From The Cover: Systematic investigation of protein phase behavior with a microfluidic formulator

Carl L. Hansen, Morten O. A. Sommer, and Stephen R. Quake

PNAS 2004;101;14431-14436; originally published online Sep 27, 2004;
doi:10.1073/pnas.0405847101

This information is current as of January 2007.

Online Information & Services
High-resolution figures, a citation map, links to PubMed and Google Scholar, etc., can be found at:
www.pnas.org/cgi/content/full/101/40/14431

Supplementary Material
Supplementary material can be found at:
www.pnas.org/cgi/content/full/0405847101/DC1

References
This article cites 16 articles, 5 of which you can access for free at:
www.pnas.org/cgi/content/full/101/40/14431#BIBL

This article has been cited by other articles:
www.pnas.org/cgi/content/full/101/40/14431#otherarticles

E-mail Alerts
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

Rights & Permissions
To reproduce this article in part (figures, tables) or in entirety, see:
www.pnas.org/misc/rightperm.shtml

Reprints
To order reprints, see:
www.pnas.org/misc/reprints.shtml
Systematic investigation of protein phase behavior with a microfluidic formulator

Carl L. Hansen, Morten O. A. Sommer, and Stephen R. Quake*

Department of Applied Physics, California Institute of Technology, MS 128-95, Pasadena, CA 91125

Communicated by Robert H. Austin, Princeton University, Princeton, NJ, August 16, 2004 (received for review February 9, 2004)

We demonstrated a microfluidic device for rapidly generating complex mixtures of 32 stock reagents in a 5-nl reactor. This “formulation chip” is fully automated and allows thousands of experiments to be performed in a single day with minimal reagent consumption. It was applied to systematically study the phase behavior of the protein xylanase over a large and complex chemical space. For each chemical formulation that demonstrated a pronounced effect on solubility, the protein phase behavior was completely mapped in the chip, generating a set of empirical phase diagrams. This ab initio phase information was used to devise a rational crystallization screen that resulted in 72-fold improvement in successful crystallization hits compared with conventional sparse matrix screens. This formulations tool allows a physics-based approach to protein crystallization that may prove useful in structural genomics efforts.

The application of x-ray crystallography to the determination of protein structure with atomic resolution was a triumph of structural biology in the 20th century. Since the first solution of the structure of myoglobin in 1958 (1), >23,000 different structures have been deposited in the Protein Data Bank, and their role in relating structure to function in biology has been profound. As structure determination efforts continue to move past the most tractable crystallization targets (typically small soluble proteins) and focus instead on more challenging macromolecules, such as large protein complexes and membrane proteins (2), the need to better understand and explore the crystallization process has become urgent. That is because once high-quality crystals are at hand, advances in x-ray sources, computer codes, and related technology have made it relatively straightforward to obtain the structure. Determining the appropriate crystallization conditions has become one of the most significant remaining bottlenecks to structure determination (3).

Understanding the phase behavior of proteins is an essential part of the crystallization process. The growth of crystals from a protein solution requires the existence of a nontrivial phase diagram, which allows the protein state to be manipulated between at least two thermodynamic phases: soluble and precipitated. The processes of crystal nucleation and growth arise on the boundary between these two phases and are governed by subtle effects in physical chemistry. There are a variety of schemes that manipulate the kinetics of the crystallization process, and all take advantage of generic features of these phase diagrams (4). However, in practice, the phase behavior of very few proteins has been studied in detail (5–12), and solubility information for a specific protein is rarely available for crystallization and optimization experiments (13, 14).

Furthermore, it is often an arduous process to find the right combination of chemicals that yields appropriate phase behavior for a given protein. Every protein is different, and even a modest subset of stock precipitating solutions comprises a vast chemical phase space that must be explored. The large amounts of sample required make systematic exploration by conventional techniques infeasible, and screening is typically directed toward an incomplete factorial or sparse-matrix approach, a brute-force process requiring large numbers of experiments (15, 16). There have been numerous attempts to rationalize this procedure, for example, by using computational approaches to predict phase behavior (17, 18) or by trying to correlate measurements of osmotic second virial coefficients (19, 20) with crystallization conditions. Practical limitations have thus far prevented these techniques from being generally applicable.

