Functional Assignment of Solute-Binding Proteins of ABC Transporters Using a Fluorescence-Based Thermal Shift Assay†

Sarah E. Giuliani, Ashley M. Frank, and Frank R. Collart*

Biosciences Division, Argonne National Laboratory, Lemont, Illinois 60439

Received August 29, 2008; Revised Manuscript Received October 29, 2008

ABSTRACT: We have used a fluorescence-based thermal shift (FTS) assay to identify amino acids that bind to solute-binding proteins in the bacterial ABC transporter family. The assay was validated with a set of six proteins with known binding specificity and was consistently able to map proteins with their known binding ligands. The assay also identified additional candidate binding ligands for several of the amino acid-binding proteins in the validation set. We extended this approach to additional targets and demonstrated the ability of the FTS assay to unambiguously identify preferential binding for several homologues of amino acid-binding proteins with known specificity and to functionally annotate proteins of unknown binding specificity. The assay is implemented in a microwell plate format and provides a rapid approach to validate an anticipated function or to screen proteins of unknown function. The ABC-type transporter family is ubiquitous and transports a variety of biological compounds, but the current annotation of the ligand-binding proteins is limited to mostly generic descriptions of function. The results illustrate the feasibility of the FTS assay to improve the functional annotation of binding proteins associated with ABC-type transporters and suggest this approach that can also be extended to other protein families.

Progress in sequencing technology has contributed to an exponential increase in the number of available genome sequences. The emergence of "next-generation" technology in recent years will continue this expansion as these machines are several orders of magnitude more efficient than the standard capillary machines that were used for the human genome project (*1, 2*). However, this ability to generate sequence has far outpaced computational and experimental methods to fully utilize the genomic data. The automated annotation process applied to genome sequence data often results in unresolved structural features and a limited ability to associate relevant biological information with the sequences (*3*−*5*). Estimates vary, but there is a general consensus that 50−70% of ORFs in newly sequenced genomes are described as having unknown or poorly characterized function (*4, 6, 7*). The magnitude of the problem will continue to grow as each new genome or metagenome generates additional uncharacterized "singletons" and novel protein families (*7*).

The assignment of functional annotation is clearly a challenge as proteins possess a diverse spectrum of functional characteristics, and the functional description for a specific protein is more properly visualized as a progression of states (*8*). Functional descriptions can vary from general to specific, and some proteins can have more than one function. However, more specific and detailed functional descriptions for proteins generally increase the scientific utility of the functional annotation. In this context, the extensive use of homology-based approaches for function prediction imposes some limitations on specificity and comprehensiveness of functional annotation. These approaches can provide a prediction of protein function for 50−70% of the ORFs in newly sequenced organisms (*7, 9*), but many of these predictions are very general in nature, and this approach cannot address novel protein function. Experimental approaches attempt to provide more specific functional characterization but are often expensive and limited in throughput. In the final analysis, a comprehensive approach for functional annotation will require the application of both computational and experimental methods to improve the extent of specific functional annotation for proteins.

In this study, we evaluate an experimental approach for improvement of functional annotation based on mapping of ligands with their binding proteins. Elucidation of protein−ligand interactions is an effective approach for functional interrogation of unknown proteins and can be applied to proteins with known binding ligands as a method to increase or extend biochemical/biophysical characterization. The NIH-sponsored Structural Genomic Programs have exemplified the utility of this approach for proteins of unknown function where discovery of protein−ligand pairs can lead to predictions for protein function. In the process leading to structure determination, ligands can copurify and crystallize with their binding partners, and analysis of the interaction can often provide clues to elucidate the functional role of the protein (*10*). Care must be exercised in this approach as many structures in the PDB have ligands but these are not cognate ligands (*11*). Other methods such as NMR (*12*), isothermal titration calorimetry (*13*), static light scattering (*14*), or affinity binding technologies (*15*) also use the property of

† The work was supported by a grant from the Office of Biological and Environmental Research, U.S. Department of Energy Office of Science to F.R.C.

* To whom correspondence should be addressed. Phone: (630) 252-4859. Fax: (630) 252-3387. E-mail: fcollart@anl.gov.

10.1021/bi801648r CCC: $40.75
Published on Web 12/08/2008

ligand binding as a route to identify or supplement protein functional annotation.

Our approach for the identification of bound ligands and assignment of function uses a fluorescence-based thermal shift (FTS)[1] assay for high-throughput screening of protein−ligand interactions. Protein melting profiles can be determined by other methods such as differential scanning calorimetry (*16*), optical methods such as CD (*17*), or intrinsic tryptophan fluorescence (*18*). However, these approaches are generally low throughput and not applicable to all proteins. The FTS approach is a target-independent assay that uses a fluorescent dye (SYPRO orange (*19*)) to monitor protein unfolding. The dye is initially quenched in the aqueous environment with the folded protein. As temperature is increased, protein unfolding exposes hydrophobic regions and the dye becomes unquenched, resulting in a large increase in fluorescence. This approach has broad applications for assessment of ligand binding (*19*−*21*) and as a HTP screen for stability optimization of proteins for structural studies (*19, 22*). This assay uses a commercially available real-time PCR instrument (*23*), where thermal melting curves of the protein/ligand combinations can be screened in a 96-well plate format.

To validate the ability of the FTS assay to map ligands with binding proteins and to determine the utility for annotation of proteins with unknown function, we focused on ligand-binding proteins from the ABC-type transporter family. These types of transport systems are widely distributed in all three kingdoms of life and can transport a variety of substrates such as metals, small ions, mono- and oligosaccharides, peptides, amino acids, iron siderophores, polyamines, and vitamins (*24*−*26*). Approximately 5% of the *Escherichia coli* genome encodes components of the ABC-type transporters (*27*). The ABC-type transporter family is diverse, but the transporters are organized in specific groups on the basis of phylogeny and function (*28*−*30*). From a practical perspective, this facilitates a ligand screening approach for functional annotation as some of the groups are characterized by the general chemical character of the binding ligands. The extracellular solute ABC transporters in bacteria are typically comprised of three component proteins: a permease, ATPase, and ligand-binding protein. In many cases, the transporter is organized in the genome as an operon encoding the component proteins, enabling the annotation to be assigned to the set of proteins.

Our FTS assay used the solute-binding protein components of the transporter to screen for potential protein/ligand interactions. In Gram-negative bacteria, these proteins are typically localized in the periplasmic space while in Gram-positive bacteria they are processed as an extracellular lipoprotein. Many of these binding proteins are able to concentrate extracellular solutes when concentration of the compound is below the micromolar range. Although most bacteria contain 20−100 genes encoding these binding proteins, the specific binding ligand is unknown for >80% of these proteins. Most of these are annotated as proteins of unknown function or generally annotated as extracellular transport proteins. The ability to provide biologically significant and specific annotation would provide insight into the functional diversity of this important class of proteins. When applied across a particular organism, the ability to decipher its transport capability would prove invaluable toward elucidating the functional and metabolic potential and contribute to a system-level understanding of the organism.

EXPERIMENTAL PROCEDURES

Target Staging. Coding sequences were extracted from the NCBI database, and the protein sequences were characterized using the SignalP (*9, 10*) and TMHMM (*31*) algorithms to identify hydrophobic sequence features and guide the selection of clonable regions. The peptides specified by cloned target sequences excluded predicted N-terminal signal peptide sequences and any overlapping N-terminal transmembrane helices. Predicted N-terminal cleavage sites were manually adjusted for input into an automated primer design tool (http://tools.bio.anl.gov:8000/bioJAVA/jsp/PeriplasmicProteinDesign/index.html) that uses the predicted cleavage information derived from SignalP 3.0 to automatically generate expression primers for the mature (cleaved) form of the protein.

