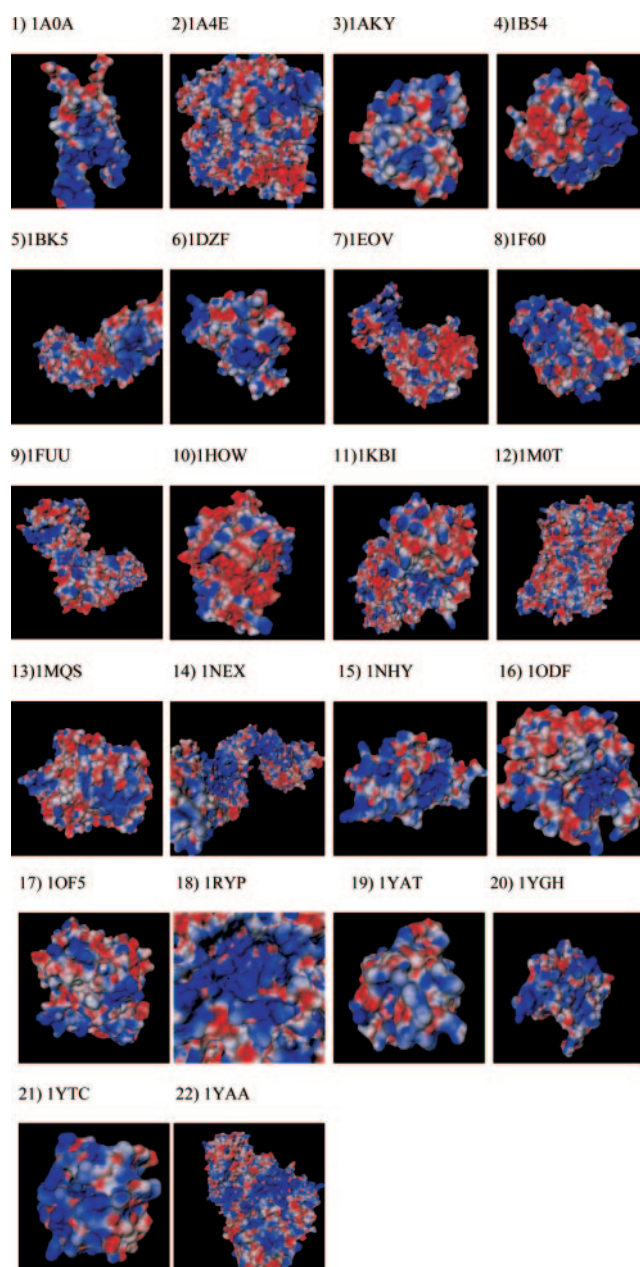


FIG. 2. Electrostatic potential surfaces of 22 PABPs from -3.0 kcal/mol (red) to $+3.0$ kcal/mol (blue). The surface potential calculations were not successful for one of 24 structures (PDB code 1YGP) due to problems with the backbone structure deposited in the PDB. PDB code 1RYP defines two structures for YGR135W (proteasome component Y13) and YOR362C (proteasome component C1). For more structural detail, see ST 7 in the supplemental material.

distinct positively charged surfaces were also observed in each case.⁵

Relationship between Phosphorylation and PABPs—Phosphorylation is a major cellular regulatory mechanism and modulates the polyanionicity of target proteins. We conducted an analysis to determine whether any PABPs were kinase substrates and whether phosphorylation could significantly alter the polyanion binding character of a PABP. Thus, we suggested that when PABPs are stabilized by polyanionic surfaces phosphorylation could alter their polycationic nature and potentially create a regulatory mechanism for alteration of their interaction with the polyanionic surfaces. Ptacek *et al.* (32) have identified 1,325 kinase substrates that belong to different functional categories in yeast. We compared their list to the list of our experimentally determined PABPs as well as proteins that did not interact with polyanionic probes. Because phosphorylation (or dephosphorylation) will have a major effect on the polyanionic nature of proteins at specific sites, we hypothesized that PABPs might be more susceptible to this particular chemical modulation. Thus, the regulatory effect of phosphorylation might modulate the charge within the polyanion binding sites of PABPs and thus alter their interaction with any possible intracellular polyanionic network (see below). This has been shown, for example, with the multisite phosphorylation of stathmin (a microtubule-binding protein), which regulates microtubule dynamics (45).

It was found that of the 3,536 unique yeast proteins that did not interact with polyanions, 1,134 were kinase substrates (32%). In contrast, of 529 proteins that did interact with polyanionic probes, 311 were kinase substrates (59%). This demonstrates a statistically significant (Fisher's exact test, $p < 0.0001$) difference between the number of kinase substrates that possess polyanion binding ability and those that do not, suggesting that PABPs are stronger candidates for phosphorylation than non-PABPs. Because protein phosphorylation is estimated to affect 30% of the proteome as a major regulatory mechanism (32), this finding may be of functional significance. This also suggests the possibility that the regulation of PABPs by phosphorylation may involve their nonspecific interaction with polyanions. For example, phosphorylation, by decreasing the positive charge in a specific region of a protein, could decrease the interactions of the protein with polyanions. An example of such an event is seen in DNA non-homologous end joining. In this case, the phosphorylation of linker histones (a PABP) by DNA-dependent protein kinase in the vicinity of a DNA break reduces the histone affinity for DNA ends and facilitates DNA ligation to perform end joining (46).

PABPs and Intrinsic Protein Disorder—The view that proteins require a well defined conformation to function has increasingly been challenged (47, 48). It is now often argued that many proteins contain a disordered region that permits

increased structural flexibility and, therefore, the option of favorable interactions with a variety of ligands or protein partners (48). Such proteins have also been proposed to be associated with "hub" activity in which they possess a central interaction role in protein-protein networks and presumably require structural flexibility for interaction with multiple partners (49). Interestingly the disordered regions of proteins appear to also be subject to increased phosphorylation (50). Therefore, we hypothesized in this work that polyanions may serve as potential interaction partners for "disordered" PABPs. We focused on the hot loop type of disorder (defined as coils/loops with a high degree of mobility as determined from crystallographic B-factors) because this identifier has demonstrated a more accurate performance evaluation (33). We also conducted a similar analysis for PABPs that interacted with more than 15 partners (hubs) to investigate whether hub proteins were disordered in nature (49) because it is known that a node with a large number of connections should serve as a hub in a protein network (51). As mentioned previously, there is no formal definition of how many protein neighbors are required for a protein to be qualified as a hub. Thus, PABPs with more than 15 interaction partners were arbitrarily chosen.

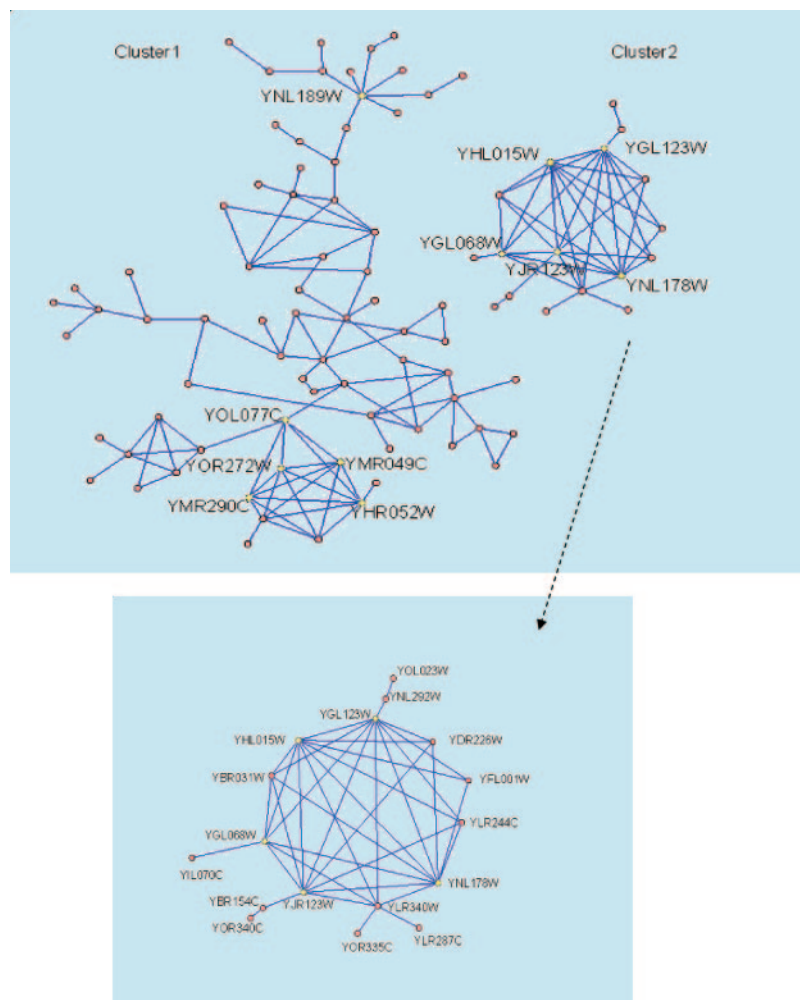
The calculations for hot loops found that 519 of 529 PABPs displayed an average value of 32% disorder (at +1S.D. level). A value of 29% was found for non-PABPs. Although the difference does not appear large, it was significant with a p value of 6×10^{-5} . These values at +2S.D. and 3S.D. levels were also very significant ($p < 0.001$). This percentage also was close to statistical significance in the case of PABPs with 15 or more partners ($p = 0.067$).

Considering the increase in the disordered nature of PABPs, we suggest that it is possible that polyanions may serve as interaction partners for these proteins to structure their more flexible, disordered regions *in vivo*. In addition, this suggests the phrase "natively disordered" may be inappropriate at least *in vivo*. Given the ubiquity of intracellular polyanions, it may well be that proteins that are PABPs fold immediately upon synthesis due to the local presence of polyanions (such as ribosomes) (52). This is consistent with recent NMR studies that have shown little evidence for unfolded proteins in cells (53). Interestingly, it has also been observed that the negatively charged cavity environment of GroEL is essential for rapid folding of some proteins (54).

Essential Genes, Protein Homologues, and PABPs—We also investigated whether PABPs were expressed by essential genes in yeast cells because some proteins expressed by these genes serve as hubs (51). Additionally we investigated the presence of protein homologues for PABPs because these proteins often possess functional importance. Such functional features are involved in the robustness of the system for tolerating errors (55). The result of a query between the yeast ORFs of essential gene products and PABPs demonstrated that 26 PABPs of a total 135 proteins that interact with

⁵ N. Salamat-Miller, J. Fang, C. W. Seidel, Y. Assenov, M. Albrecht, and C. R. Middaugh, unpublished results.

FIG. 3. **Yeast PABP networks generated using Cytoscape.** Two clusters are shown, and the yellow nodes represent the 11 PABPs that serve as hubs. Only the names of hubs are illustrated for clarity.



actin were classified as essential. In the case of tubulin, heparin, HS, and DNA, 11, 48, 78, and 16 PABPs of 60, 264, 348, and 86 total PABPs are classified as essential, respectively. Although Fisher's exact test found no enrichment in the number of essential genes in PABPs ($p > 0.05$), the two-tailed p value ($p = 0.0426$) for protein homologues demonstrated that the number of homologues of PABPs is greater than the number of homologues found in the general population of yeast proteins. The average number of homologues for the population set was 4.36, whereas this number was 4.31, 5.53, 5.61, 5.35, and 6.79 for PABPs that bind actin, tubulin, heparin, HS, and DNA, respectively. This indicates that in the case of DNA-binding proteins, on average, there are almost seven proteins that have significant sequence similarity with each DNA-binding protein. Because the sequence determines the function of proteins, it may be that such homologues perform a similar function in yeast cells.

Visualization of PABP Networks—One of the more potentially intriguing aspects of this study involves a description of the yeast proteome in terms of protein-polyanion interactions. We thus hypothesized that a network of PABPs may be pres-

ent as a direct consequence of PABP-polyanion interactions. To this end, we were able to find 260 of 529 unique PABPs in the YeastHighQuality.sif network file of Cytoscape. These 260 PABPs interacted with each other through 82 protein-protein interactions. A map of these selected PABPs displayed eight distinct clusters with 83 nodes and 82 edges (not shown). The rest of the PABPs (177) demonstrated a non-interacting cluster. We also identified 266 nodes with 168 interactions for PABPs in the dataset generated by von Mering *et al.* (37). It appears that this dataset represents a more complete protein-protein interaction map because we detected only two major clusters consisting of 90 proteins with a more complete 145 interactions identified (compared with 82 edges, Fig. 3). The remaining 176 proteins did not participate in the clusters and manifested only 23 interactions. The lack of even more interactions between these proteins probably reflects the incompleteness of the interaction databases that are currently available. As previously discussed (see "Experimental Procedures"), we propose two aims in analyzing PABPs networks. The first aim was the detection of hub PABPs and their extent of interaction with other yeast proteins. For this pur-

pose, we identified PABPs that appear to act as hubs and investigated the number of interacting partners as well as their function. The second aim was to investigate whether polyanionic surfaces induce PABPs to form functional clusters in the form of networks. Thus, we used BiNGO analysis to investigate the functionality of PABPs (see below).

We identified 11 and 51 PABPs as hubs when we defined all PABPs as requiring more than 40 and 15 interactions partners, respectively (see “Experimental Procedures”). The 11 PABPs were selected in the YeastHighQuality.sif of Cytoscape sample data and demonstrated 10 interactions primarily among five of these proteins (YMR290C, YMR049C, YOR272W, YOL077C, and YHR052W). When the von Mering *et al.* (39) dataset was used, the 11 hubs demonstrated 20 interactions in the form of two major pentagonal clusters (SF 4A).

To take this analysis one step further, we used Cytoscape to identify the first neighbors of the 11 hub PABPs. This included some PABPs and also many non-PABPs. Using the YeastHighQuality.sif and von Mering *et al.* (37) datasets, 124 proteins with 811 interactions and 287 proteins with 4,204 interactions were identified, respectively. This high degree of interaction between PABPs and the yeast proteome demonstrates the extent to which PABPs may be involved in a yeast cell. The BiNGO analysis of many of these proteins from both datasets showed that they not only perform the expected molecular functions such as RNA binding but also possess helicase activity and are structural constituents of ribosomes (see below). They are involved in biological processes as diverse as general metabolism, rRNA metabolism, and ribosome biogenesis/assembly. They appear to reside at several different cellular locations such as membrane-bound and unbound organelles, the nucleus, ribosomal subunits, and various protein complexes. These results are consistent with the results of the MEME analysis. Similar analyses can be found for hub PABPs with more than 15 interaction partners in the supplemental material (SF 4B.a and 4B.b). These exhibit additional functional roles such as transferase and polymerase activity.