Here we describe a microfluidic formulation device that allows for the combinatorial mixing of 16 buffers and 16 precipitation agents with a purified protein sample. The ability of the formulation chip to access a vast number of chemical conditions and to accurately dispense and mix fluids on the picoliter scale makes detailed characterization of macromolecule phase behavior both possible and practical. We used this device to screen 4,300 different solubility conditions of the model protein Endo-1,4-β-xylanase from Trichoderma reesei. Xylanase is a 21-kDa member of the gluconase enzyme family. For those conditions that exhibited nontrivial phase behavior (i.e., precipitation), a full phase diagram was generated. From this thorough characterization of the phase behavior, we designed a rational crystallization screen for xylanase. Comparison of this screen to four commercially available sparse matrix screens showed an increase of nearly two orders of magnitude in crystallization success and allowed new insight into the physics of crystallization.

Materials and Methods

Sample Preparation and Crystallization Protocols. Endo-1,4-β-xylanase (xylanase) from T. reesei (Hampton Research, Riverside, CA) was prepared in deionized water from stock (36 mg/ml protein/43% wt/vol glycerol/0.18 M sodium/potassium phosphate, pH 7.0) by repeated buffer exchange at 4°C using a centrifugal filter with a molecular mass cutoff of 10,000 Da (Micron Separations, Westboro, MA). Protein concentration was measured by absorption at 280 nm and adjusted to a stock concentration of 120 mg/ml. Ten-microliter aliquots were flash frozen in liquid nitrogen and stored at –80°C. To avoid sample–sample variations, a single sample preparation was used for all solubility screening, phase-space mapping, and corresponding crystallization experiments. Batch crystallization trials were actively mixed by repeated aspiration and incubated under paraffin oil. Crystalization trials were inspected daily for a period of 2 weeks. Observed crystals were confirmed to be protein crystals by staining (IZIT dye, Hampton Research) and were recorded as crystallization hits.

Experimental Setup and Data Collection. Automation of metering, mixing, and data acquisition allows for thousands of solubility experiments to be executed without the need for user intervention. In each solubility experiment, a unique mixture of the 32 reagents and the protein sample is produced in the chip; details of chip fabrication are in Supporting Text, which is published as supporting information on the PNAS web site. All device control and data acquisition were implemented by using a custom

Abbreviation: PCI, positive displacement cross-injection.

*To whom correspondence should be addressed. E-mail: quake@caltech.edu.

© 2004 by The National Academy of Sciences of the USA
software driver developed in LABVIEW (National Instruments, Austin, TX). Mixing recipes were generated by using a spreadsheet program and translated into valve actuation sequences by the software driver. Off-chip solenoid valves (Fluidigm, South San Francisco, CA), controlled by using a digital input–output card (DIO-32HS; National Instruments), were used to generate square-wave pressure signals at the device control ports. A frame-grabber card (Imagization PXC200A, CyberOptics, Minneapolis) was used to automate image acquisition from a charge-coupled device camera.

Absorption and Precipitation Measurements. Absorption measurements were taken to determine the concentration of bromophenol blue sodium salt (absorption peak at 590 nm) in the mixing ring. A 9-µm-high segment of the mixing ring (∼300 × 50 µm) having rectangular cross section was illuminated with a 590-nm diode (AND180HYP; Newark Electronics, Chicago) and imaged through a stereocope (SMZ 1500; Nikon) onto a charge-coupled device camera. Pixel intensities were averaged and compared with an identical adjacent reference channel containing the undiluted dye (2 mM bromophenol blue sodium salt/100 mM Tris-HCl, pH 8.0). In some experiments, glycerol was added to the injected dye to vary the viscosity. Dye concentrations were determined by using the Beer–Lambert relation and used to calculate the injected volume.

Precipitation of the protein was automatically detected by imaging a portion of the mixing ring, calculating the standard deviation of the pixel intensities, and comparing this value to the background (no protein added). To ensure even illumination, images were taken at ×112 magnification at a 9-µm-high section of the mixing ring having a rectangular cross section.