Gene Cloning and Protein Expression and Purification. The *E. coli* cloning and expression strategy employed 96-well plate-based methods and parallel expression strategies (*32*). Target genes were PCR amplified from genomic DNA (ATCC: *E. coli* K-12 no. 10798D-5, *Salmonella typhimurium* LT2 no. 700720D-5, *Campylobacter jejuni* NCTC11168 no. 7001819D, *Shewanella oneidensis* MR-1 no. 700550D) using a KOD HiFi DNA polymerase reaction (Novagen) combined with a touchdown PCR program designed for the GeneAmp PCR System 9700 thermocycler machine (Applied Biosystems). Amplification products included appended 5′ and 3′ ligation independent cloning (LIC) sites to enable simultaneous cloning in multiple vectors. To optimize protein solubility outcome, each gene was cloned in parallel by a LIC method (*33*) into two cytoplasmic expression vectors, pMCSG7 and pMCSG19c (*34*). Both vectors append an N-terminal hexahistidine (6×-His) fusion tag and TEV protease cleavage recognition sequence between the fusion tag and the target protein. Additionally, pMCSG19c codes for an N-terminal MBP fusion protein followed by a TVMV protease cleavage site upstream of the 6×-His tag (*34*). Each target was characterized for amplification, expression, and solubility using 96-well plate assays and high-density gel formats for denaturing gel analysis of proteins (*35*). Targets were scored as positive for expression and solubility if a detectable fusion protein of the correct molecular weight was observed on gels stained with Coomassie-based Simply Blue Safestain (Invitrogen). Targets scored as positive for solubility were sequence verified prior to purification and screening.

Clones expressing soluble proteins were scaled to 500 mL cultures and purified using standard affinity chromatography Ni-NTA bead (Qiagen) purification methods on an AKTA system (GE-Healthcare) as previously described (*36*). To remove the 6×-His fusion tag, proteins were incubated with TEV protease (0.05 mg of protease/mg of target protein) at 4 °C for 3 days on a rotator. The TEV protease also contains a 6×-His fusion tag, and the protease and target His-tag fusion were removed by adsorption on Ni-NTA beads in a

[1] Abbreviations: FTS, fluorescence-based thermal shift; COG, cluster of orthologous groups; LAO transporter, lysine/arginine/ornithine transporter; PBPs, periplasmic binding proteins; LIV protein, leucine/isoleucine/valine binding protein; LIC, ligation independent cloning.

reverse affinity chromatography procedure. The purified proteins were dialyzed for buffer exchange into an assay-compatible buffer, flash frozen in liquid nitrogen (*37*), and stored at −80 °C until the assay. Protein concentrations were initially determined by measuring absorbance at λ_{280} using UV spectrophotometry and verified by comparison with a BSA standard using SDS−PAGE denaturing gel electrophoresis and Simply Blue Safestain for protein visualization. Once thawed, proteins were stored at 4 °C and used within 2 weeks.

Ligand Reagent Stock Preparation. All ligands were dissolved in standard assay buffer (100 mM HEPES, 150 mM NaCl, pH 7.5), and solutions were stored at 4 °C. The cysteine stock solution contained 1 μM DTT to prevent oxidation during the assay.

Standard Assay Parameters. FTS assays were performed using an Mx4000 multiplex quantitative PCR instrument (Stratagene) that enabled thermal manipulations and dye fluorescence detection based on a previously published method (*38*). An environmentally sensitive dye, SYPRO orange (Invitrogen, no. S6650), was used at 5× concentration in all assays. Custom filter configurations corresponded to the optimal excitation (492 nm) and emission (610 nm) wavelengths for SYPRO orange dye. Detector gain was adjusted to 800−900 to obtain an optimal measurement of dye fluorescence, and the Mx4000 instrument's predefined SYBR green experimental method was chosen to enable software calculation of the first derivative values ($-R'(t)$) for the denaturation curve raw data points (R). Data were collected as three individual end point readings for each 1 min cycle. The temperature was increased by 1 °C each cycle over a temperature range of 25−95 °C. Individual raw data points displayed for each temperature represented the average of the three end point readings. Parameters related to automatic data analysis for qPCR-related measurements were standardized for all assays since their settings did not directly apply to raw thermal shift data. Default options used for the assay included the threshold fluorescence, baseline correction, and the smoothing option for graphical temperature resolution. Replicates were treated collectively for calculation and display of the average values for replicate data points. The FTS assay was run using the slow scan rate setting for the Mx4000 instrument. Assay reactions were performed in 96-well white PCR plates (Axygen, no. PCR-96-FLT-W), and wells were capped using optical 8×-strip caps (Axygen). All reported T_m values resulted from manual subtraction of dye background fluorescence from raw data fluorescence values.

Protein/Ligand Binding Assays: Individual Ligand Screens. The amino acid ligand set selection was based on the structural data for protein PDB models with bound native amino acids and was only limited by reagent solubility in reaction buffer at concentrations specified for the assay. The individual ligand screening set contained all of the amino acids except Phe, Trp, and Tyr as well as ornithine. The amino acids Phe and Trp were originally excluded from the standard assay mix but were incorporated into the pooled ligand screening protocols. Tyr was not included in any of the screens due to solubility constraints. All reagents were obtained from Sigma-Aldrich except histidine, which was obtained from Novabiochem. Final amino acid concentrations in the assay were 20, 200, and 1000 μM. Protein concentrations were standardized to 10 μM based on preliminary screens to optimize fluorescence detection of proteins over a range of concentrations from 2 to 50 μM. All protein and SYPRO orange dilutions were performed such that final assay buffer concentrations were 100 mM HEPES and 150 mM NaCl, pH 7.5. All thermal shift reactions were performed in triplicate in a volume of 40 μL. For each plate screened, triplicate reactions were performed for buffer and SYPRO orange dye negative controls. The variability associated with each data point derived from averaged data of triplicate reactions was consistently ±1 °C. For each plate assay, proteins at one concentration were screened against ligands each at three different concentrations, and reactions were performed in triplicate. For a positive control, the protein was screened with its respective known binding ligand for targets based on PDB models.

A series of metal ligands [sodium salts of tungstate, molybdate, and vanadate and chloride salts of cobalt(II), nickel(II), manganese(II), and iron(III)] were used to screen a selected set of proteins from *S. oneidensis*. Ligands were screened at 1000 μM final concentrations, as this was the maximum concentration for optimal detection of binding. All other assay parameters for buffer, dye, and protein were identical to those mentioned above, except final reaction volumes were 20 μL. Reagents were obtained from Sigma-Aldrich except molybdate (Mallinckrodt) and iron(III) (Fisher Healthcare).

Protein/Ligand Binding Assays: Pooled Ligand Screens. Pooled ligand libraries contained all of the amino acids except Tyr. Pools were designed to test how many ligands could be included in a pool without interfering with the fluorescence in the assay and to test binding specificity. Similar amino acids were grouped into smaller pools by two, three, or six; then these pools were combined into larger pools, including a final pool with all amino acids. For each pool, one ligand was included as a positive control which corresponded to a PDB model protein or homologue (Met) bound with that ligand. Thus, all pools were tested in the screening of all positive control proteins. Positive control ligands were as follows: pool 1, *methionine*; pool 2, *leucine*; pool 3, *cysteine*; pool 4, *glutamine*; pool 5, *arginine*; pool 6, *histidine*. Ligand concentrations were 0, 200, and 1000 μM in order to determine protein-binding dependence on ligand concentration. Since most T_m shifts were not detected significantly below 200 μM concentrations based on individual screens, the lowest concentration was not tested for ligand pools. Protein concentrations were standardized to 10 μM, and reaction volume was decreased to 20 μL. In each plate screened, each reaction was performed only once, but duplicate reactions were performed for buffer and SYPRO orange dye negative controls. For each plate assay, proteins at one concentration were screened with no ligand, against the expected single ligand at two different concentrations as a positive control, and against 10 ligand pools, each at two different concentrations, for a total of 23 reactions per protein screened. Thus, four different proteins were screened with specified ligand combinations in one plate with negative controls.