Fig. 3 shows a PABP network (266 proteins with 168 interactions) generated from the von Mering *et al.* (37) dataset using Cytoscape. This consists of two distinct clusters. Each contains a specific number of nodes and edges. For example, cluster 1, which was the larger of the two networks, contains 73 proteins with 107 interactions. Both clusters were analyzed using BiNGO (see below) to search for any functional overrepresentation present in either one.

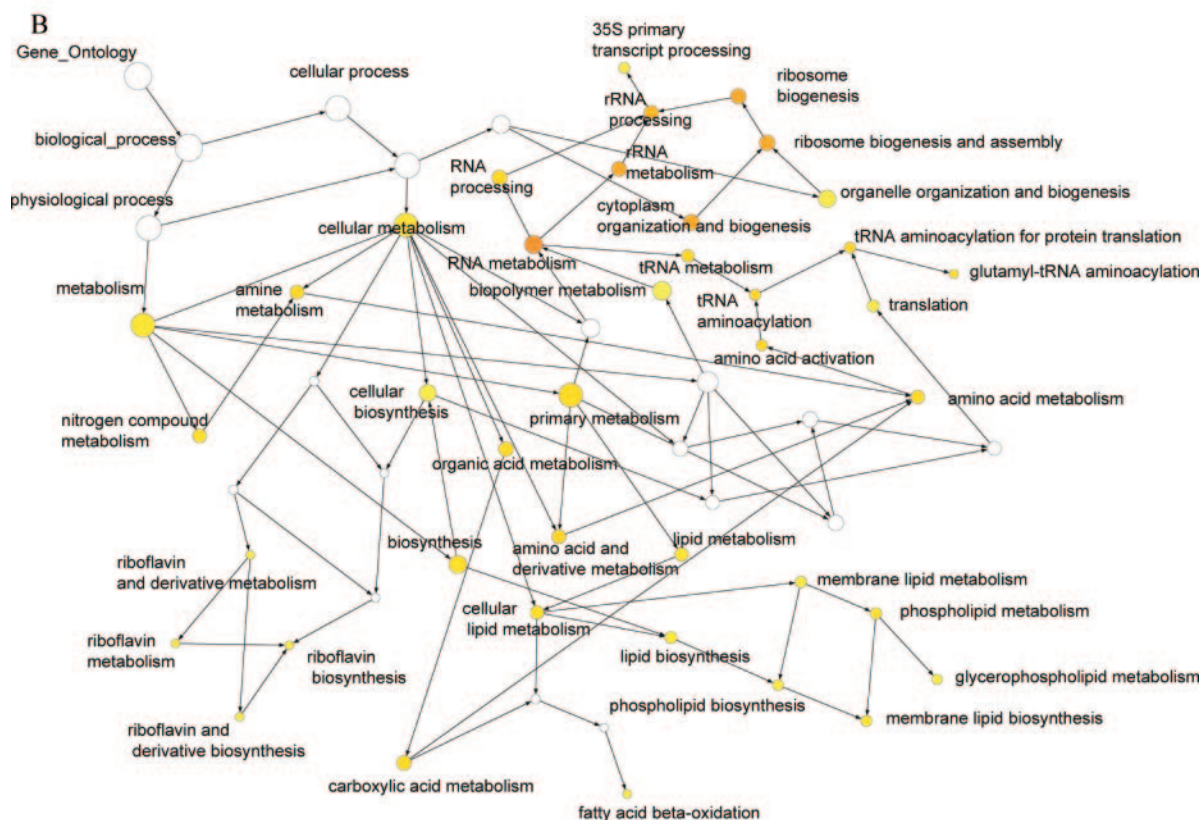
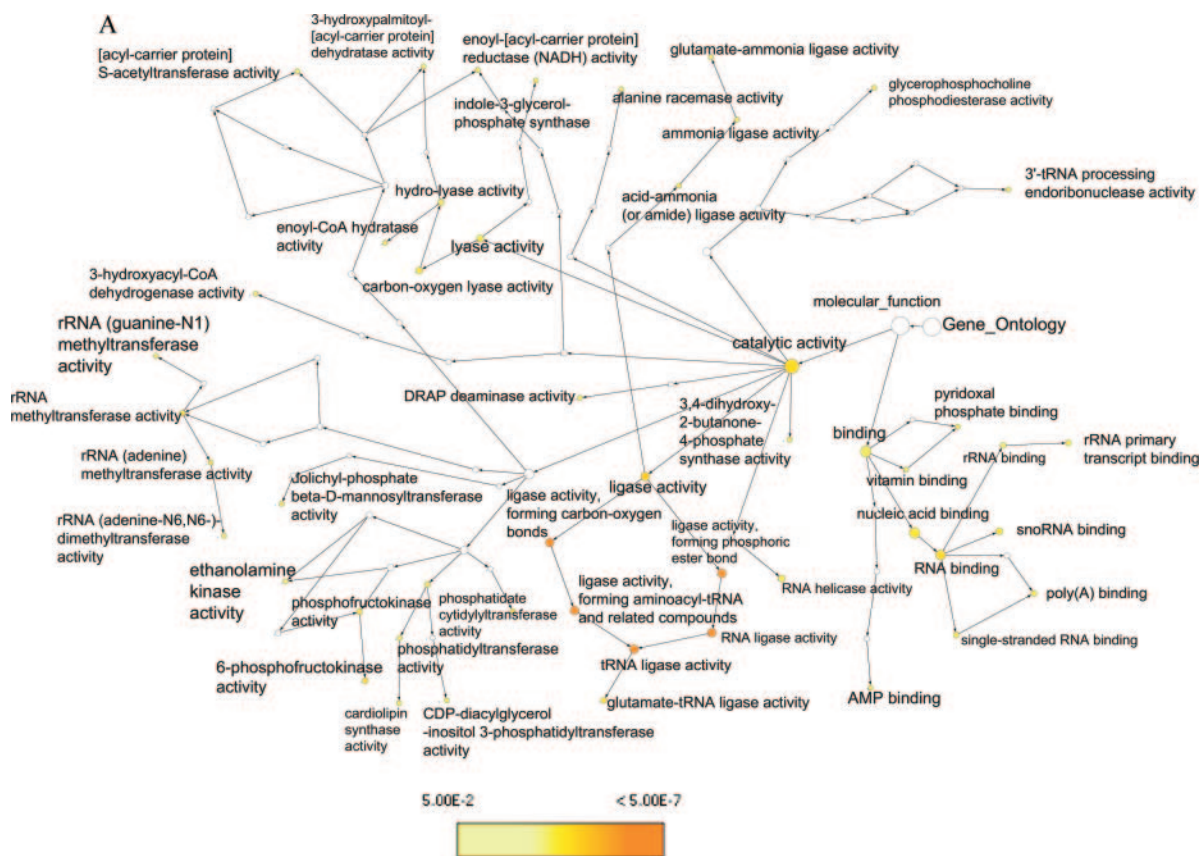
Searching for an Enriched Functional PABPs Using BiNGO—Overrepresentation of the GO categories of molecular function, biological processes, and cellular compartments of cluster 1 in yeast PABP networks is shown in Fig. 4, A, B, and C. This diagram demonstrates the statistical significance of any overrepresentation through a color scale ranging from yellow (p value = 0.01) to dark orange (p value =

10^{-7}). Because the color saturates at dark orange for p values that are more than 5 orders of magnitude smaller than the chosen significance level (0.01), the most intensely colored nodes that are farthest down the hierarchy should be considered the more significant in terms of molecular function, biological process, or the cellular location of the subject proteins (38).

If polyanionic interaction with PABPs within cells really facilitates the formation of these proteins into specific network arrangements, it can be hypothesized that they should interact with each other in a functionally important manner and should belong to a particular group(s). We used BiNGO to address this question. The resultant diagrams in Fig. 4 clearly show a functional overrepresentation of PABPs in yeast in contrast to randomly selected yeast proteins that demonstrated little or no functional enrichment. Using p values as criteria, Fig. 4A suggests that many of the yeast PABPs are involved in the expected activity categories such as nucleic acid binding as well as in other classes including ligase activity and lyase reactions. Additionally their molecular functions are overrepresented in macromolecular (RNA and rRNA) metabolism, lipid biosynthesis, ribosome biogenesis, and cellular biosynthesis (Fig. 4B). This analysis also argues that PABPs in cluster 1 are particularly located in the nucleolus, peroxisomal matrix, mitochondrion, 6-phosphofructokinase complex, and methionyl glutamyl-tRNA synthetase complex. The analysis of the second cluster can be found in SF 5.

Investigating PABP Networks—To investigate whether PABPs form a specific interaction network, we compiled a yeast interactome consisting of 36,135 physical interactions between 5,574 proteins. The rationale behind this study was that if polyanions and PABPs have an important functional role in many cellular processes, the PABPs should be involved in significantly more protein-protein interactions than a random set of proteins. We found out that the number of interactions involving PABPs was significantly higher than in random subnetworks as judged by the estimated p values (average of 0.04, see ST 8). The resulting histograms from counting the protein-protein interactions in each subnetwork is shown in SF 6. Interestingly the p values for the PABP dataset at the +3S.D. statistical level are not always higher than the ones for the corresponding +1S.D. and/or 2S.D. datasets (see SF 6).

Summary—This work suggests the existence of a network of polyanions and/or PABPs that may play a role in the structure and function of yeast cells. Thus, polyanionic surfaces could potentially align/regulate PABPs for interaction with other proteins through electrostatic and non-coulombic molecular interactions. This may be thought of as a “beads on string” type of organization, a structure previously observed for protein-polyanion complexes (56). This might also be thought of in terms of the following admittedly crude analogy. If a cell is considered a city, then polyanions behave like roads, buildings, elevators, and other routes of transportation



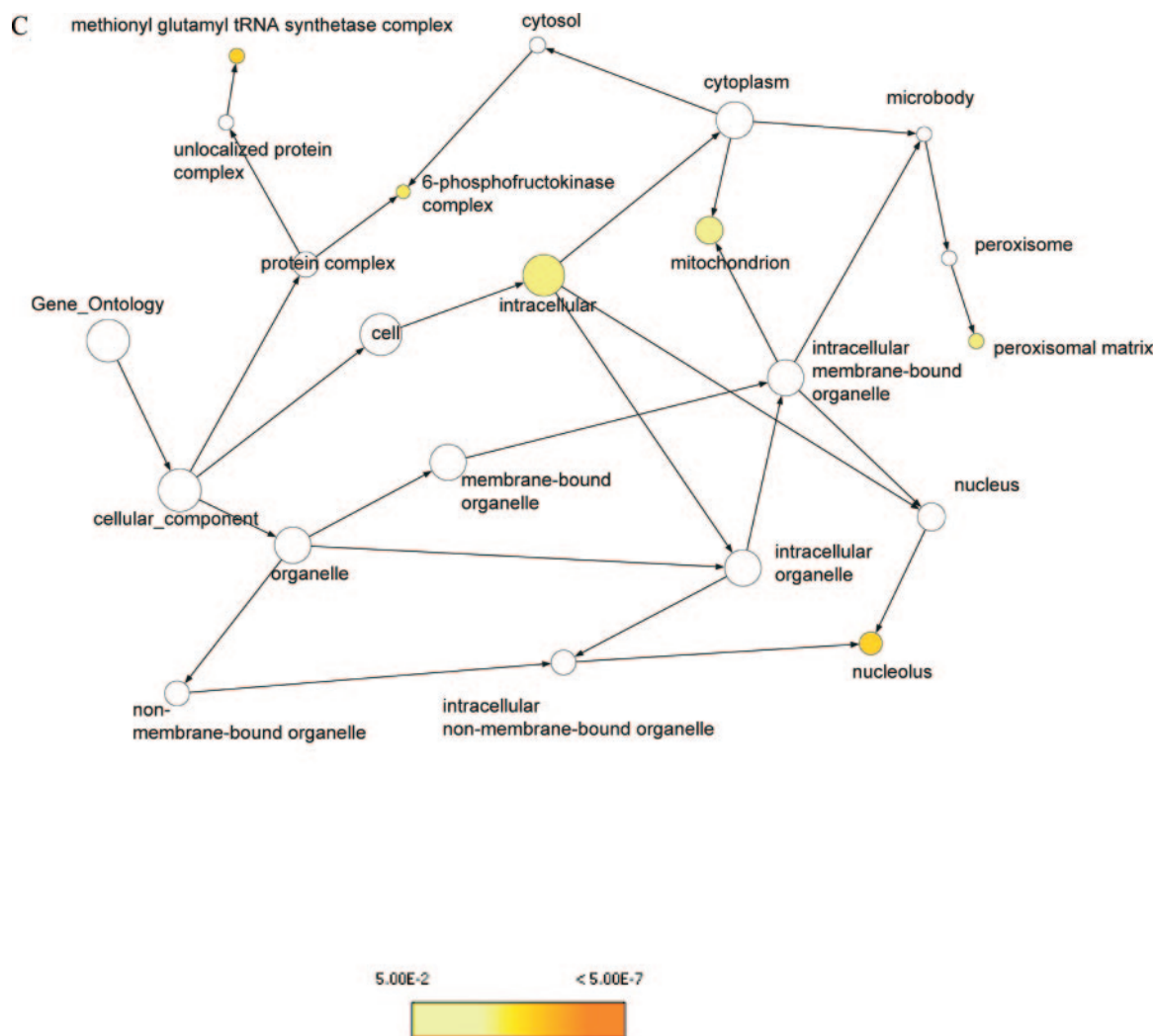


FIG. 4—continued

and sites of habitation, whereas the PABPs correspond to the individuals who interact and perform the various activities of the city. Although simplistic, this analogy captures the possible role of the proposed polyanion network.

From the analyses performed here as well as observations by others (57, 58), it appears that there is a limited stringency in the criteria for polyanion-protein interactions. A minimal requirement of basic amino acids that form localized patches of positive charge on the surface of proteins with no necessity for unique binding motifs appears to be sufficient for such interactions with little discrimination toward the source of negative charge. Several previous studies have been conducted to better define protein and glycosaminoglycan interactions based on molecular modeling and amino acids sequences (40, 43). Most of the latter studies were conducted

with either a limited number of proteins or small synthetic peptides. As we gain further knowledge of the structural basis of polyanion-protein interactions, a lack of specificity for this type of binding (despite its relatively high affinity) is becoming increasingly apparent. The question we must address here is whether such nonspecific interactions play a functional role within cells. This idea is somewhat in contrast to the usual assumption that interactions between macromolecules are directed only by stereoscopically precise binding.