Positive-Displacement Crossinjection Metering. The metering scheme allows for sequential injection of precise sample aliquots from a single microfluidic channel into an array of reaction chambers through a positive displacement crossinjection (PCI) junction. The PCI junction is formed by the combination of a three-valve peristaltic pump (21) and a four-port crossinjection junction with integrated valves on each port (Fig. L4). At each junction, two sets of valves are actuated to direct the flow either horizontally or vertically. To execute the metering task, the flow is switched vertically through the junction, charging the cross-

injection with the sample fluid (Fig. 1B). The flow is then directed horizontally through the junction, and the three valves forming the peristaltic pump are actuated in a five-state sequence to advance the fluid in the horizontal direction (Fig. 1C). Each cycle of the peristaltic pump injects a well defined volume of sample (∼80 pl) determined by the dead volume under the middle valve of the peristaltic pump. The deflection of the valve membranes when not actuated is determined by the pressure difference across the membrane. The volume injected during each cycle therefore may be tuned continuously, allowing for variable positive displacement metering. By repeating the injection sequence, the volume of injected solution may be increased in 80-pl increments, allowing for the dynamic quantized control of the final downstream sample concentration.

The active region of the microfluidic formulation chip that implements this scheme and allows for the arbitrary combinatorial mixing of 16 stock reagents into one of 16 buffer solutions is shown in Fig. 2.4. Two 16-solution multiplexer arrays, actuated by eight control lines, allow for the selection of buffers (Fig. 2A) and reagents (bottom, Fig. 2A). A PCI junction, formed by a three-valve peristaltic injection pump (red) and crossinjection valves (center green), dispenses directly into a 5-nl ring reactor. Once the reactor has been flushed, a reagent line is selected, and the crossinjection sequence is executed. The extended split channel region increases the volume of the crossinjection junction, thereby allowing for up to 15 injections between flushing steps. The maximum number of consecutive injections that may be executed before the junction needs to be refreshed depends on the Taylor dispersion (22) of the injected fluid as it is pumped down the channel and is therefore a function of the viscosity.

Fig. 2B shows the injection of four slugs, each having a volume of 80 pl, into the ring reactor. Arbitrary combinations of 16 reagents may be produced in the reactor by sequential flushing and injection steps. Fig. 2C shows a color gradient formed from injections of water and blue, green, yellow, and red dyes. In screening applications that require the interrogation of a precious sample against many premixed reagent formulations, the crossinjection flushing step is wasteful and is circumvented by the addition of a separate sample injection site (Fig. 2D). After the ring is filled with the desired reagents, they are mixed by

14432 | www.pnas.org/cgi/doi/10.1073/pnas.0405847101

Hansen et al.
Precise and robust microfluidic metering. Absorption spectroscopy was used to measure the replaced fraction of fluid in the ring reactor from PCI injections. The total ring volume is nominally 5 mL. (A) Precision and reproducibility of PCI injections. Each of the nine clusters represents 100 identical injection sequences; the standard deviation of the clusters corresponds to an injection error of \( \pm 0.6 \) pl. (B) Absorption measurements of four sets of 20 injection and mixing sequences show the metering to be robust to the viscosity of the injected fluid. Fluids contain varying amounts of glycerol and have viscosities ranging from 1 to 400 cp.

Device Characterization. The precision of metering was evaluated by injecting variable amounts of dye (bromophenol blue sodium salt; Sigma) into a reactor, mixing, and performing absorption measurements. A set of 900 sequential titration experiments (Fig. 3A) shows the metering to be both precise and reproducible, with a slope of 83.4 pl per injection cycle and a coefficient of correlation of 0.996. The standard deviation of the injected slug volume was determined to be \( \pm 0.6 \) pl. Although positive displacement metering ensures that the injected volume is robust to changes in the fluid viscosity, the viscosity of the working fluid does reduce the bandwidth of the injector. It was found that for a solution having a viscosity of 400 cp, the frequency response of the injector began to roll off at 10 Hz. When operating at an injection frequency of 5 Hz, all solutions with viscosities \(<400\) cp produced equal injection volumes. Because the metering mechanism is completely mechanical, there is no dependence on the pH or ionic strength of the injected fluid. Additionally, because the fluid is not dispensed from the chip, there is no phase interface and therefore little dependence on surface tension. The metering technique is therefore truly robust to the physical properties of the injected fluid. Titration experiments with fluids of varying glycerol concentration show the injection volume to vary by \(<5\%\) over a viscosity range of 1–400 cp without modification to the injection sequence (Fig. 3B).

Results and Discussion

Although small-scale characterization of protein solubility by a precrystallization solubility assay has been reported (15, 24, 25), this technique has not been widely adopted, because the large required sample volumes make it unsuitable for targets that cannot be expressed and purified in large quantities. Microfabricated dispensers have been used to reduce sample consumption in cases where the sequential addition of reagents to a levitated drop of microliter volume is sufficient to explore a restricted chemical space (25). Although microfluidic devices have been previously used to screen crystallization conditions using free interface diffusion (19) and microbatch formats (20), they have not been applied to systematically measure phase behavior (26, 27).