Protein/Ligand Binding Assay Screening Various Concentrations of Protein with One Concentration of Individual Ligands. To assess the effect of protein concentration on ligand binding, target proteins were screened each at three different concentrations of 2 and 10 μM for all proteins and

Table 1: FTS Assay Results for Amino Acid-Binding ABC-Type Transporter Proteins Annotated from PDB Structures

description	source organism	accession ID	PDB ID	T_m (°C) (no ligand)	ligand(s) bound in PDB	binding ligand(s)	T_m shift (°C) 1000 μM	200 μM	20 μM
cysteine-binding protein	*C. jejuni*	YP_179057	1xt8	64.5	Cys, Ser	Cys	12.0	9.0	6.0
						Ser	2.0	0.0	0.0
glutamine-binding protein	*E. coli*	NP_415332	1wdn	60.6	Gln	Gln	8.0	0.0	0.0
histidine-binding protein	*E. coli*	NP_416812	1hsl	52.5	His	His	5.0	2.0	0.0
						Arg	2.0	0.0	0.0
histidine-binding protein	*S. typhimurium*	NP_461296	1hpb	52.5	His	His	6.0	4.0	0.0
lysine-, arginine-, ornithine-binding protein	*S. typhimurium*	NP_461297	1lst	50.0	Arg, Orn, Lys	Arg	9.0	6.0	0.0
						Orn	8.0	5.0	0.0
						Lys	6.0	5.0	1.0
						Cys	5.0	5.0	0.0
leucine transporter subunit	*E. coli*	NP_417915	1usk	51.5	Leu, Ile, Met, Val	Leu	14.0	12.0	6.0
						Ile	6.0	3.0	1.0
						Met	5.0	3.0	1.0
						Val	3.0	1.0	0.0

50 μM for most proteins, both with no ligand and with 1000 μM ligand. A few proteins were screened at 30−40 μM for highest concentrations, as limited by purified stock concentration. As positive controls, only expected binding ligands were screened. Unique reaction combinations were performed only once along with triplicate reactions for negative controls. Reaction volumes were 40 μL, and all other assay conditions were standard. All proteins screened, except the LAO-binding protein from *S. typhimurium,* were 6×-His fusion tag versions.

RESULTS

Target Selection for Assay Validation. Validation targets were selected from the bacterial ABC transporter family and specifically the amino acid solute-binding proteins of the transporter complex. These proteins have a conserved structure generally consisting of two domains connected by a hinge region (*24, 39*). The two-domain structure is described as a Venus flytrap in reference to the method for ligand binding where the two domains close around the ligand (*24*). A group of six periplasmic binding proteins (PBPs) with known ligand-binding specificity were selected to evaluate the specificity/utility of the fluorescent thermal shift assay for functional assignment. The six validation targets were selected from structures deposited in the PDB and had been crystallized both with and without bound ligand(s) (Table 1). This group included four proteins with specificity restricted to one or two amino acid ligands and two proteins with broad ligand specificity for hydrophobic or positively charged amino acids. In most cases, additional data from genetic or biochemical approaches were available to support assignment of binding specificity.

Fluorescence-Based Thermal Shift Assay Results for Known Ligand Binding Proteins. The structure of a periplasmic binding protein from *C. jejuni* was determined with bound cysteine and demonstrated to be a cysteine-binding protein by fluorescence titration experiments (*40*). In the FTS assay, the melting temperatures (T_m) of this cysteine-binding protein in the absence and presence of ligand are reflected as the midpoint of the rise in each curve (Figure 1A). The native protein in the absence of ligand displays a temperature midpoint (T_m) for the unfolding transition of approximately 67.5 °C (Figure 1A). The shift in the T_m as defined by the relative midpoint of the denaturation profiles suggests that the thermal stability is dependent on ligand concentration. For data analysis, the T_m for each unique sample (average of triplicate reactions) was determined as the minimum peak in the negative first derivative ($-R'(t)$) plot of the melting curve raw fluorescence data (R) (Figure 1B). The negative derivative curve of the denaturation data provides a more defined illustration of the T_m differences between ligand concentrations. This approach more effectively illustrates the concentration-dependent increase in the thermal stability observed upon addition of ligand (compare panels A and B of Figure 1). This assay structure and data analysis procedure was reproduced for additional amino acids using a 96-well plate format. In this assay format, each well contained a maximum of a single ligand. The outcome of a screen for each protein can be represented as a histogram of the net change in the T_m relative to the protein without ligand (ΔT_m) for each unique assay sample (Figure 1C). The data in the figure represent the assay outcome for the highest concentration of ligand (1000 μM) as this provided the greatest level of discrimination. In the case of the cysteine-binding protein, a significant increase in the T_m was only observed for cysteine, the specific binding ligand for this protein (Figure 1C). Changes in the T_m of less than 2 °C were observed for all other amino acids.

A similar approach was applied to the other five PBPs with known ligand-binding specificity. In the selection of these targets, we attempted to maximize the functional and biochemical diversity of the targets and thereby obtain a practical assessment of the general utility of the assay for functional characterization of the amino acid-binding protein family. The native proteins (less signal sequence for periplasmic localization) varied in molecular mass from 25 to 36 kDa and showed an ~15 °C difference in the calculated midpoint T_m for the unfolding transition (Table 1). In the evaluation of the ligand binding, each of these targets was screened in triplicate using the screening set of amino acid ligands listed in the Experimental Procedures section. The FTS assay identified specific binding ligands for each protein in the validation set (Table 1). We considered an amino acid a "binding ligand" if the calculated ΔT_m (relative to the T_m without ligand) was 2 °C or greater for an amino acid

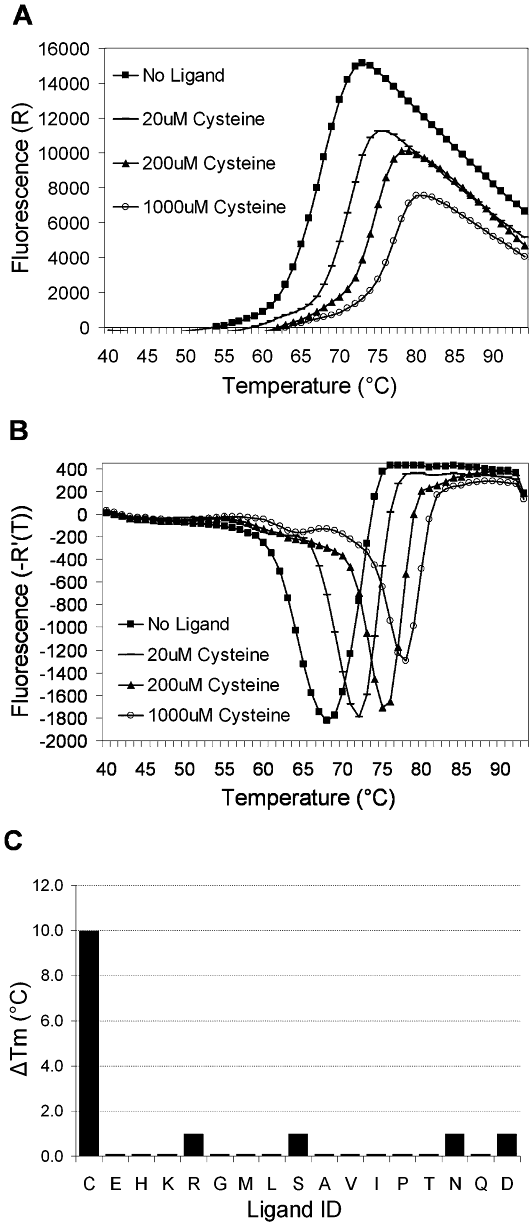

FIGURE 1: Generation of ligand-binding profiles for a cysteine-binding protein from *C. jejuni* using the FTS assay. (A) Thermal denaturation curve for the *C. jejuni* protein with 20, 200, and 1000 μM cysteine. Proteins and protein/ligand combinations were screened in triplicate and the data points averaged to generate individual thermal melt curves for data comparison. (B) Negative derivative curve of the denaturation data in (A). (C) Histogram display of the screening results with the amino acid library. The bars represent the difference in calculated T_m for the ligand plus protein samples relative to the positive control protein. Shifts of 1 °C for some amino acids within the replicate variability range of $\pm$1 °C are not considered significant. Data where T_m shifts were 0 °C are shown as 0.1 °C to indicate ligands were tested.