Chu *et al.* (22) have defined “unique” and “common” binding sites for fibroblast growth factor-2 (FGF-2) and heparin-binding epidermal growth factor-like growth factor (HB-EGF) through competition studies. Their analysis found that the number of common binding sites for FGF-2 and HB-EGF were 13- and 30-fold greater than the unique ones, respectively.

Fig. 4. **Diagrams produced by BiNGO.** A, molecular function; B, biological process; and C, cellular compartment. The *inset* demonstrates the intensity of the *p* values generated using a hypergeometric test. *snoRNA*, small nucleolar; *DRAP*, 2,5-diamino-6-ribitylamino-4(3*H*)-pyrimidinone 5'-phosphate.

Although as suggested by Chu *et al.* (22) it is plausible that the common sites facilitate the access of HS to the unique sites, we propose that such common sites may also represent a functional surface for nonspecific binding to polyanionic surfaces in a cell that can regulate and/or stabilize FGF-2 or HB-EGF (and similarly other PABPs) and/or align them for further interactions with other proteins. Such less specific sites on PABPs should be more available for interaction with polyanions from simply a statistical point of view.

We showed that PABPs that are functionally important (through gene duplication and BiNGO analysis) exhibit high affinity toward polyanions not because of novel sequence signatures but rather, at least partially, through nonspecific electrostatic interactions. We also demonstrated a higher probability of regulation of PABPs by phosphorylation, suggesting further functional modulation by nonspecific interactions between PABPs and polyanions. These findings as well as those of Chu *et al.* (22) that PABPs have common sites for nonspecific interactions suggest that such interactions are of functional importance in a cell. A similar view has been proposed recently in which electrolyte pathways and pools “wire” the cytoplasm and organize the crowded environment of the cell (59). In this work, however, we suggest a different “wiring” system in which the cellular polyanions provide a matrix within which some of the functional activity of the cell occurs. The possibility of the simultaneous presence of multiple systems of such organization seems quite plausible.

Although it is apparent that protein arrays are useful tools for high throughput proteomics studies, several limitations of this approach should be noted. Binding to a static or “just-in-time-assembly” complex of proteins (60) and transient protein-protein interactions cannot be easily detected by this stationary approach. Thus, this approach would be expected to significantly underestimate the extent of any interactions probed. Protein arrays, however, are perhaps the most effective tool for rapid proteomics studies where post-translational modifications of proteins are present.

In summary, we identified 529 PABPs in the yeast proteome. It was confirmed that charge-charge interactions as well as other types of non-covalent binding are responsible for these polyanion-protein interactions. Highly localized positively charged patches of basic amino acids result on average for a slightly basic character for the 529 yeast PABPs. No unique sequence signature, however, was identified within and among PABPs, but some confirmation of previously known binding motifs was produced. The most interesting outcomes of this study were the observations that (a) many PABPs were subjected to phosphorylation, implying that regulation of such proteins through their charge-charge interactions with polyanions may be possible; (b) a high content of disorder was present in many PABPs that may be a target for stabilization and structure induction (e.g. folding) by polyanions; (c) there was an extensive interaction between PABPs and other yeast proteins in the form of a

network; and (d) there is a potential role of some PABPs as protein hubs. The functional, biological, and locational overrepresentation of PABPs networks in yeast, which is clearly not the result of chance, suggests a degree of critical interaction between polyanions and proteins of widespread functional significance.

Acknowledgments—We extend a special thank you to Richard Himes, Fernando Ausar, and Sangeeta Joshi at the University of Kansas for helpful discussions. Kim Bure, Clea D’sa, Kevin Clancy, and Gregory Michaud at Invitrogen Inc. helped with the use of the protein arrays. Chris Putnam at the Scripps Research Institute generously supplied the code for pl calculations. Gary Bader, Steven Maere, the members of Cytoscape discussion group, Maria Miteva, and Karen Christie are gratefully acknowledged for help with Cytoscape, BiNGO analysis, PCE calculations, and the suggestion to use the batch SGD databases, respectively.

* This work was supported in part by Kansas IDeA Network of Biomedical Research Excellence Bioinformatics Core, National Institutes of Health Grant P20 RR016475. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ The on-line version of this article (available at <http://www.mcponline.org>) contains supplemental material.

** To whom correspondence should be addressed: Dept. of Pharmaceutical Chemistry, University of Kansas, 2030 Becker Dr., Lawrence, KS 66047. Tel.: 875-864-5813; Fax: 875-864-5814; E-mail: middaugh@ku.edu.

REFERENCES

1. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., and Blundell, T. L. (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature* **407**, 1029–1034
2. Tumova, S., Woods, A., and Couchman, J. R. (2000) Heparan sulfate proteoglycans on the cell surface: versatile coordinators of cellular functions. *Int. J. Biochem. Cell Biol.* **32**, 269–288
3. Dabora, J. M., Sanyal, G., and Middaugh, C. R. (1991) Effect of polyanions on the refolding of human acidic fibroblast growth factor. *J. Biol. Chem.* **266**, 23637–23640
4. Manna, T., Sarkar, T., Poddar, A., Roychowdhury, M., Das, K. P., and Bhattacharyya, B. (2001) Chaperone-like activity of tubulin: binding and reactivation of unfolded substrate enzymes. *J. Biol. Chem.* **276**, 39742–39747
5. Rentzeperis, D., Jonsson, T., and Sauer, R. T. (1999) Acceleration of the refolding of Arc repressor by nucleic acids and other polyanions. *Nat. Struct. Biol.* **6**, 569–573
6. Volkin, D. B., Tsai, P. K., Dabora, J. M., Gress, J. O., Burke, C. J., Linhardt, R. J., and Middaugh, C. R. (1993) Physical stabilization of acidic fibroblast growth factor by polyanions. *Arch. Biochem. Biophys.* **300**, 30–41
7. Tsai, P. K., Volkin, D. B., Dabora, J. M., Thompson, K. C., Bruner, M. W., Gress, J. O., Matuszewska, B., Keogan, M., Bondi, J. V., and Middaugh, C. R. (1993) Formulation design of acidic fibroblast growth factor. *Pharm. Res.* **10**, 649–659
8. Mach, H., Ryan, J. A., Burke, C. J., Volkin, D. B., and Middaugh, C. R. (1993) Partially structured self-associating states of acidic fibroblast growth factor. *Biochemistry* **32**, 7703–7711
9. Wiedlocha, A., Madhus, I. H., Mach, H., Middaugh, C. R., and Olsnes, S. (1992) Tight folding of acidic fibroblast growth factor prevents its translocation to the cytosol with diphtheria toxin as vector. *EMBO J.* **11**, 4835–4842
10. Kuelto, L. A., and Middaugh, C. R. (2000) Potential use of non-classical pathways for the transport of macromolecular drugs. *Expert Opin. Investig. Drugs* **9**, 2039–2050
11. Kuelto, L. A., Normand, N., O’Hare, P., and Middaugh, C. R. (2000)

- Conformational lability of herpesvirus protein VP22. *J. Biol. Chem.* **275**, 33213–33221
12. Jones, L. S., Yazzie, B., and Middaugh, C. R. (2004) Polyanions and the proteome. *Mol. Cell. Proteomics* **3**, 746–769
13. Visentin, G. P., Moghaddam, M., Beery, S. E., McFarland, J. G., and Aster, R. H. (2001) Heparin is not required for detection of antibodies associated with heparin-induced thrombocytopenia/thrombosis. *J. Lab. Clin. Med.* **138**, 22–31
14. Benezra, M., Vogel, T., Ben-Sasson, S. A., Panet, A., Sehayek, E., Al-Haideiri, M., Decklbaum, R. J., and Vlodavsky, I. (2001) A synthetic heparin-mimicking polyanionic compound binds to the LDL receptor-related protein and inhibits vascular smooth muscle cell proliferation. *J. Cell. Biochem.* **81**, 114–127
15. Fennewald, S. M., and Rando, R. F. (1995) Inhibition of high affinity basic fibroblast growth factor binding by oligonucleotides. *J. Biol. Chem.* **270**, 21718–21721
16. Kohler, J. J., and Schepartz, A. (2001) Effects of nucleic acids and polyanions on dimer formation and DNA binding by bZIP and bHLHZip transcription factors. *Bioorg. Med. Chem.* **9**, 2435–2443
17. Folkman, J., Weisz, P. B., Joullie, M. M., Li, W. W., and Ewing, W. R. (1989) Control of angiogenesis with synthetic heparin substitutes. *Science* **243**, 1490–1493
18. Middaugh, C. R., Mach, H., Burke, C. J., Volkin, D. B., Dabora, J. M., Tsai, P. K., Bruner, M. W., Ryan, J. A., and Marfia, K. E. (1992) Nature of the interaction of growth factors with suramin. *Biochemistry* **31**, 9016–9024
19. Pineda-Lucena, A., Jimenez, M. A., Nieto, J. L., Santoro, J., Rico, M., and Gimenez-Gallego, G. (1994) ¹H-NMR assignment and solution structure of human acidic fibroblast growth factor activated by inositol hexasulfate. *J. Mol. Biol.* **242**, 81–98
20. Yeh, B. K., Eliseenkova, A. V., Plotnikov, A. N., Green, D., Pinnell, J., Polat, T., Gritli-Linde, A., Linhardt, R. J., and Mohammadi, M. (2002) Structural basis for activation of fibroblast growth factor signaling by sucrose octasulfate. *Mol. Cell. Biol.* **22**, 7184–7192
21. Sedlak, E., and Antalík, M. (2002) Effect of polyanion on the acidic conformational transition of native and denatured ferricytochrome c. Circular dichroism study. *Gen. Physiol. Biophys.* **21**, 175–188
22. Chu, C. L., Goerges, A. L., and Nugent, M. A. (2005) Identification of common and specific growth factor binding sites in heparan sulfate proteoglycans. *Biochemistry* **44**, 12203–12213
23. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science* **293**, 2101–2105
24. Yu, Q., and Toole, B. P. (1995) Biotinylated hyaluronan as a probe for detection of binding proteins in cells and tissues. *BioTechniques* **19**, 122–124, 126–129
25. Apgar, J., Tseng, Y., Fedorov, E., Herwig, M. B., Almo, S. C., and Wirtz, D. (2000) Multiple-particle tracking measurements of heterogeneities in solutions of actin filaments and actin bundles. *Biophys. J.* **79**, 1095–1106
26. Templeton, D. M. (1988) General occurrence of isobestic points in the metachromatic dye complexes of sulfated glycosaminoglycans. *Int. J. Biol. Macromol.* **10**, 131–136
27. Wollin, A., and Jaques, L. B. (1973) Metachromasia: an explanation of the color change produced in dyes by heparin and other substances. *Thromb. Res.* **2**, 377–382
28. Fang, J., Haas, R. J., Dong, Y., and Lushington, G. H. (2005) Discover protein sequence signatures from protein-protein interaction data. *BMC Bioinformatics* **6**, 277–284
29. Bailey, T. L., and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36
30. Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bradley, P., Bork, P., Bucher, P., Cerutti, L., Copley, R., Courcelle, E., Das, U., Durbin, R., Fleischmann, W., Gough, J., Haft, D., Harte, N., Hulo, N., Kahn, D., Kanapin, A., Krestyaninova, M., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McDowall, J., Mitchell, A., Nikolskaya, A. N., Orchard, S., Pagni, M., Ponting, C. P., Quevillon, E., Selengut, J., Sigrist, C. J., Silventoinen, V., Studholme, D. J., Vaughan, R., and Wu, C. H. (2005) InterPro, progress and status in 2005. *Nucleic Acids Res.* **33**, D201–D205
31. Miteva, M. A., Tuffery, P., and Villoutreix, B. O. (2005) PCE: web tools to compute protein continuum electrostatics. *Nucleic Acids Res.* **33**, W372–W375
32. Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R., McCartney, R. R., Schmidt, M. C., Rachidi, N., Lee, S. J., Mah, A. S., Meng, L., Stark, M. J., Stern, D. F., De Virgilio, C., Tyers, M., Andrews, B., Gerstein, M., Schweitzer, B., Predki, P. F., and Snyder, M. (2005) Global analysis of protein phosphorylation in yeast. *Nature* **438**, 679–684
33. Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003) Protein disorder prediction: implications for structural proteomics. *Structure (Lond.)* **11**, 1453–1459
34. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., El Bakkoury, M., Foury, F., Friend, S. H., Gentale, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D. J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J. L., Riles, L., Roberts, C. J., Ross-MacDonald, P., Scheerens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R. K., Veronneau, S., Voet, M., Volckaert, G., Ward, T. R., Wysocki, R., Yen, G. S., Yu, K., Zimmermann, K., Philippsen, P., Johnston, M., and Davis, R. W. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906
35. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504
36. Bader, G. D., Donaldson, I., Wolting, C., Ouellette, B. F., Pawson, T., and Hogue, C. W. (2001) BIND: The Biomolecular Interaction Network Database. *Nucleic Acids Res.* **29**, 242–245
37. von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., and Bork, P. (2002) Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**, 399–403
38. Maere, S., Heymans, K., and Kuiper, M. (2005) BINGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449
39. Albert, R. (2005) Scale-free networks in cell biology. *J. Cell Sci.* **118**, 4947–4957
40. Hileman, R. E., Fromm, J. R., Weiler, J. M., and Linhardt, R. J. (1998) Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *BioEssays* **20**, 156–167
41. Sedlak, E., and Antalík, M. (1998) Coulombic and noncoulombic effect of polyanions on cytochrome c structure. *Biopolymers* **46**, 145–154
42. Cahuzac, B., Berthouneau, E., Birlirakis, N., Guittet, E., and Mirande, M. (2000) A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. *EMBO J.* **19**, 445–452
43. Cardin, A. D., and Weintraub, H. J. (1989) Molecular modeling of protein-glycosaminoglycan interactions. *Arteriosclerosis* **9**, 21–32
44. Hall, D. A., Zhu, H., Zhu, X., Royce, T., Gerstein, M., and Snyder, M. (2004) Regulation of gene expression by a metabolic enzyme. *Science* **306**, 482–484
45. Honnappa, S., Jahnke, W., Seelig, J., and Steinmetz, M. O. (2006) Control of intrinsically disordered stathmin by multisite phosphorylation. *J. Biol. Chem.* **281**, 16078–16083
46. Kysela, B., Chovanec, M., and Jeggo, P. A. (2005) Phosphorylation of linker histones by DNA-dependent protein kinase is required for DNA ligase IV-dependent ligation in the presence of histone H1. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1877–1882
47. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function. *Biochemistry* **41**, 6573–6582
48. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331
49. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., and Uversky, V. N. (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **272**, 5129–5148
50. Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., and Dunker, A. K. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049
51. Jeong, H., Mason, S. P., Barabasi, A. L., and Oltvai, Z. N. (2001) Lethality

