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,
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/μl) 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 polyanion and exposed to streptavidin-Alexa Fluor 647 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 μl of biotin hydrazide in dry DMSO and 12.5 μl 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 μl 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 μM 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 × 10⁴ (25) and 1.15 × 10⁴ m⁻¹ cm⁻¹, 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 × 10⁻⁵ M. One milliliter of the dye stock solution was titrated with 1–10 μl 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 μg/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-His₅ 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 μl of a 25–50 μg/ml concentration of each biotinylated polysaccharide 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-μm 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 (+1 S.D.). This +1 S.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, +2 S.D. and +3 S.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 +1 S.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 Aminos 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.yeastgenome.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 (QuickCalc, 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.

**pl Calculations**

The code for the program for theoretical calculations of pl values for yeast PABPs was obtained from the code author. This program was used to generate and compare the theoretical pl 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 Interg8 web portal (www.ebi.ac.uk/Interg8/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 again 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 disorderedness 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-sample 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 Blossom 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 Meiring 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 +3D.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).
The information provided in the table is extracted from the SGD at db.yeastgenome.org/cgi-bin/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                      | 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-1α   | 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                                 | Cytoplolic large ribosomal subunit (sensu Eukaryotae) (TAS) | Unknown                                  | Unknown                                     |
| YHL033C                       | Ribosomal protein L8A (rp6) (YL5) (L4A) | Unknown                           | 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 (IGI)                  | 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-RNA, 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 +1 S.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 +1 S.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,
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 (±1 S.D.) compared with 264 proteins at the level of ±1 S.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 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 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 ±1 S.D. level. | |
| Polyion | 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 μg/ml. | |
| NaCl (M) | 0 | 0.05 | 0.15 | 0.75 |
| Total proteins with positive interaction | 4247 | 4282 | 4276 | 4278 |
| Total proteins with significant interaction |
| ±1 S.D. | 254 | 261 | 264 | 258 |
| ±2 S.D. | 107 | 105 | 107 | 104 |
| ±3 S.D. | 60 | 58 | 59 | 60 |

**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 ±1 S.D. (A), ±2 S.D. (B), and ±3 S.D. (C).
TABLE IV

PABPs are enriched in basic amino acids

| Polyanion | Mean percentage of | Mean no. positive pitches |
|-----------|--------------------|--------------------------|
|           | Lys (experimental) | Arg (experimental) | His (experimental) | All positive residues | pl | pitches |
| 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   | 7.60               | 4.70                    | 2.18               | 14.48                    | 7.40 | 3.33 |
| Population|                    |                         |                    |                          |      |      |
| Difference|                    |                         |                    |                          |      |      |
P value    | 0.0081             | 0.067                   | 0.050               | 0.0137                    | 0.002 | 0.0079 |

*Indicates a significant difference (p < 0.05).

The number 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 pitches 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 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.

pl Calculations—If PABPs are presumably enriched in basic amino acids, their pl should demonstrate this fact. Thus, we investigated the pl values for PABPs as well as 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 pl 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-BXBX (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
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

| Table V | PABPs and the known polyanion binding domains |
|---------|-----------------------------------------------|
|         | DNA binding (423) | RNA binding (117) | ATP binding (369) | Actin binding (19) |
| 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 |

**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 surfaces phosphorylation could alter their polycationic nature and potentially create a regulatory mechanism for alteration of their interaction with the polyanion 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 polycationic 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 polycationic 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% disorderedness (at \(+1\text{S.D. level}\)). A value of 29% was found for non-PABPs. Although the difference does not appear large, it was significant with a \(p\) value of \(6 \times 10^{-5}\). These values at \(+2\text{S.D. and 3S.D. levels}\) 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.
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 present 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 ribosomal 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.
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(3H)-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 disorderedness 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 over-representation 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 Ausr, and Sangeeta Joshi at the University of Kansas for helpful discussions. Kim Bure, Cieta 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. Tunova, 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, 289–288
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. Rentzepis, 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., Matuszewski, B., Keoghan, 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., Madshus, I. H., Mach, H., Middaugh, C. R., and Olssen, S. (1990) Tight folding of acidic fibroblast growth factor prevents its translocation to the cytosol with diphertheria toxin as vector. EMBO J. 9, 4835–4842
10. Kuetzlo, 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. Kuetzlo, L. A., Normand, N., O’Hare, P., and Middaugh, C. R. (2000)
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. Spitze, J. J., and Poolman, B. (2005) Electrochemical structure of the crowded cytoplasm Trends Biochem. Sci 30, 526–541
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