concentration of 1000 μM. The 2 °C ΔT_m significance threshold was selected based on the screening outcome for the validation set of targets as well as a qualitative assessment of the observed variation for proteins between the various assays (individual ligand screens, pooled screens, and replicates). Using this criterion, all of the ligands identified as specific binding ligands displayed a concentration-dependent increase in the T_m when incubated with the target protein. The ligand-binding specificity determined by the FTS assay corresponds extremely well with the functional description derived from the crystallographic studies. Essentially, all of the ligands identified in the crystallographic studies (and complementary experimental approaches) were also predicted by the thermal shift assay. Two of the proteins in the validation set (Lys/Arg/Orn-binding protein and leucine transporter subunit) were known to exhibit broad substrate specificity and bind a variety of chemically similar ligands. For both of these PBPs, the principle binding ligands characterized by other experimental approaches were also identified using the FTS approach. In several cases, additional binding ligands were identified using the FTS approach. This is perhaps not unexpected since previous experimental studies were not designed to evaluate all amino acid-binding ligands and there is no evidence that these were included as part of the structure determination process. However, other published experimental data provide support for the assignment of the additional binding ligands observed in this study (reviewed in the Discussion section). Overall, the functional assignments obtained by the FTS assay are consistent with the generally accepted experimental functional assignments.

Fluorescence-Based Thermal Shift Assay Results for Unknown Ligand Binding Proteins. We extended our assessment of the FTS assay for functional annotation by evaluating binding ligands for a second set of targets. The targets in this evaluation set had various degrees of experimental functional characterization (Table 2), but none of the individual proteins were represented as structures in the PDB. Three proteins in this group have specific functional assignments while the remaining six have general functional descriptions limited to a cluster of orthologous groups (COG) classification. The ArtJ PBP of the ABC transporter from *E. coli* was demonstrated to bind arginine by equilibrium dialysis and gel filtration analysis using [^{14}C]-L-arginine (*41*). These analyses examined all amino acids and the amines: putrescine, citrulline, and ornithine. The data demonstrated specific binding only for arginine with a K_d ~0.4 μM determined by equilibrium dialysis. These results were confirmed by the FTS assay (Table 2) with specific binding observed only for arginine. In a parallel experiment, the periplasmic binding protein of the *metD* locus in *E. coli* had been suggested to encode an ABC transporter for methionine uptake (*42*). This functional assignment is based on metabolite competition experiments (*43*), genetic deletions (*43*), and analysis of transcription regulation (*42*). The periplasmic binding protein (MetQ) of this operon was screened against our amino acid library using the FTS assay and displayed specific binding to methionine (Table 2) with no evidence of binding to other ligands in the screening library. The PBP for the lysine/arginine/ornithine (LAO) transporter from *E. coli* is highly similar to the PBP from *S. typhimurium* from the validation target set (Tables 1 and 2) and is annotated as a LAO-binding protein. The protein from *E. coli* had a slightly lower T_m than the protein in the validation set (Table 1) but exhibited an identical ligand-binding profile (Tables 1 and 2) for the characterized ligands. Cysteine was not identified as a binding ligand for the *E. coli* LAO transporter since the calculated ΔT_m was less than 2 °C at the highest amino acid concentration.

The set of targets from *S. oneidensis* in this group is clustered into two COGs. Five of the *Shewanella* proteins are categorized in COG083ET, a cluster which contains amino acid-binding and signal transduction proteins. These

Table 2: FTS Assay Results for Solute-Binding ABC-Type Transporter Proteins Varying in Degree of Experimental Functional Annotation

					T_m shift (°C)		
description	source organism	accession ID	T_m (°C) (no ligand)	binding ligand(s)	1000 μM	200 μM	20 μM
arginine transporter subunit	*E. coli*	NP_415381	59.6	Arg	8.0	5.0	0.0
DL-methionine transporter subunit	*E. coli*	NP_414739	68.5	Met	6.0	4.0	1.0
lysine-, arginine-, ornithine-binding protein	*E. coli*	NP_416813	49.5	Arg	9.0	6.0	3.0
				Orn	7.0	4.0	1.0
				Lys	6.0	4.0	0.0
COG0834ET protein	*S. oneidensis MR1*	NP_716672	52.5	Arg	11.0	7.0	1.0
COG0834ET protein	*S. oneidensis MR1*	NP_717120	58.6	none			
COG0834ET protein	*S. oneidensis MR1*	NP_717447	65.5	none			
COG0834ET protein	*S. oneidensis MR1*	NP_719392	57.6	none			
COG0834ET protein	*S. oneidensis MR1*	NP_716883	53.5	none			
COG2998H protein	*S. oneidensis MR1*	NP_720235	51.5	tungstate	22.0		
				molybdate	5.0		

five proteins were annotated as periplasmic binding proteins but only one (NP_716672) was specifically annotated as an amino acid-binding protein. The remaining protein is grouped in COG2998H, a cluster with the description of "ABC-type tungstate transport system, permease component", but is annotated as a hypothetical protein (NCBI database). However, this protein is part of an operon containing two proteins annotated as an "ABC transporter permease protein" and "ABC transporter, ATPase protein". The set of targets from *S. oneidensis* was screened for binding using the amino acid library and a series of metal salts. Specific binding to arginine was detected for one of the proteins (NP_716672) which was originally annotated as an amino acid-binding protein (Table 2). For the protein assigned to the tungstate transport COG (NP_720235), the addition of tungstate or molybdate salts resulted in a ΔT_m of 22 and 5 °C, respectively, in the calculated midpoint T_m relative to the control protein, suggesting that this protein is a component of a tungstate transport system.

Pooled Ligand Approach for Assessment of Candidate Ligands. The diversity of ligands bound by ABC transporters and the inability of current bioinformatic approaches to predict specific ligand binding are significant constraints to high-throughput utilization of the current FTS approach for functional annotation. To improve the screening throughput, we evaluated the feasibility of a pooled ligand approach for assessment of ligand binding. Our strategy was to conduct a primary screen using ligand pools to identify candidate binding ligands from various ligand classes. The specific ligand binding could then be determined in a set of secondary screens that employed individual ligands representative of the respective pool(s) that elicited a thermal shift in the primary screen.

To evaluate this strategy, we rescreened several of the PBPs whose binding ligands had been confirmed through the FTS assay, using the assigned binding ligand and various pools of amino acids. The amino acid pools were organized so that pools 1−6 represented unique combinations of all the amino acids plus ornithine. Pools 7−9 contain a combination of two single pools while pool 10 contained the complete amino acid screening set plus ornithine (Table 3). The assay outcome is represented as the temperature shift of the target ligand or ligand pool relative to the protein with no ligands (Figure 2). Each assay included a positive control comprising the target protein with the amino acid-binding ligand identified from the individual ligand screens (sample on the far left of each bar graph in Figure 2). Comparison of the T_m shifts of the positive control with the ligand pools indicates that the largest shifts were associated with pools containing the specific binding ligand. These T_m shifts observed with the ligand pools were similar to those observed when only the known binding ligand was present. In the case of the methionine-binding protein (Figure 2A), the observed T_m shift confirms that the protein is stabilized by its annotated ligand, methionine ($\Delta T_m = 8.0$ °C). The thermal shifts with ligand pools containing methionine were consistent with this 8.0 °C shift (pool 1, $\Delta T_m = 7.0$ °C; pool 7, $\Delta T_m = 8.0$ °C; pool 10, $\Delta T_m = 8.0$ °C), indicating that pooled ligand screening provides a valid method for dense ligand assays. The thermal shifts of known histidine-binding (B) and lysine/arginine/ornithine-binding (C) proteins from *E. coli* provide further evidence that pooled screens offer comparable data to individual ligand screening, as the difference in thermal shift between individual binding ligand and pool shifts was consistently less than 1 °C. For these proteins, we observed some decrease in T_m shift relative to the control protein, but in most cases this shift was within the experimental variability for the assay. The exception was a 4 °C decrease in the T_m relative to the control observed for the His-binding protein incubated with pool 4 (Figure 2B) which contained asparagine and glutamine. This decrease was not observed for pool 8 which also contained these amino acids in addition