- and centrality in protein networks. *Nature* **411**, 41–42
52. Das, B., Chattopadhyay, S., Bera, A. K., and Dasgupta, C. (1996) In vitro protein folding by ribosomes from *Escherichia coli*, wheat germ and rat liver: the role of the 50S particle and its 23S rRNA. *Eur. J. Biochem.* **235**, 613–621
 53. Dedmon, M. M., Patel, C. N., Young, G. B., and Pielak, G. J. (2002) FigM gains structure in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12681–12684
 54. Tang, Y. C., Chang, H. C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., Hartl, F. U., and Hayer-Hartl, M. (2006) Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* **125**, 903–914
 55. Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., and Li, W. H. (2003) Role of duplicate genes in genetic robustness against null mutations. *Nature* **421**, 63–66
 56. Mach, H., Volkin, D. B., Burke, C. J., Middaugh, C. R., Linhardt, R. J., Fromm, J. R., Loganathan, D., and Mattsson, L. (1993) Nature of the interaction of heparin with acidic fibroblast growth factor. *Biochemistry* **32**, 5480–5489
 57. Ashikari-Hada, S., Habuchi, H., Kariya, Y., Itoh, N., Reddi, A. H., and Kimata, K. (2004) Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. *J. Biol. Chem.* **279**, 12346–12354
 58. Kreuger, J., Jemth, P., Sanders-Lindberg, E., Eliahu, L., Ron, D., Basilico, C., Salmivirta, M., and Lindahl, U. (2005) Fibroblast growth factors share binding sites in heparan sulphate. *Biochem. J.* **389**, 145–150
 59. Spitzer, J. J., and Poolman, B. (2005) Electrochemical structure of the crowded cytoplasm *Trends Biochem. Sci.* **30**, 536–541
 60. de Lichtenberg, U., Jensen, L. J., Brunak, S., and Bork, P. (2005) Dynamic complex formation during the yeast cell cycle. *Science* **307**, 724–727

Supplementary Material

A Network-Based Analysis of Polyanion Binding Proteins Utilizing Yeast Protein Arrays

Nazila Salamat-Miller¹, Jianwen Fang², Christopher Seidel³, Aaron Smalter¹, Yassen Assenov⁴, Mario Albrecht⁴, and C. Russell Middaugh^{1,*}

¹Department of Pharmaceutical Chemistry, University of Kansas, 2030 Becker Dr., Lawrence, KS 66047.

² Bioinformatics Core Facility, University of Kansas, Lawrence, KS 66045.

³ Stowers Institute for Medical Research, 1000 E. 50th St., Kansas City, MO 64110.

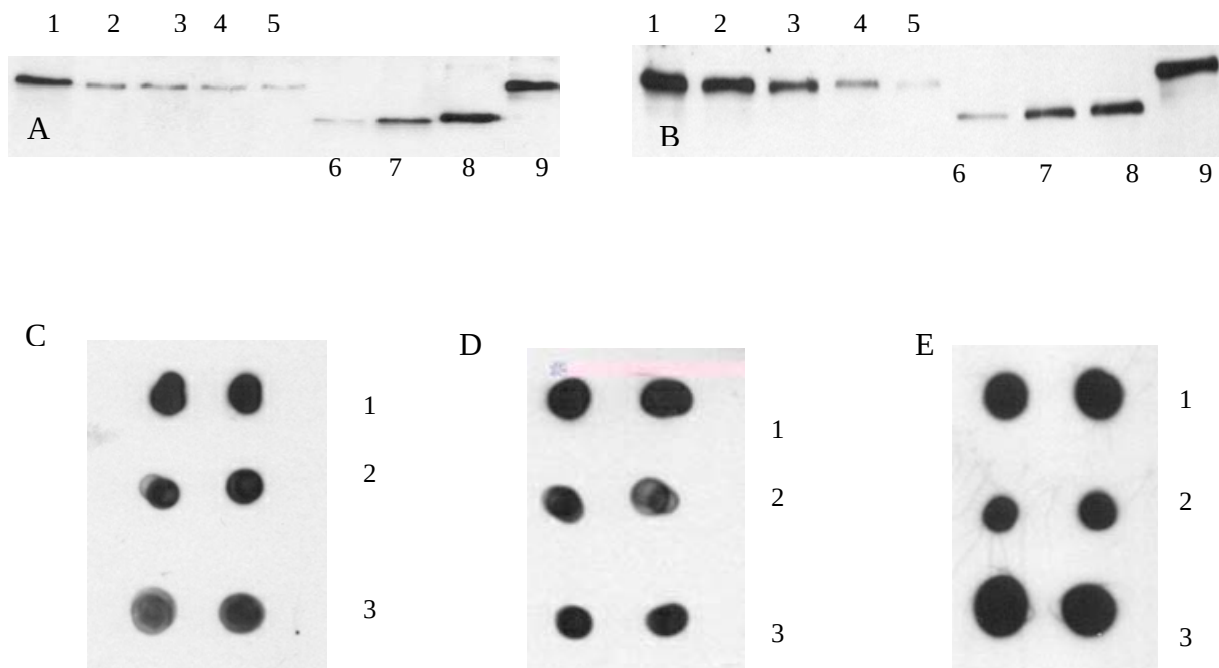
⁴ Max-Planck-Institut for Informatics, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany

*Corresponding author: C. Russell Middaugh, Phone: 875-864-5813, Fax: 875-864-5814, e-mail: middaugh@ku.edu

Supplementary Figures

SF 1

Western blots of biotinylated A) actin and B) tubulin as well as dot-blots of C) heparin, D) heparan sulfate (HS), and E) DNA. Lines 1 to 5 in panels A and B represent biotinylated standards (200-12.5 fmole biotin) and lines 6-8 represent biotinylated proteins with a 3/1, 9/1, and 27/1 ratio of biotin to protein. The last line (9) represents the biotinylated BSA standard at a ratio of 9/1. The row numbers in panels C, D, and E represent 200 fmole of; 1) biotinylated actin (3/1), 2) biotinylated BSA (9/1), 3) desalted biotinylated heparin, HS, and DNA.



SF 2

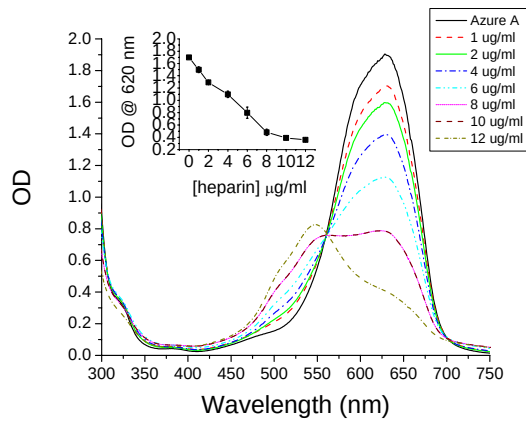
Azure A, in the absence of heparin or HS, exhibited a UV peak maximum at approximately 630 nm. Addition of heparin to the dye solution decreased the intensity of this peak and produced a new peak at a lower wavelength of 530 nm (characteristic of substances which produce a complex with this dye and demonstration of metachromacy, Figs. 2A and 2B) (Jaques 1979). This assay was linear only at low concentrations of heparin since excess heparin interfered with color formation of the complex (Jaques *et al.* 1947). Although, solutions of HS and Azure A did not produce a new peak at 530 nm, the intensity of the peak maximum of the dye was reduced following the addition of HS to the dye solution in a concentration dependent manner (this Fig. 2C and 2D).

Traditionally, two different dyes, Azure A and 1,9-dimethylmethylene blue (DMMB), have been used to quantify heparin and HS concentrations. In our experiments, while Azure A formed a spectroscopically detectable complex with heparin and HS in both biotinylated and non-biotinylated forms, DMMB failed to bind to the biotinylated form of HS (data not shown). Therefore, Azure A was selected to quantify biotinylated heparin and HS. This dye, similar to other thiazine dyes, is believed to form dimers through stacking interactions. It has been shown that certain structural arrangements of polysaccharides produce charge densities and geometries that permit the dimer to form a complex with the polysaccharide. This is detectable at lower wavelengths when compared with the wavelength of the original dye (Wollin and Jaques 1973). This band was observed for heparin-dye interactions at 530 nm and is shown in this section (2A and 2B). Azure A in the presence of HS did not manifest this new, single band but demonstrated a decrease in its peak maximum intensity in a HS concentration dependent fashion (Fig. 2C and 2D). This fact and the change in the color of dye solution upon addition

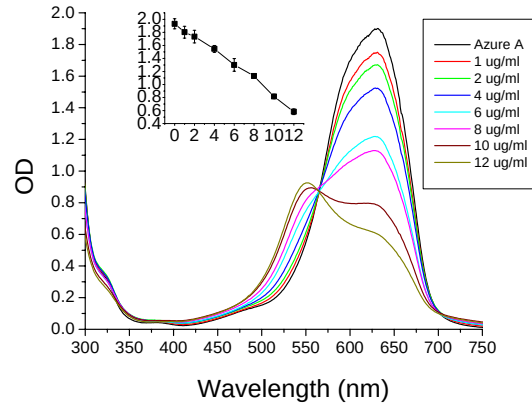
of HS were both indicative of complex formation. Sample heterogeneity has been shown to suppress the formation of a new band when compared with more homogenous solutions (Templeton 1988). Thus, we speculate that the HS solutions used here may be more heterogeneous than heparin. Additionally, the lower charge density of HS may play a role in this phenomenon (Wollin and Jaques 1973).

The ultraviolet absorption spectra are shown for A) heparin, B) biotinylated heparin, C) HS, and D) biotinylated HS in presence of azure A. The experiments were conducted in triplicate.

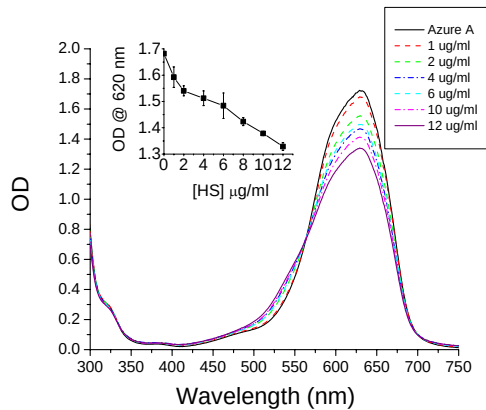
A)



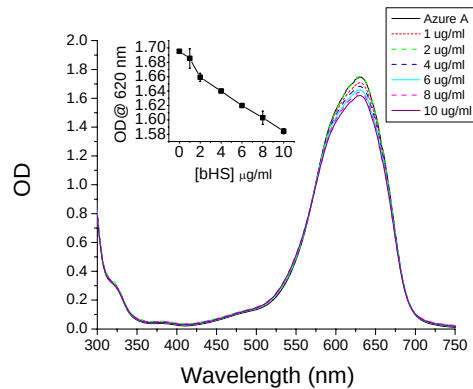
B)



C)

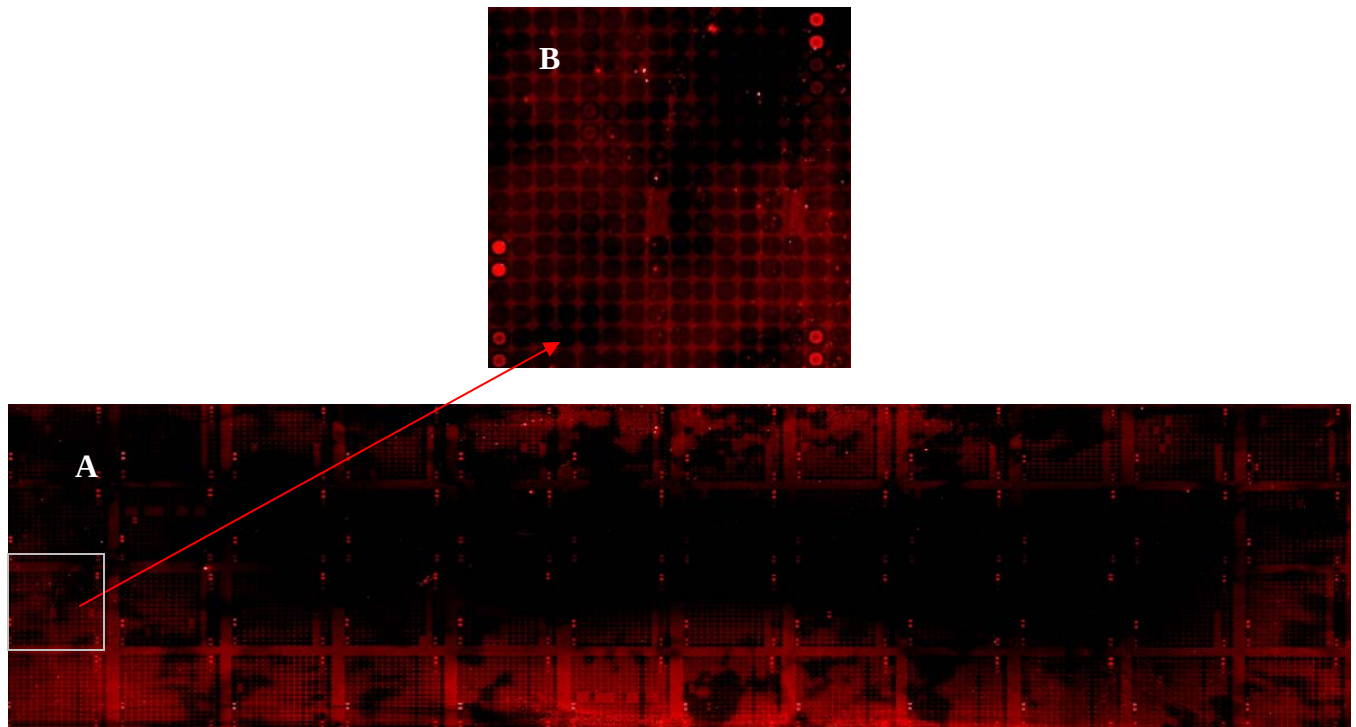


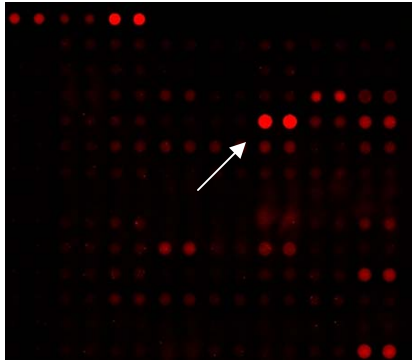
D)



SF 3

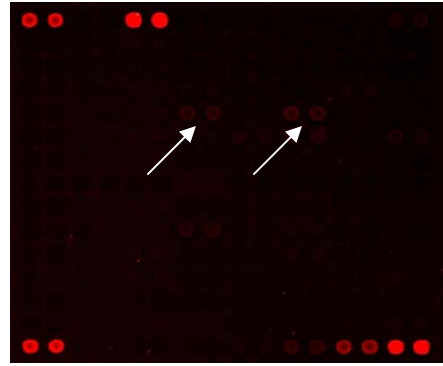
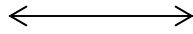
An example of images produced when probing the yeast protein array with biotinylated polyanions. A) The array containing ~5000 unique doubly spotted proteins on a nitrocellulose-coated slide was probed by tubulin. B) An enlarged image of one of the blocks is depicted above the proteome chip. The fluorescent spots represent the controls as well as the proteins which interacted with tubulin, C) a KSP slide probed with heparin, and D) a nitrocellulose-coated protein array probed with actin. The white arrows indicate representative proteins interacting with the probe. A larger number of interactions can be seen in the array which was probed with heparin compared to actin.