Thorough characterization of protein solubility behavior requires accessing a vast chemical space through the combinatorial mixing of a limited number of stock reagents. The conventional reagents used in crystallization exhibit a large variation in physical properties, such as viscosity, surface tension, ionic strength, and pH. This variation presents a formidable challenge for fluid-handling systems that must allow for arbitrary fluid combinations and proportioning. We developed a PCI metering method that overcomes this obstacle, allowing for variable dispensing to be dynamically programmed by the user in 80-pl increments with \(<5\%\) variation over a broad range of fluid properties. By combining this method with microfluidic mixing (23) and multiplexing elements (28), large-scale combinatorial screening has been achieved on chip for the first time. The flexibility, precision, and small volume requirements of this device make feasible the systematic mapping of crystallization phase space.

To demonstrate the utility of ab initio protein solubility characterization, we used the formulator chip to exhaustively explore the solubility behavior of a commercially available crystallization standard, Endo-1,4-\( \beta \)-xylanase (xylanase) from T. reesei (Hampton Research) (29, 30). Trial solutions were imaged in the chip’s reactor ring, and the standard deviation of the intensity of the imaged pixels was used as a metric of precipitation; this allowed for distinction between precipitated and soluble conditions and provided a rough quantitative measure of the degree of precipitation. Beyond the precipitation limit, the pixel standard deviation increases linearly with the protein concentration and therefore is proportional to the concentration of precipitated protein present in the solution (Fig. 4A). A video of on-chip protein titration and precipitation is included as Movie 1, which is published as supporting information on the PNAS web site.

A coarse search of solubility space was performed initially, in which the protein sample was mixed with a large number of chemical formulations to identify conditions that produce nontrivial phase behavior. This generates a solubility fingerprint of the crystallization target in which each precipitation peak indicates a chemical formulation that exerts a pronounced effect on solubility. Fingerprints were generated for xylanase in four independent runs, each of which consists of \( \sim 4,000 \) titration experiments (Fig. 4B). These fingerprints are highly reproducible and are characteristic of the protein studied. For example, although sodium chloride is a strong precipitating agent for chicken egg white lysozyme, it does not produce a precipitation peak in the solubility fingerprint of xylanase over the pH range studied.

Each solubility fingerprint was generated over a period of 20 h and consumed \( \sim 8 \) \( \mu \)L of protein sample. Chemical formulations were created by flushing the ring with one of 16 buffers, injecting a precipitating agent (salt or polymer), diluting the ring with water, and then mixing. Protein sample was then introduced at a variety of concentrations and mixed before data acquisition. When a polymer was used as the major precipitating agent, a small amount of salt was also introduced as an additive. Experiments in Fig. 4B were grouped by the identity of the major precipitating agent so that each peak represents the effect of this reagent over a range of pH values and concentrations. The large width of these peaks indicates a high level of experimental redundancy, suggesting that a more efficient search could be conducted using fewer related chemical conditions.

The solubility fingerprint of xylanase reveals five salts (sodium citrate, dipotassium phosphate, ammonium sulfate, and sodium/potassium tartrate) as likely crystallizing agents. A high molecular-weight polymer (polyethylene glycol 8000) in combination with various salt additives was also identified to be a strong precipitating agent at high pH values. The high isoelectric point of xylanase suggests that the reduced effectiveness of this precipitant at low pH values is due to two-body
deviation exhibits an approximately linear dependence on protein concentration. Below the precipitation limit, the standard deviation of pixel intensities provides a quantitative metric of protein solubility. (A) Conditions. The top solubility fingerprint (red) is generated by using a sample having elevated protein concentration (90 mg/ml) and exhibits both higher signal-to-noise ratio and additional peaks not present in the other data series (70 mg/ml). The two center solubility fingerprints were generated sequentially on a single device (first green, then blue) with the same loaded sample, demonstrating the stability of the protein over the time of the experiment (~40 h). The bottom solubility fingerprint (pink) was generated on a separate device by using the same protein sample as the blue fingerprint, showing reproducibility across devices. (C) Comparison of xylanase phase mapping done on chip and in microbatch experiments. Conditions that gave rise to precipitation in microbatch format and in chip are represented by red squares and overlaid yellow circles, respectively. Conditions that did not produce precipitation in either format are shown as blue circles. (D) Reversibility of precipitation and solubility hysteresis for lysozyme observed by outward and return titrations from the origin. The red region shows the area of phase space in which precipitation was observed in both outward and return titrations. The yellow region is the area of hysteresis in which the protein was soluble for the outward titrations but precipitated in the return titrations. The blue region is the area in which the protein was soluble for both the outward and return titrations.