Table 3: Composition of Amino Acid Pools for FTS Assay Pooled-Ligand Screens

pool no.	ligands
1	Gly, Ala, Met, Trp, Phe, Pro
2	Leu, Ile Val
3	Ser, Thr, Cys
4	Asn, Gln
5	Lys, Arg, Orn
6	His, Asp, Glu
7	Gly, Ala, Met, Trp, Phe, Pro, Leu, Ile, Val
8	Ser, Thr, Cys, Asn, Gln
9	Lys, Arg, Orn, His, Asp, Glu
10	all amino acids + Orn

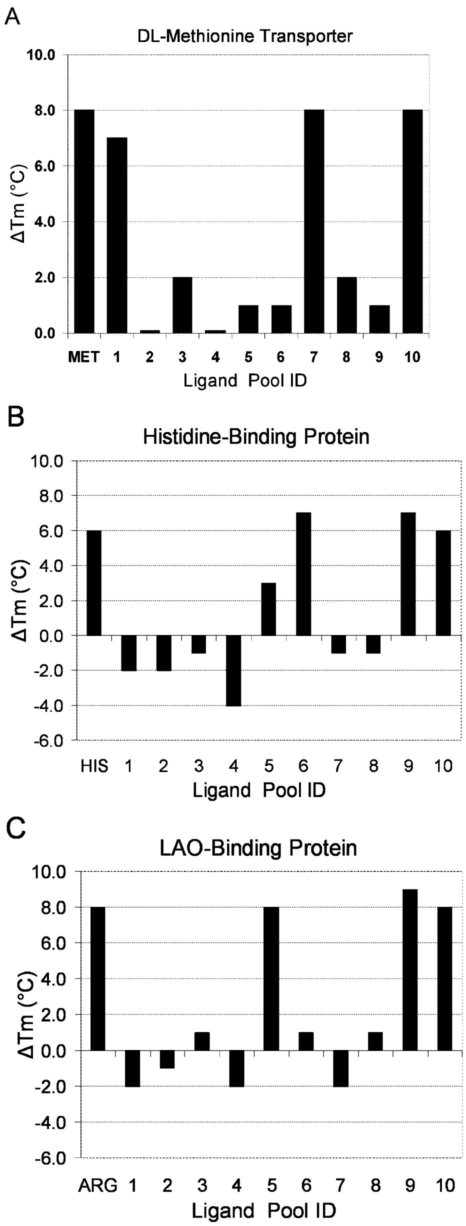

FIGURE 2: Histogram representation of outcomes for positive control proteins screened with known binding ligands and with ligand pools. In each assay the protein was independently screened with its expected binding ligand (1000 μM) and with the amino acid pools. The representative proteins were the (A) DL-methionine transporter protein from *E. coli*, (B) histidine-binding protein from *E. coli*, and (C) lysine/arginine/ornithine-binding protein from *E. coli*.

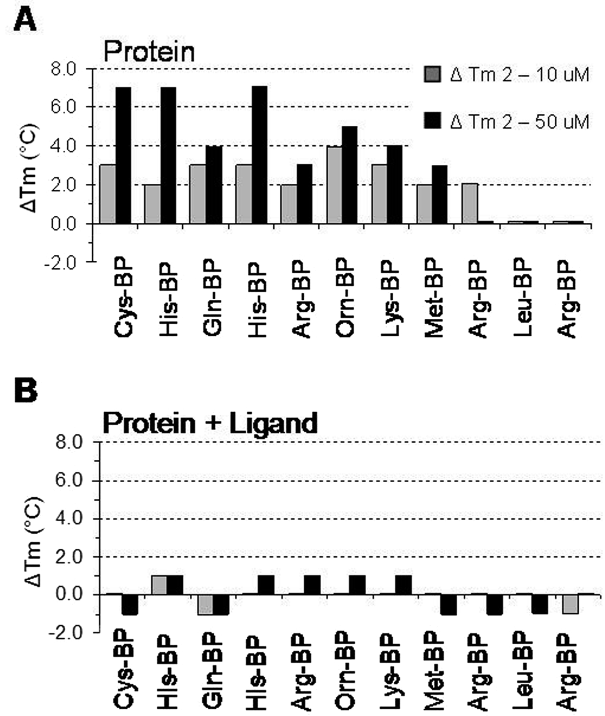

FIGURE 3: Ligand and protein concentration dependence of the calculated midpoint T_m. Protein T_m values were obtained from amino acid-binding proteins screened over a range of protein concentrations, 2−50 μM, without ligand (A) and with 1000 μM ligand (B). The bars representing the difference in T_m values (ΔT_m, °C) between 10 and 2 μM protein (gray bars) and between highest concentrations (27−50 μM) and 2 μM protein (black bars) are shown.

to serine, threonine, and cysteine. This pooled ligand screen was applied to several other PBPs listed in Tables 1 and 2 and in each case provided results consistent with the individual ligand screening approach. These results indicate that a pooled ligand approach for assessment of ligand binding is a viable initial strategy for identification of ligand candidates for validation by additional screening methods.

Characteristics of the Thermal Shift Assay. Implementation of the FTS approach as a HTP screen for functional annotation requires selection of a uniform set of assay conditions. To assess intrinsic variability in the melting temperature, we surveyed the dependence of T_m on the ligand and protein concentration for a selected set of proteins. The calculated midpoint T_m for the unfolding transition of the protein without ligand spanned a range of over 20 °C (Tables 1 and 2). This calculated midpoint T_m was concentration dependent for approximately half of the target proteins in this study (Figure 3). Using the midpoint T_m calculated at 2 μM protein as a reference, shifts of up to 7 °C were observed for some of the target proteins. However, there was not an obvious correlation of the absolute value of the ΔT_m with the melting temperature of the native protein. The native form of the cysteine-binding protein has a calculated midpoint T_m of 63.5 °C at 2 μM. A 7 °C increase in the calculated midpoint T_m was observed for this cysteine-binding protein at 50 versus 2 μM. The two histidine-binding proteins have a calculated midpoint T_m of ~49 °C at 2 μM but also show a 7 °C increase in the calculated midpoint T_m between the lowest and highest concentrations. This is in contrast to the lack of effect observed for the same proteins with the binding ligand present. In this case, only minimal shifts were observed. The data suggest that, in the presence of ligand, the characteristic thermal shift is largely independent of protein concentration. These observations are consistent with other studies which suggest the occurrence of multiple, but related open forms (*44*−*46*) which are in equilibrium with a closed form(s) of protein not bound with ligand.

Many heterologous protein production strategies utilize fusion tags to facilitate purification of expressed proteins for *in vitro* analysis with the 6×-His tag commonly employed (*47*). Since this approach was used to prepare proteins for the FTS study, we compared assay outcomes for four targets with tagged and untagged proteins. A total of five confirmed binding ligands were screened, three of which were ligands for the lysine/arginine/ornithine-binding proteins from *E. coli* and *S. typhimurium*. The remaining two were cysteine and glutamine (Tables 1 and 2). Proteins for this study were purified, and the His tag was removed from a portion of the purified sample by cleavage with TEV protease as described in the Experimental Procedures section. For both tagged and untagged versions, the proteins were screened at 10 μM against ligands of various concentrations. A slight decrease in the calculated midpoint T_m (no more than 2 °C) was noted for all tagged proteins without ligand, with the exception of the LAO protein from *E. coli*, which exhibited a T_m increase of 1 °C (data not shown). No significant changes were observed in the T_m between the tagged and untagged protein versions when samples were incubated with their binding ligands. The results indicate the assignment of the binding ligands and relative degree of the shifts were not significantly different for tagged vs native proteins.