C)

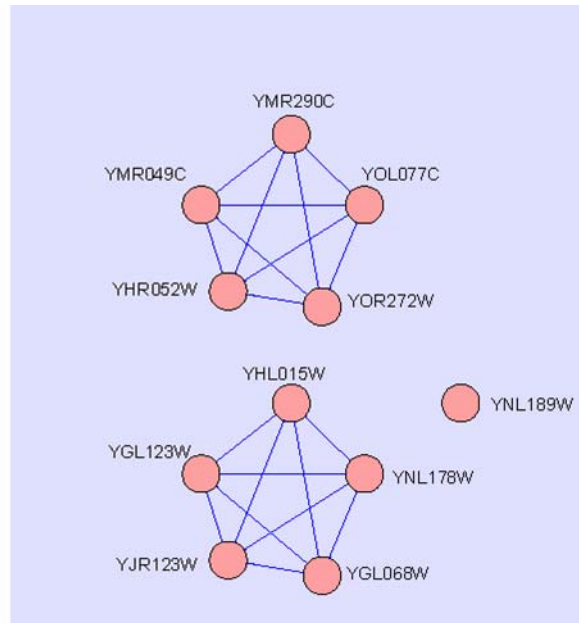
Control spots



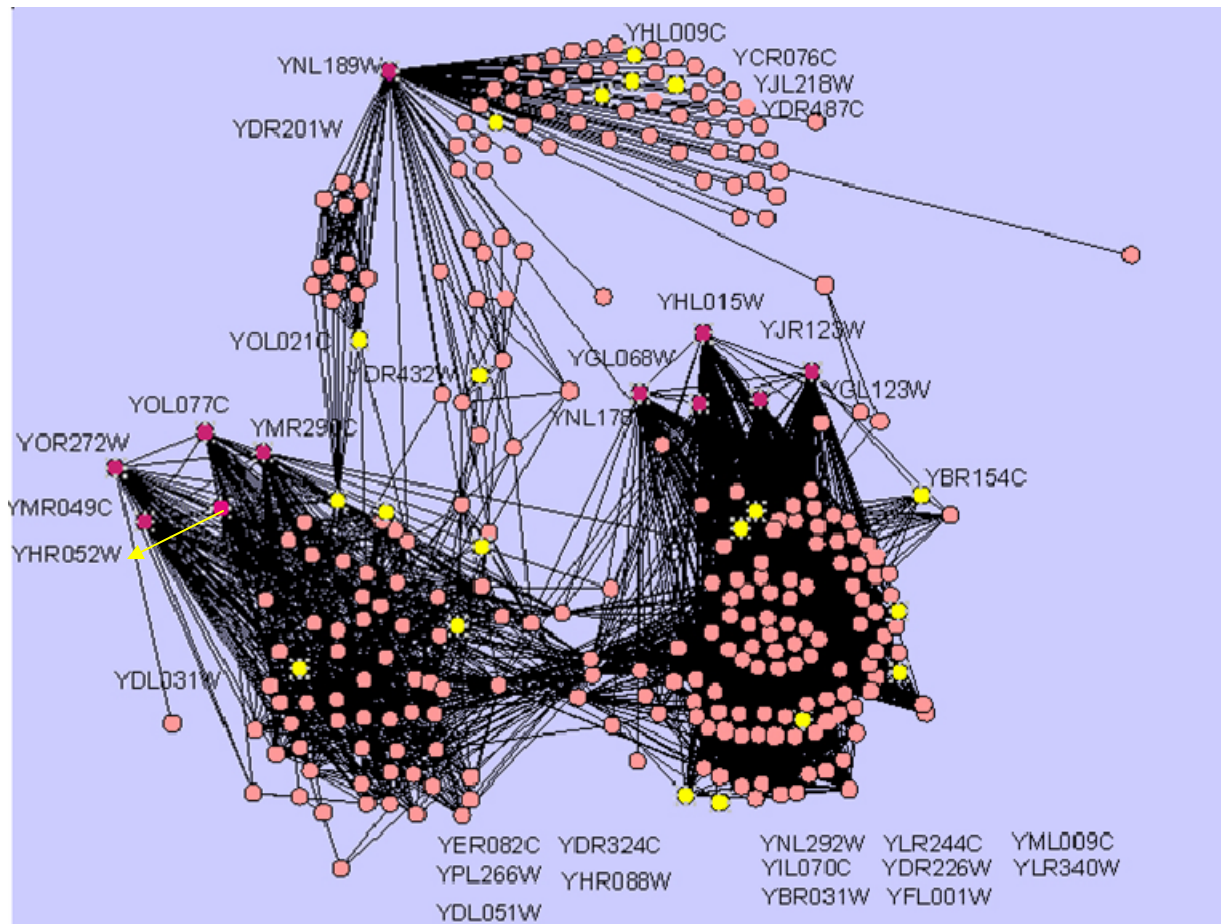
D)

SF 4

A) The PABPs which served as hubs (with more than 40 interaction partners) were selected from von Mering *et al.* (von Mering *et al.* 2002) database and demonstrated 20 interactions.



Upon further analysis of these hubs and after obtaining their first interacting neighbors from the same dataset, 287 proteins with 4,204 interactions were found, which is illustrated in the following figure. The 11 hubs and the available PABPs are represented by magenta and yellow nodes, respectively. The names of the non-hubs and non-PABPs are omitted from this figure for simplification purposes.



B) Visualization of 51 hubs in Cytoscape:

a) We investigated the number of interacting partners for PABPs which had more than 15 interaction partners and found 51 such proteins. Only 42 of these proteins were present in the YeastHighQuality.sif of Cytoscape data which were connected through 28 edges. When the von Mering *et al.* (von Mering *et al.* 2002) dataset was selected, all 51 hubs were present and demonstrated 79 interactions.

b) To take this analysis one step further, we used Cytoscape to identify the first neighbors of the 51 hub PABPs which included some PABPs and many non-PABPs. Using the

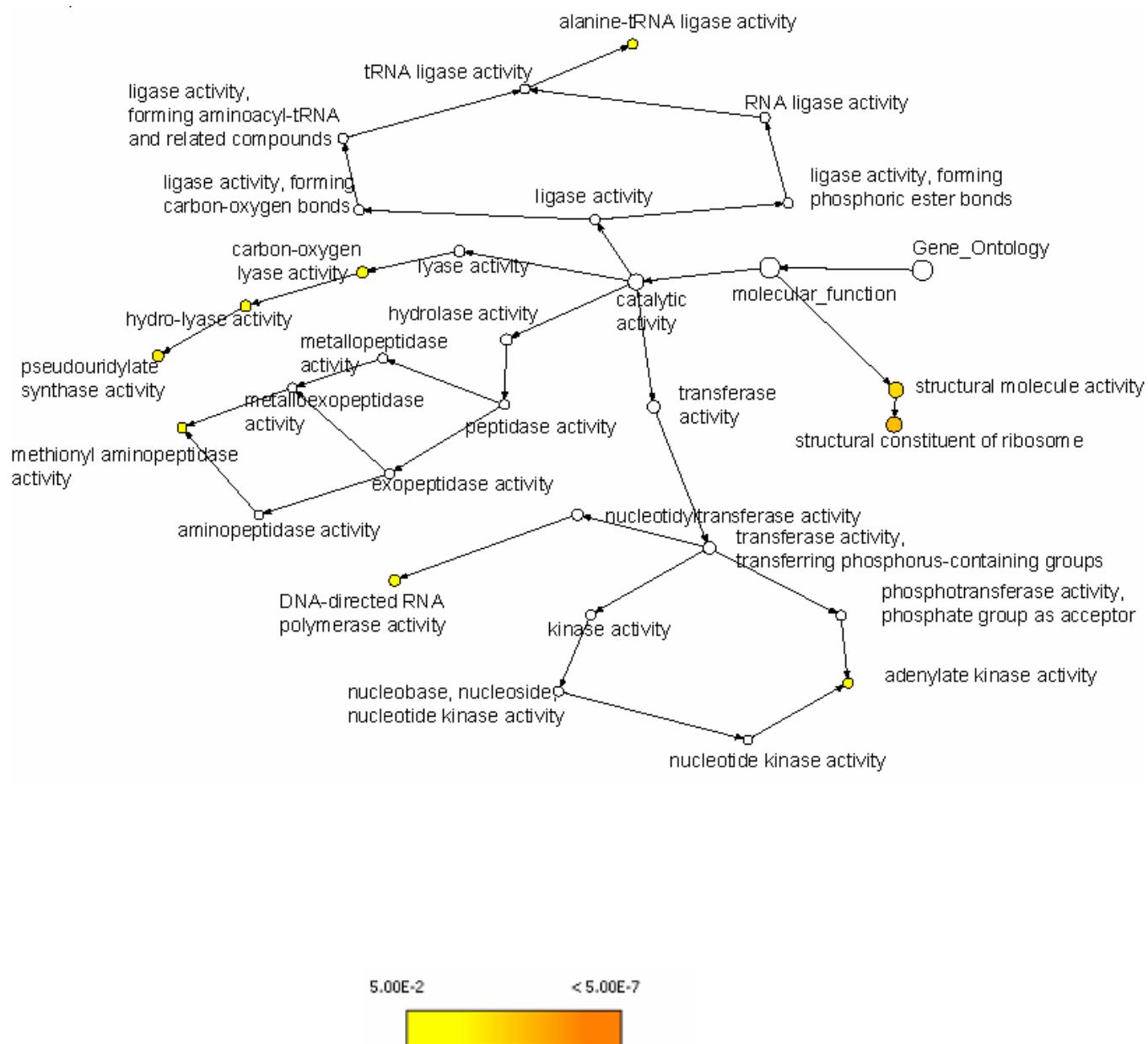
YeastHighQuality.sif and von Mering *et al.* dataset, 344 proteins with 1,639 interactions and 671 proteins with 6,712 interactions were identified, respectively. Similarly, a high degree of interaction between these proteins and the rest of yeast proteome argues for an important role of the PABPs in yeast.

The use of BiNGO to investigate the GO classification of 671 proteins with 6,712 interactions, obtained from the von Mering *et al.* dataset, resulted in a complicated and extensive array of molecular functions including small nucleolar RNA (sno-RNA) binding, transferase activity, DNA-directed RNA polymerase activity, and RNA polymerase II transcription mediator activity among others. Ribosomal subunit assembly, rRNA processing, translation, and macromolecular metabolism were the common biological processes identified using BiNGO. The 51 hub PABPs were located in the ribosomal subunits, proteasome regulatory particle, RNA polymerase complex, and many other cell locations.

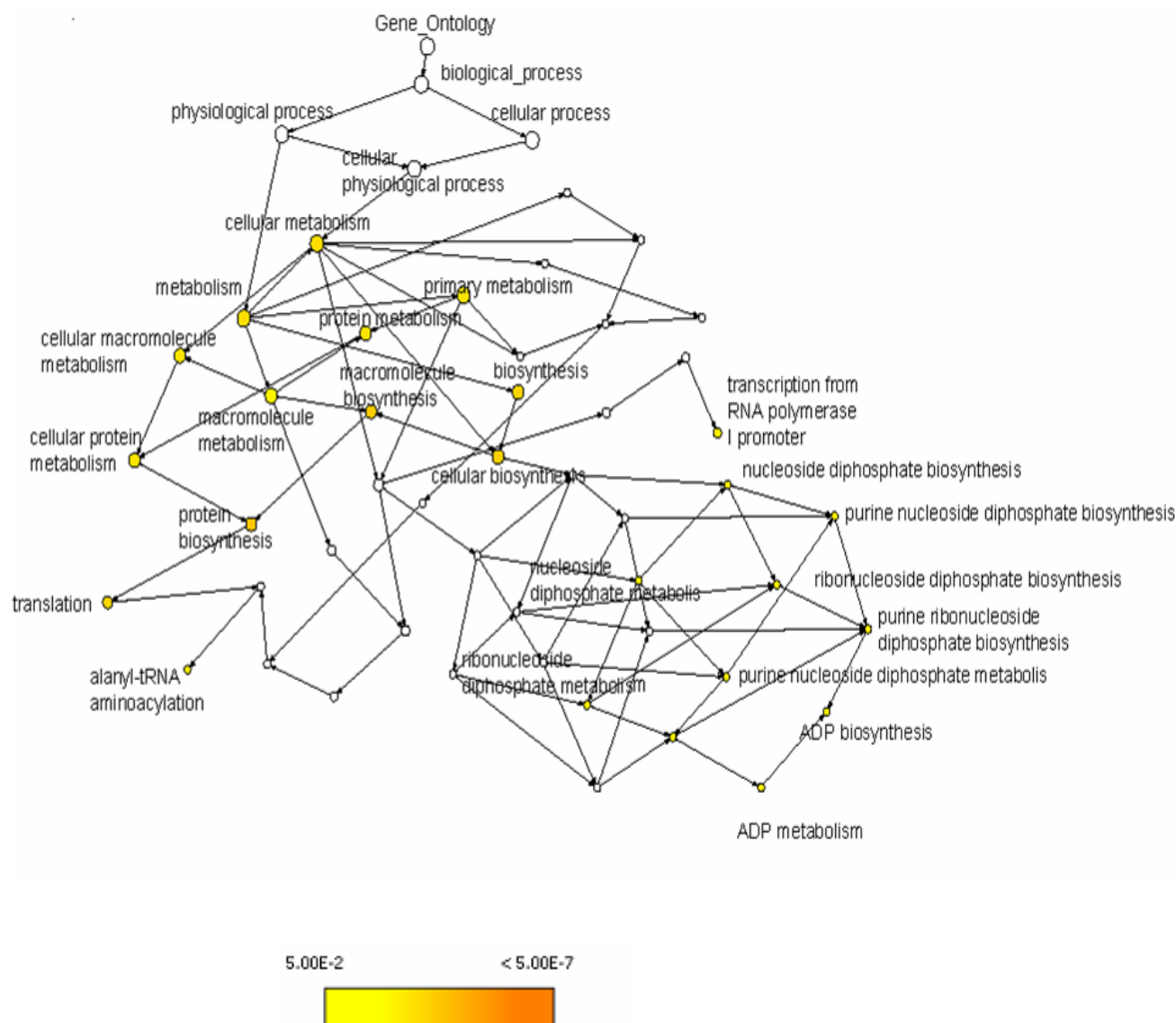
SF 5

A physical interaction network map of the yeast PABPs was generated using Cytoscape. We utilized the dataset based on high and medium confidence levels produced by von Mering et al. (von Mering *et al.* 2002). This dataset consists of 2,617 proteins (nodes) and 11,855 interactions (edges). We were able to identify 266 out of 529 PABPs in this network which exhibited 168 interactions. Cytoscape demonstrated that these 266 proteins interacted with each other in the form of two major clusters. The functional analysis of the first and larger cluster) is described in the text. The functional analysis of the second cluster (17 nodes and 38 edges) was conducted in a similar manner to the first. Panels A, B, and C represent the molecular function, biological process, and cellular compartment associated with this cluster of proteins. In some cases captions have been removed for the non-statistical protein categories (represented as white circles) for better clarity of the figures.