Electrostatic repulsion. A smaller molecular-weight polymer (polyethylene glycol 3350) was found to be a much weaker precipitating agent and was not investigated further in phase-space mapping experiments.

The reporter chip was then used to measure full-phase diagrams for the 24 most promising precipitation conditions. Each phase diagram consisted of a grid of 72 separate mixing experiments over all accessible protein and precipitant concentrations. All 24 phase spaces were generated sequentially on a single device by using <3 µl of protein sample (~100 nl per phase space) and are included in Fig. 6, which is published as supporting information on the PNAS web site. A comparison of precipitation phase spaces measured for xylanase in chip (5-ml reactions) and in microbatch format under paraffin oil (5-µl reactions) shows good agreement in detecting the precipitation boundary (Fig. 4C). Because measurements of precipitation are made immediately after mixing (within 3 sec), the locus of points that separate the precipitated and soluble regions of the graph generates a precipitation curve that is distinct from the thermodynamic solubility curve.

Conditions that reside just below the precipitated region may be in a metastable state conducive to crystallization, so that a detailed knowledge of protein solubility behavior provides an empirical basis for the design of maximum likelihood crystallization trials. The 24 phase spaces generated for xylanase were used to rationally design an optimal crystallization screen consisting of 48 reagents. The formulations and protein concentrations used in the optimized screen are in Table 1, which is published as supporting information on the PNAS web site. A single-batch crystallization trial using the optimal screen was set up by combining relative amounts of protein and precipitant stock so that the final condition was located on the boundary of the precipitation region. The efficiency of this screen was evaluated by comparison with standard commercially available sparse matrix screens (Crystal Screens I and II, Hampton Research; Wizards I and II; Emerald Biostructures, Bainbridge Island, WA). Two batch crystallization trials of 48 unique conditions were prepared for each of the four sparse matrix screens for a total of 384 individual assays.

A list of the successful conditions from the rationally designed crystallization screen and from the sparse matrix screens is provided as Table 2, which is published as supporting information on the PNAS web site. Twenty-seven crystallization conditions were observed in the optimal screen compared with a total of three crystallization conditions in the eight standard sparse matrix screens. The use of ab initio solubility information therefore resulted in a 72-fold enrichment in crystallization...
success (Fig. 5A). A surprising result was that xylanase crystals were observed in the optimal screen for all of the major salt and polymer precipitants identified in the solubility fingerprint. This striking finding suggests that achieving optimal levels of supersaturation is more important in the crystallization of xylanase than is the broad sampling of chemical space. In cases such as this, systematic screening for crystallization using a reduced chemical space may prove more effective than sparse matrix strategies. Additionally, crystallization conditions were identified in the rationally designed screen that gave large single three-dimensional crystals suitable for diffraction studies without further optimization (Fig. 5B). In contrast, only flat plate crystals were observed in the sparse matrix screens.

To evaluate the influence of lot variability on these crystallization results, crystallization trials based on the optimal screen were repeated by using a new protein sample obtained from the same vendor and prepared identically as the original sample. Fourteen of 17 polyethylene glycol conditions that gave crystals in the original experiment were reproduced by using the second sample, compared with 1 of 10 for the salt-based conditions. To determine whether the variable crystallization behavior observed in salt-based conditions was due to variations in phase-space behavior, a complete phase space of one chemical formulation (sodium/potassium tartrate, Tris-HCl, pH 8.5) was measured in microbatch format for both samples (Fig. 5C and D). It was discovered that, although both samples exhibited similar phase-space behavior, they produced different crystallization results. Eleven conditions produced crystals in the original sample compared with only one successful condition in the second sample. The reason for this difference in behavior is unclear but may be due to variable degrees of proteolysis or trace amounts of chemical contaminant introduced during purification or concentration steps. It is interesting that some crystallization conditions (e.g., those based on polyethylene glycol 8000) are more robust to batch-dependent perturbations than others.