DISCUSSION

This study evaluated the utility of the FTS assay as an experimental approach for specific functional characterization of proteins. The assay has been utilized in several variations as a rapid and inexpensive approach to assess changes in protein thermal stability associated with ligand binding (*21, 23, 38, 48*). This preliminary study focused on solute-binding proteins in the bacterial ABC transporter family. These transporters are essential membrane transport components in many organisms and transport a diverse range of ligands, but a specific functional role has not been assigned for a majority of these proteins in most organisms. The availability of several structures with bound ligand enables a rigorous assessment of the ability of the FTS approach to identify binding ligands for this class of proteins. Our validation set was representative of the most common bacterial amino acid-binding proteins and included proteins with binding specificity limited to a single amino acid and proteins with the ability to bind multiple amino acids. For the positive control group of amino acid-binding proteins (Table 1), the FTS assay was consistently able to map proteins to their known binding ligands. Similarly, the FTS assay was used to unambiguously identify preferential binding for several homologues of amino acid-binding proteins with known specificity and to functionally annotate two proteins of unknown binding specificity. An excellent correlation with previous ligand-binding assignments was observed for those PBPs binding only one or two ligands. The only exception was the case of the two histidine-binding proteins where the binding of arginine was detected for the protein from *E. coli* but not for the *S. typhimurium* protein. The *S. typhimurium* protein was screened using a pooled ligand strategy similar to that of the *E. coli* protein (Figure 2B) and demonstrated a temperature shift of 5 °C for a pool (pool 5) with arginine but without histidine (data not shown); however, the individual amino acid screens did not indicate arginine binding and were not repeated at this time. The inability to reliably detect arginine binding for this protein may be due in part to the lower affinity for arginine. The *S. typhimurium* protein has been demonstrated to bind histidine with a dissociation constant of 0.03 μM relative to a value of 0.7 μM for arginine (*49*). Additionally, we also cannot rule out the presence of native ligand in the purified proteins used for the assay which may contribute to a lower sensitivity for certain ligands. However, the proteins from *E. coli* and *S. typhimurium* differ by only three amino acids that are distal to the binding pocket, suggesting that additional tests may be needed to verify this result. Overall, the assay results obtained with the set of validation targets demonstrated a dose response for specific binding ligands and a ligand-binding specificity consistent with the historical annotation (Table 1). These observations support the utility for functional assignment but suggest some ligands with low binding affinity may not be detected using the present format.

Proteins experimentally validated for multiple ligand-binding specificity include the lysine/arginine/ornithine (LAO) proteins (*50*−*52*) and the histidine-binding protein from *E. coli* (*42*), as well as the cysteine (*40*) and leucine/isoleucine/valine-binding (LIV) proteins (*53*). For most of the targets, the spectrum of binding ligands identified by the FTS assay was very similar to the ligands identified by other experimental approaches with some minor differences. Proteins which bound multiple ligands displayed varying capabilities for thermal stabilization dependent on the particular binding ligand. We noted a general correlation between relative ΔT_m values and ligand affinity constants reported in the literature. However, direct comparisons of relative binding constants based on comparison of the ΔT_m values are not appropriate as the binding affinity is a function of both the ΔT_m and the binding enthalpy. Several studies have compared the linkage between ligand binding and protein stability (*21, 23, 48*) and also provide a process for the rigorous assessment of ligand affinity using thermal denaturation data.

For two of the targets, an additional binding ligand that had not been previously reported was detected in the FTS assay. The binding of cysteine was consistently detected for the LAO protein from *S. typhimurium* in both the individual and pooled ligand screens. The historical ligand-binding assignments for the *S. typhimurium* protein were based on crystal structures obtained with four different amino acids (lysine, arginine, ornithine, and histidine). In contrast to the screening approach applied in this study, the crystals of the *S. typhimurium* protein were obtained by a directed approach with the desired amino acid mixed with the purified protein prior to crystallization with no indication that additional amino acids ligands were incorporated into the crystallization process. However, the crystal structures with the various ligands provide additional information to evaluate the feasibility for cysteine binding. In these structures, the orientation of amino acids in the binding pocket is fairly well conserved for all of the bound ligands. The various ligands are accommodated by a large cavity size, ligand-mediated relocation of protein-bound water molecules in the binding site, and the movement of the Asp-11 side chain. In this respect there are some differences in the mode of binding for each of the ligands characterized by a crystal structure. For example, histidine binds with a K_d 30-fold less than lysine which is attributed to the inability to form an ionic interaction with the Asp-11 residue (*50*). These considerations suggest that the cavity size is sufficient to enable cysteine incorporation in the binding pocket. However, a different mode of binding would be anticipated due to the nucleophilic character of the thiol group with the relative binding affinity determined by the ability to accommodate the two water molecules in the active site. The thermal shift data obtained with the LAO protein from *E. coli* were suggestive of cysteine binding, but the magnitude of the shift did not provide sufficient confidence to assign this amino acid as a binding ligand. The LAO proteins from *E. coli* and *S. typhimurium* are similar, but several residues differ, and these changes may impact the relative binding ability for cysteine.

The FTS assay also identified methionine as a binding ligand for the LIV protein from *E. coli*, an observation not previously reported. The *E. coli* LIV protein has been crystallized in the ligand-free and ligand-bound (leucine, valine, and isoleucine) forms. The amino acids in the binding pocket show little conformational change with the three different ligands. The aliphatic side chains of the ligands are accommodated in a hydrophobic pocket formed by Tyr-18, Leu-77, Ala-100, and Phe-276 from one domain and Tyr-202 from the second domain (*53*). The majority of the contacts with the side chain are van der Waals and are suggested to account for the differences in the binding constants (apparent K_m) of Leu (2.3 μM), Ile (0.9 μM), and Val (4 μM) (*54*). Although not reported as a component of the crystal structures, the LIV protein from *E. coli* has also been shown to bind phenylalanine with a binding constant of 19 μM. The broad substrate specificity and features of the binding pocket of the LIV protein provide some support for our observation of methionine as a binding ligand.

The strength of the FTS assay is the capability to quickly survey ligand candidates for proteins of unknown or suspected function. Both of these applications are illustrated by the results presented in Table 2. For proteins where a function can be inferred by sequence homology or where functional data exists from biochemical or genetic analysis, the FTS approach provides a mechanism to validate the anticipated function. The microwell plate-based format and the capability to screen multiple ligands mean the FTS assay can rapidly identify a spectrum of candidate binding ligands. Our results also suggest the usefulness of this application for screening proteins of unknown function. In view of the extensive number of biological ligands (>15000 ligands currently listed in the KEGG Ligand Database for Chemical Compounds, Release 47.0), some constraints on ligand library size and target pools must be imposed for practical reasons. Our approach for this study was to select a set of targets from a specific COG category. Each COG consists of individual proteins or groups of paralogues that are related by sequence. Proteins in the COG0834ET cluster are characterized as "ABC-type amino acid transport/signal transduction systems, periplasmic component/domain" and are appropriate for screening with an amino acid ligand library. The *S. oneidensis* genome contains 12 proteins (NCBI annotation) clustered in the COG0834ET category. The target selected from COG2998H has a signal peptide and is part of an operon encoding an ABC transporter, and the sequence displays limited homology to the crystal structure of a periplasmic putative metal-binding protein (e-value $= 6.3 \times 10^{-19}$). The *S. oneidensis* target selected from COG2998H was strongly stabilized by tungstate and showed some stabilization with molybdate. The assignment of this protein as a tungstate-binding component of the ABC transporter is consistent with the COG clustering process. The ability of the FTS assay to verify the metal-binding annotation of this protein suggests that this assay may not just be suitable for amino acid-binding proteins but may also be amenable to various protein classes with different binding ligands.