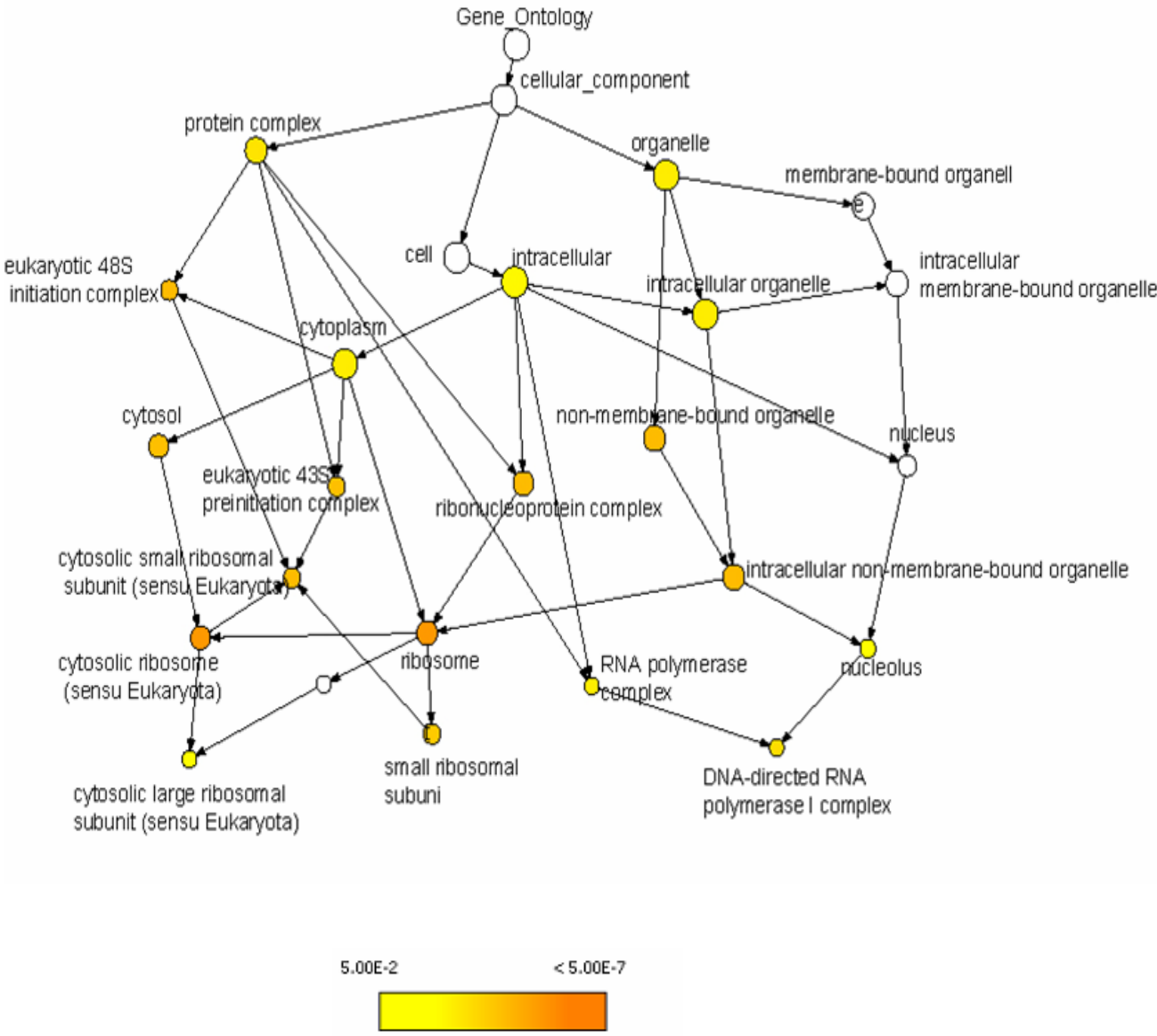
A)



B)



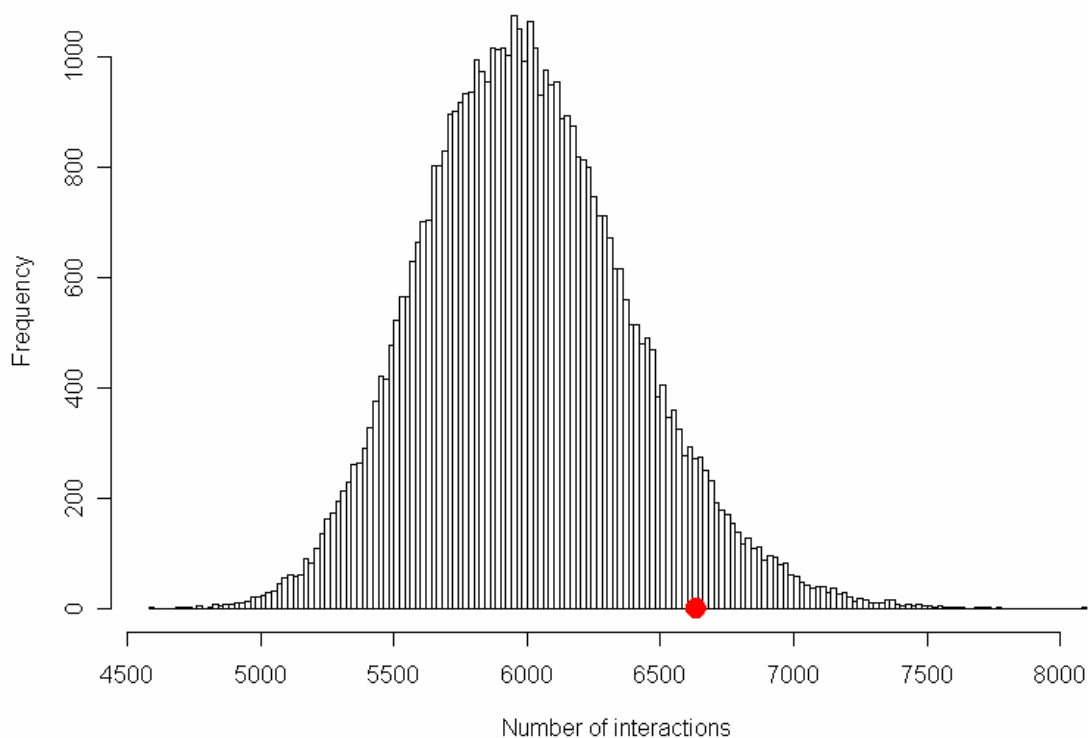
C)



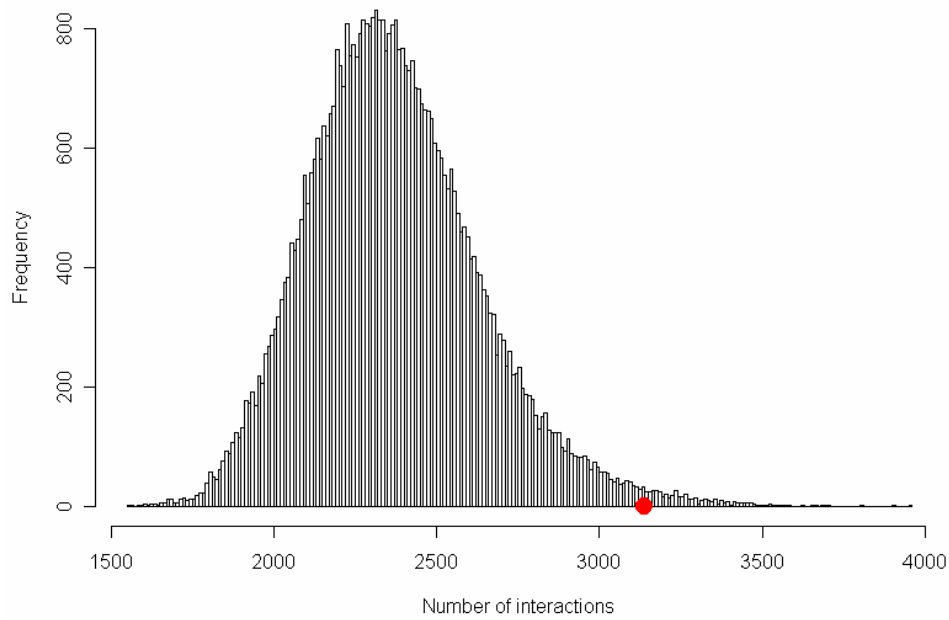
SF 6

Histograms for the number of protein-protein interactions in random sub-networks. These histograms are produced for three different statistical levels of +1, +2, and +3 SD of all PABPs detection to estimate the p -value shown in ST 8. The red circle represents the number of interactions in the sub-network formed by PABPs and their direct interaction partners.

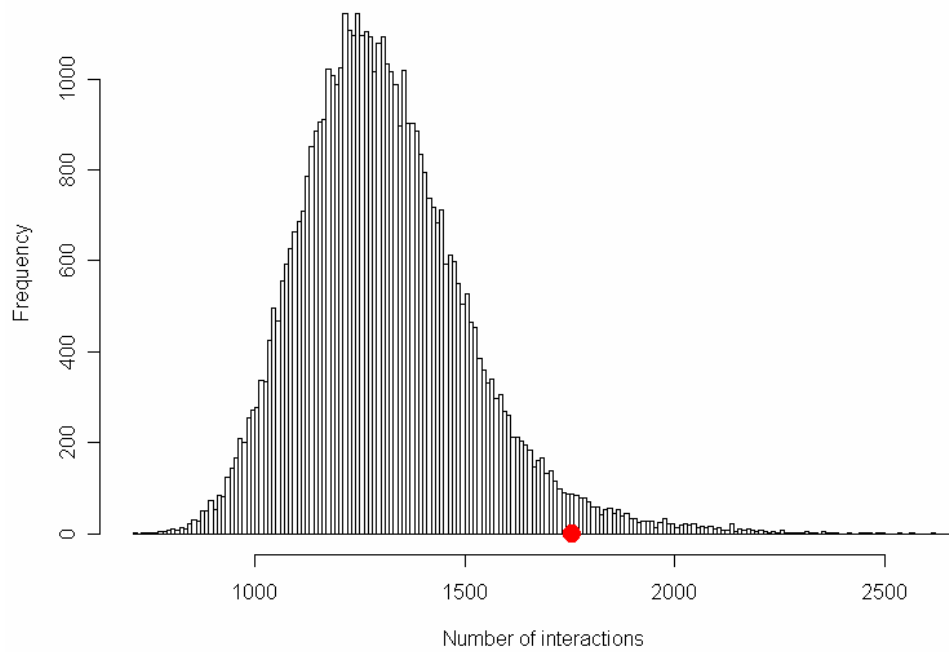
1SD



2SD



3SD



Supplementary Tables

ST 1

Definition of types of experiments, data, and statements in Table 1 (From <http://db.yeastgenome.org/cgi-bin/GO/goEvidence.pl>).

Code	Definition	Examples
IC	Inferred <u>B</u> y <u>C</u> urator	Used when an annotation is not supported by any evidence, but can be reasonably inferred by a curator from other GO annotations for which evidence is available.
IDA	Inferred from <u>D</u> irect <u>A</u> ssay	- Enzyme assays - In vitro reconstitution - Immunofluorescence - Cell fractionation - Physical interaction/binding assay
IEA	Inferred from <u>E</u> lectronic <u>A</u> nnotation	Annotations based on "hits" in sequence similarity searches, if not curator-reviewed. Annotations transferred from database records, if not curator-reviewed.
IEP	Inferred from <u>E</u> xpression <u>P</u> attern	- Transcript levels [e.g. Northern, microarray data] - Protein levels [e.g. Western blots]
IGI	Inferred from <u>G</u> enetic <u>I</u> nteraction	- Traditional genetic interactions such as suppressors, synthetic lethals, etc. - Functional complementation - Rescue experiments
IMP	Inferred from <u>M</u> utant <u>P</u> henotype	- Any gene mutation/knockout - Overexpression/ectopic expression of wild-type or mutant genes - Anti-sense experiments - Specific protein inhibitors
IPI	Inferred from <u>P</u> hysical <u>I</u> nteraction	2-hybrid interactions - Co-purification - Co-immunoprecipitation - Ion/protein binding experiments
ISS	Inferred from <u>S</u> equence or structural <u>S</u> imilarity	- Sequence similarity [homolog of/most closely related to] - Recognized domains - Structural similarity - Southern blotting
NAS	<u>N</u> on-traceable <u>A</u> uthor <u>S</u> tatement	- Database entries that do not cite a paper [e.g. SwissProt records, YPD protein reports] - Statements in papers [abstract, introduction, or discussion] that a curator cannot trace to another publication

ND	<u>N</u>o <u>B</u>iological <u>D</u>ata <u>A</u>vailable	used for annotations to unknown molecular function, biological process, or cellular component.
NR	<u>N</u>ot <u>R</u>ecorded	Used for annotations done before the SGD curators began tracking evidence types. It should not be used for new annotations
RCA	<u>R</u>eviewed <u>C</u>omputational <u>A</u>nalysis	Used for predictions based on large-scale experiments, integration of large-scale datasets, text-based computation
TAS	<u>T</u>raceable <u>A</u>uthor <u>S</u>tatement	Anything in a review article where the original experiments are traceable through that article, or in a textbook or dictionary [e.g. "everybody" knows that enolase is a glycolytic enzyme]

ST 2

The standard ORF list of PABPs generated using yeast protein arrays. The reported data are generated at 0.15 M NaCl and contain PABPs with low, medium, and high interaction confidence level.