A single ligand-binding event with an amino acid was identified for the protein from *S. oneidensis*. This protein was selective for Arg and did not show any shifts with other amino acids. This profile is the same as the ArtJ protein from *E. coli* which transports only Arg with high affinity. Sequence alignments generated using BLAST against the NCBI nonredundant database indicate the presence of many homologues. The protein is highly conserved in the 15 sequenced genomes from *Shewanella* species that are deposited in the NCBI database. The sequence from *Shewanella benthica* has the lowest *e*-score for the alignments but still retains 77% identity over the length of the sequence alignment (225 residues). It is interesting that the top 50 BLAST hits are annotated as proteins of unknown function or are assigned very generic functional annotation. Sequence alignments generated using BLAST to identify homologous sequences in the PDB indicate similarity with sequences annotated as PBPs with various amino acid specificities. The alignment with the highest *e*-score is with the arginine-, lysine-, and histidine-binding protein from *Geobacillus stearothermophilus* with an identity of 39% over a span of 221 amino acids (*55*). The next set of alignments with sequence identity of 25% or greater includes a number of uncharacterized proteins and amino acid-binding proteins with specificities for histidine, glutamate, cysteine, glutamine, lysine, and arginine.

Two structures in the PDB were crystallized with arginine and are annotated as periplasmic binding proteins. One of these is the arginine/lysine/histidine-binding protein from *G. stearothermophilus* identified in the global BLAST search. The second is the ArtJ protein from *E. coli* which has been characterized for selective binding to arginine. In the ligand-bound form of these proteins, arginine makes contact with nine amino acids in the *E. coli* protein and 10 amino acids in the protein from *G. stearothermophilus*. Alignment of these sequences indicates 50% of the arginine contact residues are identical for the two proteins with a generally conservative change at other positions. Inclusion of the *S. oneidensis* arginine-binding protein sequence in the alignments is consistent with the conservation of residues involved in arginine binding. Although the contact residues are not identical for the three proteins, the aligned amino acids of the protein from *S. oneidensis* are identical to the corresponding binding amino acids of one or both of the other sequences from the PDB at all of the binding positions. These comparisons provide persuasive support for the assignment of this protein as an arginine-binding protein by the FTS assay.

Analysis of the genome region of the *Shewanella* gene indicates that the Arg transporter is part of an operon containing the permease and ATPase component of the transporter, suggesting the annotation assigned to the binding protein can be effectively extended to other components of the transporters. This does not exclude other functional roles as some binding proteins can release bound ligands to multiple permease proteins. The *hisJ* and *argT* genes are part of a single operon also containing genes encoding a transporter and ATPase. These *hisJ* and *argT* genes encode PBPs with high affinity for histidine and arginine, respectively, and both interact with the same transporter (*56*). A similar situation exists in *E. coli* where the *hisJ* and *argT* genes share the same transporter (*57*). Although the genes are adjacent to each other in the *E. coli* genome, they have different promoters.

The overall success rate for determination of bound ligands for the unknown protein set from *Shewanella* was relatively low (33%). This is not unexpected in view of the diverse spectrum of ligands known to bind this class of proteins. The scope of this study was intentionally narrowed to focus on amino acid-binding proteins with the intention to validate specificity and reproducibility of FTS assay. However, there are few constraints to increasing the size of the ligand library. The data from the screening set using pooled ligands demonstrate the feasibility for expansion of the ligand library while maintaining throughput capabilities. An asset of the FTS approach is the capability to survey a large number of targets for individual organisms. The limited annotation provided for the group of extracellular solute receptors in *S. oneidensis* is not an exception. In the majority of sequenced genomes, annotation of this family and many other ligand-binding proteins is limited to mostly generic descriptions of the capability. This is attributed to the limitation imposed by the current annotation process since BLAST alignments often do not allow assignment of a binding specificity for this class of proteins. Proteins in this family have a similar fold with sometimes subtle differences in the amino acids of the binding cleft determining the specificity. This was elegantly illustrated in a comparison of the cysteine-binding protein Cj0982 from *C. jejuni* and the glutamine-binding protein of *E. coli* (*40*). The alignments of the two structures have an rms deviation of 1.15 Å with ligands occupying a similar position in the binding pocket. BLAST searches using these sequences yield a roughly equivalent set of hits with the exception that hits exclusive to Cj0982 have an invariant Arg-74 residue and an invariant GlyThrThrAla motif at residues 137−140. This position-specific combination of amino acids provides a set of putative cysteine transport proteins. This approach for identification of putative function complements the capabilities of the FTS assay which could provide a method for validation of the proposed function. An advantage of the FTS approach is the ability to rapidly survey many proteins which provides the possibility to elucidate the particular residues that determine specificity. We suggest an application of the FTS assay would be the accumulation of sufficient functional data for a set of homologues that can be used to define position-specific sequence motifs associated with specific ligand binding. These motifs can improve annotation assignments based on sequence alignments, resulting in a more comprehensive understanding of an organism's metabolism and functional capabilities.

ACKNOWLEDGMENT

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

REFERENCES

1. Metzker, M. L. (2005) Emerging technologies in DNA sequencing. *Genome Res. 15*, 1767–1776.
2. Schuster, S. C. (2008) Next-generation sequencing transforms today's biology. *Nat. Methods 5*, 16–18.
3. Stothard, P., and Wishart, D. S. (2006) Automated bacterial genome analysis and annotation. *Curr. Opin. Microbiol. 9*, 505–510.
4. Brent, M. R. (2008) Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nat. Rev. Genet. 9*, 62–73.
5. Valencia, A. (2005) Automatic annotation of protein function. *Curr. Opin. Struct. Biol. 15*, 267–274.
6. Jones, C. E., Brown, A. L., and Baumann, U. (2007) Estimating the annotation error rate of curated GO database sequence annotations. *BMC Bioinf. 8*, 170.
7. Raes, J., Harrington, E. D., Singh, A. H., and Bork, P. (2007) Protein function space: viewing the limits or limited by our view? *Curr. Opin. Struct. Biol. 17*, 362–369.
8. Shrager, J. (2003) The fiction of function. *Bioinformatics 19*, 1934–1936.
9. Reeves, G. A., and Thornton, J. M. (2006) Integrating biological data through the genome. *Hum. Mol. Genet. 15* (Spec. No. 1), R81–R87.
10. Yakunin, A. F., Yee, A. A., Savchenko, A., Edwards, A. M., and Arrowsmith, C. H. (2004) Structural proteomics: a tool for genome annotation. *Curr. Opin. Chem. Biol. 8*, 42–48.
11. Bashton, M., Nobeli, I., and Thornton, J. M. (2006) Cognate ligand domain mapping for enzymes. *J. Mol. Biol. 364*, 836–852.
12. Rudisser, S., and Jahnke, W. (2002) NMR and in silico screening. *Comb. Chem. High Throughput Screening 5*, 591–603.
13. Weber, P. C., and Salemme, F. R. (2003) Applications of calorimetric methods to drug discovery and the study of protein interactions. *Curr. Opin. Struct. Biol. 13*, 115–121.
14. Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Jr., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U.S.A. 103*, 15835–15840.
15. Makara, G. M., and Athanasopoulos, J. (2005) Improving success rates for lead generation using affinity binding technologies. *Curr. Opin. Biotechnol. 16*, 666–673.
16. Freire, E. (1995) Differential scanning calorimetry. *Methods Mol. Biol. 40*, 191–218.
17. Kelly, S. M., and Price, N. C. (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochim. Biophys. Acta 1338*, 161–185.
18. Royer, C. A. (1995) Fluorescence spectroscopy. *Methods Mol. Biol. 40*, 65–89.
19. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem. 357*, 289–298.
20. Cummings, M. D., Farnum, M. A., and Nelen, M. I. (2006) Universal screening methods and applications of ThermoFluor. *J. Biomol. Screening 11*, 854–863.
21. Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screening 6*, 429–440.
22. Chung, C. W. (2007) The use of biophysical methods increases success in obtaining liganded crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 63*, 62–71.
23. Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem. 332*, 153–159.
24. Acher, F. C., and Bertrand, H. O. (2005) Amino acid recognition by Venus flytrap domains is encoded in an 8-residue motif. *Biopolymers 80*, 357–366.
25. Dassa, E. (2007) Periplasmic ABC Transporters, in *The Periplasm* (Ehrmann, M., Ed.) pp 271−303, ASM Press, Washington, DC.
26. Davidson, A. L., and Maloney, P. C. (2007) ABC transporters: how small machines do a big job. *Trends Microbiol. 15*, 448–455.
27. Linton, K. J., and Higgins, C. F. (1998) The Escherichia coli ATP-binding cassette (ABC) proteins. *Mol. Microbiol. 28*, 5–13.

28. Dassa, E., and Bouige, P. (2001) The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Res. Microbiol. 152*, 211–229.
29. Saier, M. H., Jr. (1999) A functional-phylogenetic system for the classification of transport proteins. *J. Cell. Biochem., Suppl. 32−33*, 84–94.
30. Tam, R., and Saier, M. H., Jr. (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev. 57*, 320–346.
31. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol. 6*, 175–182.
32. Abdullah, J., Joachimiak, A., and Collart, F. R. (2008) "System 48" High Throughput Cloning and Protein Expression Analysis, in *High Throughput Protein Expression and Purification* (Doyle, S., Ed.) pp 117−128, Humana Press, Totowa, NJ.
33. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expression Purif. 25*, 8–15.
34. Donnelly, M. I., Zhou, M., Millard, C. S., Clancy, S., Stols, L., Eschenfeldt, W. H., Collart, F. R., and Joachimiak, A. (2006) An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expression Purif. 47*, 446–454.
35. Dieckman, L. J., Zhang, W., Rodi, D. J., Donnelly, M. I., and Collart, F. R. (2006) Bacterial expression strategies for human angiogenesis proteins. *J. Struct. Funct. Genomics*.
36. Lin, C. T., Moore, P. A., Auberry, D. L., Landorf, E. V., Peppler, T., Victry, K. D., Collart, F. R., and Kery, V. (2006) Automated purification of recombinant proteins: combining high-throughput with high yield. *Protein Expression Purif. 47*, 16–24.
37. Deng, J., Davies, D. R., Wisedchaisri, G., Wu, M., Hol, W. G., and Mehlin, C. (2004) An improved protocol for rapid freezing of protein samples for long-term storage. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 203–204.
38. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc. 2*, 2212–2221.
39. Hollenstein, K., Dawson, R. J., and Locher, K. P. (2007) Structure and mechanism of ABC transporter proteins. *Curr. Opin. Struct. Biol. 17*, 412–418.
40. Muller, A., Thomas, G. H., Horler, R., Brannigan, J. A., Blagova, E., Levdikov, V. M., Fogg, M. J., Wilson, K. S., and Wilkinson, A. J. (2005) An ATP-binding cassette-type cysteine transporter in Campylobacter jejuni inferred from the structure of an extracytoplasmic solute receptor protein. *Mol. Microbiol. 57*, 143–155.
41. Wissenbach, U., Six, S., Bongaerts, J., Ternes, D., Steinwachs, S., and Unden, G. (1995) A third periplasmic transport system for L-arginine in Escherichia coli: molecular characterization of the artPIQMJ genes, arginine binding and transport. *Mol. Microbiol. 17*, 675–686.
42. Merlin, C., Gardiner, G., Durand, S., and Masters, M. (2002) The Escherichia coli metD locus encodes an ABC transporter which includes Abc (MetN), YaeE (MetI), and YaeC (MetQ). *J. Bacteriol. 184*, 5513–5517.
43. Kadner, R. J. (1977) Transport and utilization of D-methionine and other methionine sources in Escherichia coli. *J. Bacteriol. 129*, 207–216.
44. Kreimer, D. I., Malak, H., Lakowicz, J. R., Trakhanov, S., Villar, E., and Shnyrov, V. L. (2000) Thermodynamics and dynamics of histidine-binding protein, the water-soluble receptor of histidine permease. Implications for the transport of high and low affinity ligands. *Eur. J. Biochem. 267*, 4242–4252.
45. Magnusson, U., Chaudhuri, B. N., Ko, J., Park, C., Jones, T. A., and Mowbray, S. L. (2002) Hinge-bending motion of D-allose-binding protein from Escherichia coli: three open conformations. *J. Biol. Chem. 277*, 14077–14084.
46. Magnusson, U., Salopek-Sondi, B., Luck, L. A., and Mowbray, S. L. (2004) X-ray structures of the leucine-binding protein illustrate conformational changes and the basis of ligand specificity. *J. Biol. Chem. 279*, 8747–8752.
47. Esposito, D., and Chatterjee, D. K. (2006) Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol. 17*, 353–358.
48. Matulis, D., Kranz, J. K., Salemme, F. R., and Todd, M. J. (2005) Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry 44*, 5258–5266.
49. Wolf, A., Lee, K. C., Kirsch, J. F., and Ames, G. F. (1996) Ligand-dependent conformational plasticity of the periplasmic histidine-binding protein HisJ. Involvement in transport specificity. *J. Biol. Chem. 271*, 21243–21250.
50. Oh, B. H., Ames, G. F., and Kim, S. H. (1994) Structural basis for multiple ligand specificity of the periplasmic lysine-, arginine-, ornithine-binding protein. *J. Biol. Chem. 269*, 26323–26330.
51. Oh, B. H., Pandit, J., Kang, C. H., Nikaido, K., Gokcen, S., Ames, G. F., and Kim, S. H. (1993) Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J. Biol. Chem. 268*, 11348–11355.
52. Rosen, B. P. (1973) Basic amino acid transport in Escherichia coli. II. Purification and properties of an arginine-specific binding protein. *J. Biol. Chem. 248*, 1211–1218.
53. Trakhanov, S., Vyas, N. K., Luecke, H., Kristensen, D. M., Ma, J., and Quiocho, F. A. (2005) Ligand-free and -bound structures of the binding protein (LivJ) of the Escherichia coli ABC leucine/isoleucine/valine transport system: trajectory and dynamics of the interdomain rotation and ligand specificity. *Biochemistry 44*, 6597–6608.
54. Koyanagi, T., Katayama, T., Suzuki, H., and Kumagai, H. (2004) Identification of the LIV-I/LS system as the third phenylalanine transporter in Escherichia coli K-12. *J. Bacteriol. 186*, 343–350.
55. Vahedi-Faridi, A., Eckey, V., Scheffel, F., Alings, C., Landmesser, H., Schneider, E., and Saenger, W. (2008) Crystal structures and mutational analysis of the arginine-, lysine-, histidine-binding protein ArtJ from Geobacillus stearothermophilus. Implications for interactions of ArtJ with its cognate ATP-binding cassette transporter, Art(MP)2. *J. Mol. Biol. 375*, 448–459.
56. Higgins, C. F., and Ames, G. F. (1981) Two periplasmic transport proteins which interact with a common membrane receptor show extensive homology: complete nucleotide sequences. *Proc. Natl. Acad. Sci. U.S.A. 78*, 6038–6042.
57. Caldara, M., Minh, P. N., Bostoen, S., Massant, J., and Charlier, D. (2007) ArgR-dependent repression of arginine and histidine transport genes in Escherichia coli K-12. *J. Mol. Biol. 373*, 251–267.

BI801648R