Yeast				
Actin	Tubulin	Heparin	HS	DNA
YBL024W	YAL060W	YAL009W	YAL009W	YBL032W
YBL055C	YBL011W	YBL024W	YBL007C	YBL055C
YBR031W	YBL065W	YBL030C	YBL024W	YBR031W
YBR042C	YBL103C	YBL055C	YBL030C	YBR118W
YBR118W	YBR031W	YBL057C	YBL036C	YCL014W
YBR222C	YBR118W	YBL060W	YBL055C	YCR016W
YBR299W	YBR159W	YBL082C	YBL057C	YCR024C
YCR016W	YBR222C	YBL098W	YBL082C	YCR028C-A
YCR024C	YCR016W	YBR025C	YBL084C	YCR072C
YCR072C	YCR076C	YBR031W	YBR025C	YDL002C
YDL027C	YDL051W	YBR042C	YBR029C	YDL051W
YDL051W	YDL053C	YBR068C	YBR031W	YDL078C
YDL078C	YDL078C	YBR118W	YBR042C	YDL153C
YDL153C	YDL101C	YBR121C	YBR043C	YDL177C
YDL177C	YDL171C	YBR159W	YBR056W	YDR120C
YDR120C	YDL173W	YCR016W	YBR118W	YDR257C
YDR256C	YDL177C	YCR024C	YBR121C	YDR432W

YDR257C	YDR467C	YCR028C-A	YBR154C	YDR440W
YDR376W	YER091C	YCR072C	YBR210W	YER034W
YDR432W	YER144C	YDL002C	YBR224W	YER082C
YDR516C	YGL029W	YDL007W	YBR231C	YER101C
YER056C-A	YGL046W	YDL027C	YBR237W	YER122C
YER070W	YGL196W	YDL051W	YCL014W	YER125W
YER082C	YGR026W	YDL078C	YCL035C	YFR034C
YER087W	YHL033C	YDL107W	YCL061C	YGL062W
YER091C	YHR117W	YDL128W	YCR016W	YGL105W
YER101C	YIL019W	YDR120C	YCR024C	YGR026W
YER125W	YIL084C	YDR153C	YCR028C-A	YHL013C
YER152C	YIR008C	YDR161W	YDL002C	YHL033C
YFL001W	YJL031C	YDR188W	YDL027C	YHR090C
YFL054C	YJL173C	YDR231C	YDL031W	YHR117W
YFL058W	YKL177W	YDR234W	YDL051W	YIL010W
YFR010W	YKL186C	YDR257C	YDL070W	YIL016W
YGL046W	YKR009C	YDR312W	YDL078C	YIL035C
YGL062W	YKR092C	YDR323C	YDL099W	YIL055C
YGL184C	YLR003C	YDR376W	YDL107W	YIL091C
YGL207W	YLR129W	YDR411C	YDL114W	YIR001C
YGL245W	YLR139C	YDR429C	YDL130W-A	YIR034C
YGR026W	YLR150W	YDR513W	YDL142C	YJL033W
YGR256W	YLR192C	YDR538W	YDL153C	YJL080C
YGR264C	YLR287C	YEL006W	YDL177C	YJL218W
YGR287C	YMR039C	YEL015W	YDL183C	YKL023W
YGR292W	YMR112C	YEL053C	YDR021W	YKR009C
YHL013C	YMR153W	YER036C	YDR023W	YKR060W
YHL033C	YNL009W	YER058W	YDR120C	YKR092C
YHR009C	YNL117W	YER082C	YDR145W	YLL018C
YHR063C	YNL178W	YER101C	YDR147W	YLR003C
YHR067W	YOL033W	YER122C	YDR161W	YLR150W
YHR112C	YOL049W	YER125W	YDR173C	YLR192C
YHR117W	YOL072W	YER151C	YDR181C	YLR204W
YHR183W	YOL123W	YER182W	YDR201W	YLR244C
YHR189W	YOR056C	YFL001W	YDR226W	YLR363C
YIL010W	YOR162C	YFR034C	YDR231C	YLR410W
YIL035C	YOR196C	YFR038W	YDR234W	YML034W
YIL075C	YOR198C	YFR042W	YDR244W	YMR039C
YIL082W-A	YOR309C	YGL010W	YDR257C	YMR140W
YIL091C	YPL027W	YGL055W	YDR312W	YMR153W
YIL096C	YPR086W	YGL080W	YDR323C	YMR229C
YIR034C	YPR115W	YGL101W	YDR328C	YMR290C
YJR013W	YPR139C	YGL105W	YDR335W	YMR316W
YJR019C		YGL119W	YDR344C	YNL112W
YJR096W		YGL148W	YDR346C	YNL175C
YJR148W		YGL160W	YDR405W	YNL186W
YKL006W		YGL185C	YDR429C	YNL196C
YKL023W		YGL207W	YDR440W	YNL227C

YKL155C		YGL245W	YDR487C	YNL292W
YKL182W		YGL254W	YDR499W	YOL077C
YKR009C		YGL256W	YDR513W	YOL123W
YKR060W		YGR012W	YDR514C	YOR056C
YKR079C		YGR021W	YDR530C	YOR078W
YKR092C		YGR026W	YEL006W	YOR162C
YLL026W		YGR053C	YEL017W	YOR198C
YLR003C		YGR083C	YEL039C	YOR288C
YLR017W		YGR135W	YER014W	YOR309C
YLR022C		YGR157W	YER044C	YOR386W
YLR027C		YGR196C	YER056C-A	YPL037C
YLR069C		YGR235C	YER058W	YPL133C
YLR134W		YGR252W	YER082C	YPL146C
YLR150W		YGR256W	YER120W	YPL150W
YLR225C		YHL009C	YER125W	YPL192C
YLR244C		YHL013C	YER182W	YPL249C-A
YLR258W		YHL015W	YFL001W	YPL266W
YLR363C		YHL019C	YFL016C	YPR008W
YLR460C		YHL033C	YFR007W	YPR086W
YML048W		YHL048W	YFR029W	YPR115W
YML073C		YHR003C	YFR034C	YPR139C
YML080W		YHR057C	YFR038W	
YMR004W		YHR073W	YFR042W	
YMR102C		YHR088W	YGL010W	
YMR120C		YHR117W	YGL038C	
YMR140W		YHR131C	YGL043W	
YMR153W		YHR133C	YGL050W	
YMR182C		YHR206W	YGL068W	
YMR205C		YHR207C	YGL079W	
YMR290C		YIL010W	YGL105W	
YMR304W		YIL016W	YGL119W	
YNL009W		YIL035C	YGL122C	
YNL045W		YIL051C	YGL123W	
YNL112W		YIL053W	YGL126W	
YNL117W		YIL057C	YGL160W	
YNL175C		YIL070C	YGL207W	
YNL186W		YIL072W	YGL245W	
YNL196C		YIL078W	YGL250W	
YNL219C		YIL082W-A	YGL256W	
YNL220W		YIL089W	YGR017W	
YNL227C		YIL091C	YGR026W	
YNL280C		YIL096C	YGR053C	
YOL021C		YIL124W	YGR116W	
YOL033W		YIL156W	YGR130C	
YOL066C		YIR008C	YGR179C	
YOL077C		YJL033W	YGR196C	
YOR002W		YJL080C	YGR205W	
YOR056C		YJL097W	YGR207C	

YOR078W		YJL110C	YGR235C	
YOR209C		YJL117W	YGR240C	
YOR234C		YJL159W	YGR249W	
YOR281C		YJL218W	YGR264C	
YOR286W		YJR013W	YGR276C	
YOR288C		YJR019C	YGR285C	
YOR309C		YJR032W	YHL009C	
YOR335C		YJR072C	YHL013C	
YOR386W		YJR077C	YHL015W	
YPL023C		YJR100C	YHL019C	
YPL133C		YJR149W	YHL033C	
YPL146C		YKL008C	YHR008C	
YPL222W		YKL023W	YHR017W	
YPL244C		YKL064W	YHR052W	
YPL249C-A		YKL081W	YHR065C	
YPL266W		YKL146W	YHR081W	
YPR035W		YKL150W	YHR090C	
YPR086W		YKL155C	YHR117W	
YPR115W		YKL212W	YHR131C	
YPR135W		YKL219W	YHR133C	
YPR139C		YKR009C	YHR156C	
YPR160W		YKR030W	YHR182W	
		YKR041W	YHR183W	
		YKR048C	YHR191C	
		YKR049C	YHR193C	
		YKR056W	YHR201C	
		YKR060W	YHR206W	
		YKR084C	YIL008W	
		YKR094C	YIL010W	
		YLL014W	YIL016W	
		YLL018C	YIL035C	
		YLL026W	YIL044C	
		YLR003C	YIL057C	
		YLR022C	YIL072W	
		YLR034C	YIL075C	
		YLR050C	YIL089W	
		YLR059C	YIL091C	
		YLR069C	YIL096C	
		YLR089C	YIL103W	
		YLR098C	YIL124W	
		YLR132C	YIL140W	
		YLR139C	YJL033W	
		YLR150W	YJL053W	
		YLR192C	YJL080C	
		YLR211C	YJL097W	
		YLR244C	YJL138C	
		YLR259C	YJL159W	
		YLR262C	YJL218W	

		YLR270W	YJR006W	
		YLR293C	YJR013W	
		YLR323C	YJR019C	
		YLR326W	YJR085C	
		YLR392C	YJR100C	
		YLR410W	YJR123W	
		YLR429W	YJR141W	
		YLR451W	YKL008C	
		YML030W	YKL023W	
		YML035C	YKL044W	
		YML048W	YKL064W	
		YML054C	YKL079W	
		YML126C	YKL099C	
		YMR004W	YKL114C	
		YMR039C	YKL149C	
		YMR102C	YKL150W	
		YMR110C	YKL155C	
		YMR115W	YKL160W	
		YMR140W	YKL182W	
		YMR153W	YKL195W	
		YMR205C	YKL197C	
		YMR216C	YKL211C	
		YMR229C	YKR009C	
		YMR262W	YKR018C	
		YMR290C	YKR021W	
		YMR316W	YKR022C	
		YNL001W	YKR030W	
		YNL007C	YKR041W	
		YNL009W	YKR042W	
		YNL045W	YKR048C	
		YNL048W	YKR049C	
		YNL077W	YKR060W	
		YNL101W	YKR084C	
		YNL104C	YKR094C	
		YNL135C	YLL010C	
		YNL145W	YLL022C	
		YNL147W	YLL026W	
		YNL155W	YLR003C	
		YNL186W	YLR022C	
		YNL196C	YLR026C	
		YNL217W	YLR043C	
		YNL219C	YLR050C	
		YNL220W	YLR069C	
		YNL227C	YLR089C	
		YNL272C	YLR117C	
		YNL292W	YLR125W	
		YNL293W	YLR126C	
		YNL304W	YLR132C	

	YNL315C	YLR139C	
	YNR001C	YLR150W	
	YNR013C	YLR192C	
	YNR018W	YLR211C	
	YNR019W	YLR225C	
	YNR041C	YLR241W	
	YOL021C	YLR244C	
	YOL027C	YLR251W	
	YOL033W	YLR259C	
	YOL093W	YLR265C	
	YOL147C	YLR314C	
	YOR002W	YLR323C	
	YOR056C	YLR326W	
	YOR059C	YLR340W	
	YOR070C	YLR363C	
	YOR100C	YLR390W	
	YOR162C	YLR410W	
	YOR208W	YLR429W	
	YOR228C	YLR451W	
	YOR232W	YLR460C	
	YOR271C	YML030W	
	YOR272W	YML036W	
	YOR276W	YML054C	
	YOR286W	YML055W	
	YOR288C	YML074C	
	YOR309C	YML080W	
	YOR362C	YML107C	
	YOR386W	YML126C	
	YPL030W	YMR004W	
	YPL041C	YMR049C	
	YPL048W	YMR071C	
	YPL050C	YMR102C	
	YPL094C	YMR110C	
	YPL104W	YMR113W	
	YPL110C	YMR138W	
	YPL123C	YMR139W	
	YPL133C	YMR140W	
	YPL146C	YMR153W	
	YPL244C	YMR165C	
	YPL249C-A	YMR178W	
	YPL266W	YMR182C	
	YPR086W	YMR205C	
	YPR113W	YMR216C	
	YPR115W	YMR229C	
	YPR128C	YMR268C	
	YPR139C	YMR282C	
	YPR174C	YMR290C	
	YPR183W	YMR322C	

		YNL006W	
		YNL007C	
		YNL101W	
		YNL112W	
		YNL125C	
		YNL175C	
		YNL181W	
		YNL186W	
		YNL189W	
		YNL196C	
		YNL217W	
		YNL219C	
		YNL225C	
		YNL227C	
		YNL272C	
		YNL280C	
		YNL292W	
		YNL293W	
		YNR001C	
		YNR013C	
		YNR015W	
		YNR018W	
		YNR019W	
		YNR041C	
		YOL023W	
		YOL027C	
		YOL033W	
		YOL077C	
		YOL141W	
		YOL147C	
		YOR002W	
		YOR029W	
		YOR045W	
		YOR056C	
		YOR059C	
		YOR070C	
		YOR071C	
		YOR078W	
		YOR100C	
		YOR103C	
		YOR162C	
		YOR177C	
		YOR189W	
		YOR201C	
		YOR208W	
		YOR217W	
		YOR228C	
		YOR241W	

			YOR271C	
			YOR286W	
			YOR288C	
			YOR309C	
			YOR335C	
			YOR340C	
			YOR362C	
			YOR386W	
			YPL030W	
			YPL042C	
			YPL048W	
			YPL050C	
			YPL057C	
			YPL074W	
			YPL094C	
			YPL104W	
			YPL133C	
			YPL137C	
			YPL140C	
			YPL143W	
			YPL150W	
			YPL210C	
			YPL244C	
			YPL249C-A	
			YPL266W	
			YPR004C	
			YPR008W	
			YPR019W	
			YPR024W	
			YPR031W	
			YPR082C	
			YPR086W	
			YPR125W	
			YPR128C	
			YPR139C	
			YPR143W	
			YPR174C	
			YPR183W	
			YPR195C	

ST 3

Number of actin binding proteins at zero and incrementally increasing concentrations of NaCl.

Actin 25 µg/ml					
NaCl (M)		0	0.05	0.15	0.75
Total proteins with positive interaction		2315	1780	1725	-
Total proteins with significant interaction	1SD	135	151	135	-
	2SD	55	69	63	-
	3SD	26	40	27	-

ST 4

Number of tubulin binding proteins at zero and incrementally increasing concentrations of NaCl.

Tubulin 50 µg/ml					
NaCl (M)		0	0.05	0.15	0.75
Total proteins with positive interaction		1140	220	216	207
Total proteins with significant interaction	1SD	326	60	60	52
	2SD	88	15	13	14
	3SD	9	7	10	9

ST 5

Number of HS binding proteins at zero and incrementally increasing concentrations of NaCl.

HS 50 µg/ml					
NaCl (M)		0	0.05	0.15	0.75
Total proteins with positive interaction		4266	4274	4269	4272
Total proteins with significant interaction	1SD	376	350	348	346
	2SD	129	121	124	123
	3SD	70	65	66	64

ST 6

Number of DNA binding proteins at zero and incrementally increasing concentrations of NaCl.

DNA 50 µg/ml					
NaCl (M)		0	0.05	0.15	0.75
Total proteins with positive interaction		927	777	887	944
Total proteins with significant interaction	1SD	101	93	86	85
	2SD	54	48	44	44
	3SD	31	30	29	29

ST 7

List of the yeast PABPs with solved three dimensional structures.

The surface potential calculations were not successful for the 1YGP structure (designated in red), due to the problems with the backbone structure deposited in the PDB.

Number	ORF	Swiss Prot	Selected PDB structure
1	YBL036C	P38197	1B54
2	YBR118W	P02994	1F60
3	YBR154C	P20434	1DZF
4	YDR226W	P07170	1AKY
5	YDR256C	P15202	1A4E
6	YDR328C	P52286	1NEX
7	YEL039C	P00045	1YTC
8	YFR034C	P07270	1A0A
9	YGR135W	P23638	1RYP, Similar to P21242
10	YGR205W	P42938	
11	YGR252W	Q03330	1YGH
12	YJL138C	P10081	1FUU
13	YKL186C	P34232	1OF5
14	YLL018C	P04802	1EOV
15	YLR026C	Q01590	1MQS
16	YLR027C	P23542	1YAA
17	YML054C	P00175	1KBI
18	YMR216C	Q03656	1HOW
19	YNL135C	P20081	1YAT
20	YNL189W	Q02821	1BK5
21	YOL049W	Q08220	1M0T
22	YOR362C	P21242	1RYP
23	YPL048W	P29547	1NHY
24	YPR160W	P06738	1YGP

Selection of crystal structures of yeast proteins

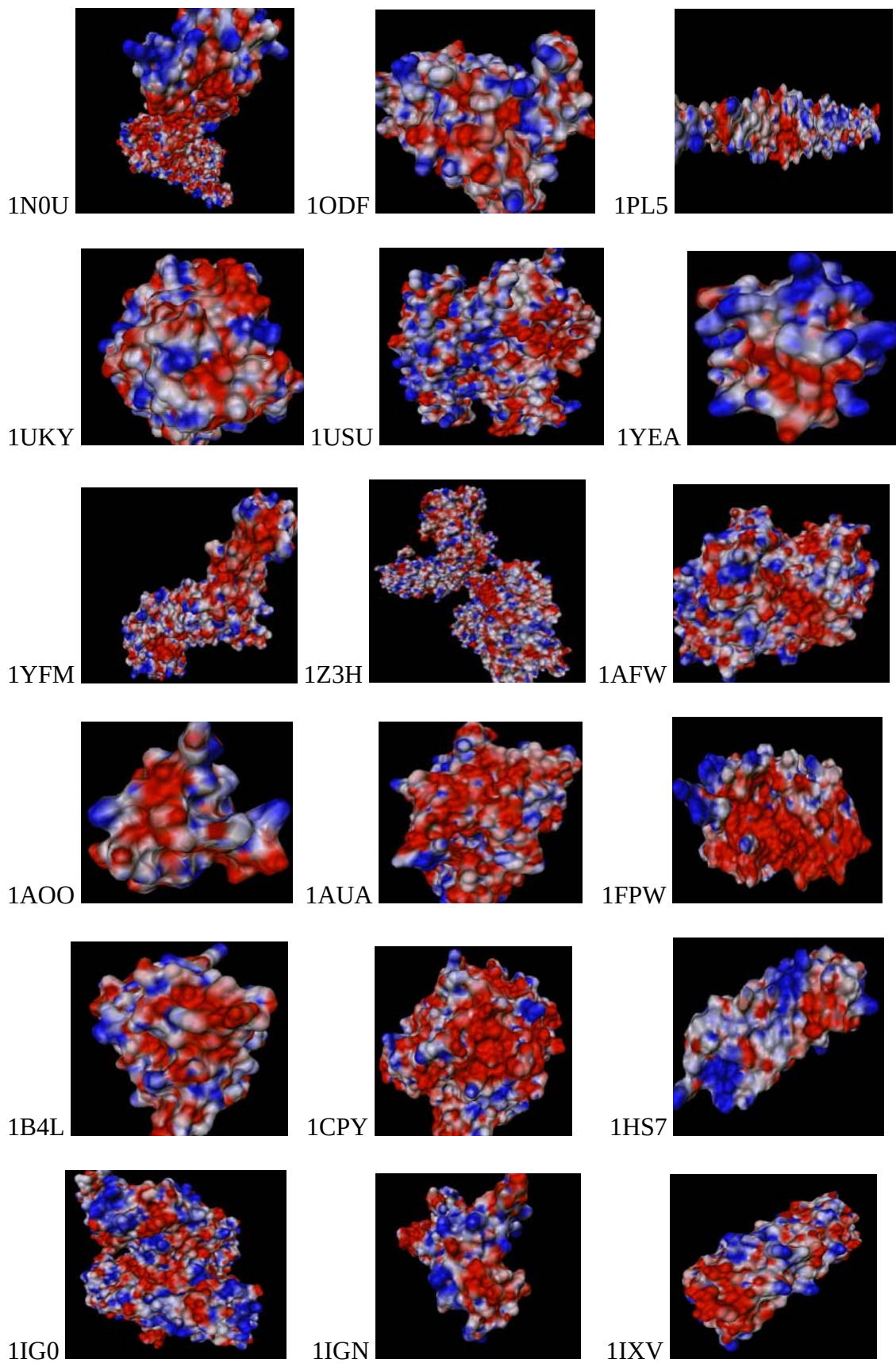
When possible, each crystal structure was selected to contain the most native-like conformation possessing the longest polypeptide chains with no ligands and/or mutations. Additionally, the highest possible resolution and R-factors were employed. The following table demonstrates more detailed information about each selected structure. The surface

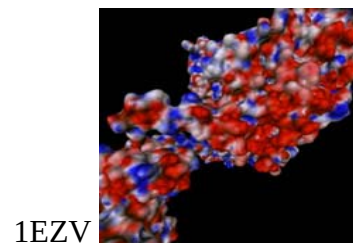
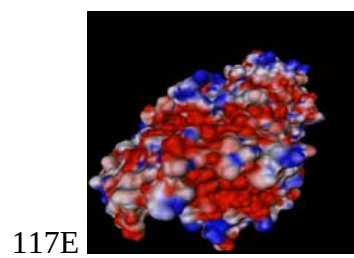
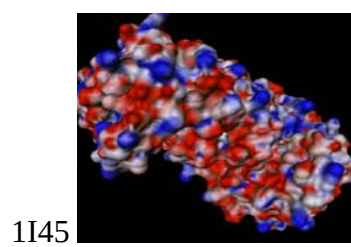
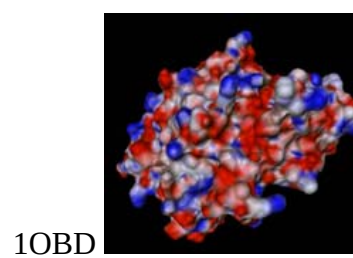
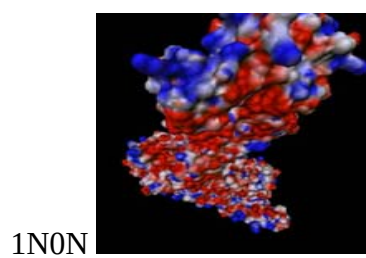
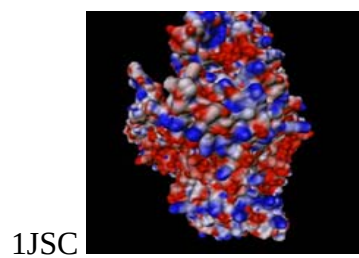
potential calculations were not successful for structures 22-24, due to the problems with the backbone structure deposited in PDB. Each resulting structure was visually inspected for the best representation of the polyanion binding site(s) and this image is represented here and in the body of the paper.

Number	PDB code	Protein Name	Notes
1	1A0A	PHOSPHATE SYSTEM POSITIVE REGULATORY PROTEIN PHO4/DNA COMPLEX	
2	1A4E	CATALASE A FROM SACCHAROMYCES CEREVISIAE	
3	1AKY	ADENYLATE KINASE FROM YEAST	Liganded to AP5A. Other structures are either mutated or liganded
4	1B54	YEAST HYPOTHETICAL PROTEIN-A	Selected by a Structural Genomics Approach, FROM BNL'S HUMAN PROTEOME PROJECT
5	1BK5	KARYOPHERIN ALPHA	
6	1DZF	RNA POLYMERASE SUBUNIT RPB5	
7	1EOV	FREE ASPARTYL-TRNA SYNTHETASE (ASPRS) (E.C. 6.1.1.12)	
8	1F60	ELONGATION FACTOR	Complex EEF1A:EEF1BA
9	1FUU	INITIATION FACTOR 4A	
10	1HOW	SKY1P, AN SR PROTEIN KINASE	
11	1KBI	Recombinant Flavin-binding Domain	1FCB was selected initially. Due to a Gly100B missing backbone, PCE calculation was not possible. 1KBI was selected

12	1M0T	Glutathione Synthase	
13	1MQS	Sly1p in complex with an N-terminal peptide of Sed5p	replaced MSE to MET required for PCE calculations
14	1NEX	ScSkp1-ScCdc4-CPD	
15	1NHY	GST-like Domain of Elongation Factor 1-gamma	replaced MSE to MET required for PCE calculations
16	1ODF	YGR205W PROTEIN	
17	1OF5	MEX67-MTR2	
18	1YAA	ASPARTATE AMINOTRANSFERASE	Cytosolic
19	1YAT	12-kDa FK-506-binding protein	Complexed with the immunosuppressive drug FK-506
20	1YGH	HAT DOMAIN OF GCN5 (transcriptional coactivator)	
21	1YTC	ISO-2 CYTOCHROME C	
22,23	1RYP	20S PROTEASOME	Same structure for two proteins, YGR135W (Proteasome component Y13) and YOR362C (Proteasome component C1)
23	1YGP	PHOSPHORYLATED FORM OF YEAST GLYCOGEN PHOSPHORYLASE (Phosphate bound in active site)	PCE problems in calculations due to a missing atom

A random selection of 26 proteins demonstrated a surface with more pronounced negative electrochemical potentials compared to the PABPs, as can be seen in the following images.





ST 8

This table lists the p -values estimated for the number of protein-protein interactions involving PABPs, showing the significance of the corresponding sub-network for each statistical PABP detection level (defined by the standard deviation).

<i>PABP dataset</i>	<i>Standard deviation</i>	<i>Number of PABPs</i>	<i>PABPs present in the interactome</i>	<i>p-value</i>	<i>95% confidence interval</i>	
Actin-binding	1	135	126	0.0829	±	0.0024
Actin-binding	2	63	62	0.0663	±	0.0022
Actin-binding	3	27	26	0.0491	±	0.0019
DNA-binding	1	86	81	0.0032	±	0.0005
DNA-binding	2	44	40	0.0149	±	0.0011
DNA-binding	3	29	27	0.0209	±	0.0013
Heparin-binding	1	264	240	0.0613	±	0.0021
Heparin-binding	2	107	100	0.0339	±	0.0016
Heparin-binding	3	59	56	0.0308	±	0.0015
HS-binding	1	348	325	0.0361	±	0.0016
HS-binding	2	124	110	0.0173	±	0.0011
HS-binding	3	66	62	0.0242	±	0.0013
Tubulin-binding	1	60	55	0.0565	±	0.0020
Tubulin-binding	2	13	13	0.0094	±	0.0008
Tubulin-binding	3	10	10	0.0311	±	0.0015
All	1	529	489	0.0635	±	0.0021
All	2	205	188	0.0091	±	0.0008
All	3	108	103	0.0282	±	0.0015
<i>Average</i>		<i>126.50</i>	<i>117.39</i>	<i>0.0355</i>	±	<i>0.0015</i>

References

- Jaques, L. B. (1979). "Heparins-anionic polyelectrolyte drugs." *Pharmacol Rev* **31**(2): 99-166.
- Jaques, L. B., Bruce-Mitford, M. and Ricker, A. G. (1947). "The metachromatic activity of heparin." *Rev. Can. de Biol.* **6**: 740-754.
- Templeton, D., M. (1988). "General occurrence of isobestic points in the metachromatic dye complexes of sulfated glycosaminoglycans." *Int. J. Biol. Macromol.* **10**: 131-136.
- von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S. and Bork, P. (2002). "Comparative assessment of large-scale data sets of protein-protein interactions." *Nature* **417**(6887): 399-403.
- Wollin, A. and Jaques, L. B. (1973). "Metachromasia: an explanation of the colour change produced in dyes by heparin and other substances." *Thromb Res* **2**: 377-382.