Another application of protein solubility phase-space mapping is in transporting successful crystallization conditions from one experimental format to another. The successful crystallization of a protein is determined both by the established thermodynamic variables and the kinetic trajectory of an experiment. For this reason, experiments conducted with different crystallization kinetics (e.g., hanging-drop vapor diffusion, microbatch, and free-interface diffusion) and the same precipitating agents will not necessarily produce similar results. For example, the hydroxydecomposition domain of a cytochrome P450 alkaline hydroxylase (Mutant 139–3 of BM-3) did not produce crystals in initial hanging-drop trials but was found to crystallize readily by microfluidic free interface diffusion (26) (1 part protein 20 mg/ml/1 part 30% m/v polyethylene glycol 8000/0.2 M sodium acetate/0.1 M Tris-HCl, pH 7.0). However, this condition was unsuccessful when set in hanging-drop vapor diffusion format, resulting only in amorphous precipitate. The microfluidic formulator was used to generate a phase space at constant buffer and salt concentration (100 mM Tris-HCl, pH 7.3/200 mM sodium acetate) with polyethylene glycol concentration and protein concentration as variables. Two hanging-drop experiments were designed to equilibrate near the solubility limit determined from the phase-space map. One condition (8 µl of 35 mg/ml protein sample mixed with 6.7 µl of 10% polyethylene glycol/100 mM sodium acetate/50 mM Tris-HCl, pH 7.3, and equilibrated at 20°C against 1 mM 20% polyethylene glycol/200 mM sodium acetate/100 mM Tris-HCl, pH 7.3) produced crystals within 3 days. This success demonstrates the usefulness of solubility mapping in transporting conditions across crystallization formats.

Finally, we used the formulator to make a direct observation of the supersaturation region of chicken egg white lysozyme. The concentrations of salt and lysozyme were manipulated while keeping the buffer concentration constant to evolve the chemical state of the mixing ring radially out from the origin and then back again. Measurements of precipitation were taken at ~1-min intervals. The addition of a family of such radial titrations was used to generate two phase-space diagrams for chicken egg white lysozyme: one for the outward and one for the return titrations (Fig. 4D). The first observation of protein precipitation appears at higher salt and protein concentration during the outward trajectory than on the return path, thereby exhibiting solubility hysteresis. The intersection of the soluble region of the outward phase space with the precipitated region of the return phase space provides a direct observation of a metastable regime in which the aggregate phase is thermodynamically stable but not observed at short times. The observation of the reversible formation of a protein aggregate may be used to distinguish between denatured and well folded protein aggregates. Additionally identified metastable regions in phase space provide likely candidates for crystal seeding and growth experiments.

Conclusion
We have shown that complex sample processing at the nanoliter scale allows for a practical implementation of automated protein solubility characterization. Ab initio solubility information obtained through systematic protein phase-space mapping provides a physical basis for the design of optimal crystallization screens, giving rise to dramatic enrichment in crystallization success. It must be noted that chemical conditions and phase behavior are not the only variables that can be adjusted in the search for good crystals; it is often equally important to tune the properties of the protein by creating point mutants with or truncated constructs. However, the path to crystallization always includes extensive chemical screening with precious protein sample, and for this step it appears that microfluidic formulation devices can play an important role. Beyond applications in protein crystallization, the formulation capability of this device should find diverse applications in areas such as combinatorial chemistry, chemical synthesis, and cell culture studies.

We thank Sebastian Maerkl and James Berger for helpful conversations and insights. This work was supported by the National Science Foundation, the Defense Advanced Research Planning Agency, and the Army Research Office. C.L.H. was partially funded by the Natural Sciences and Engineering Research Council (Canada) (Julie Payette Fellowship).

Phase knowledge enables rational screens for protein crystallization

Megan J. Anderson, Carl L. Hansen, and Stephen R. Quake

PNAS 2006;103;16746-16751; originally published online Oct 30, 2006;
doi:10.1073/pnas.0605293103

This information is current as of January 2007.

Online Information & Services
High-resolution figures, a citation map, links to PubMed and Google Scholar, etc., can be found at:
www.pnas.org/cgi/content/full/103/45/16746

References
This article cites 13 articles, 7 of which you can access for free at:
www.pnas.org/cgi/content/full/103/45/16746#BIBL
This article has been cited by other articles:
www.pnas.org/cgi/content/full/103/45/16746#otherarticles

E-mail Alerts
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

Rights & Permissions
To reproduce this article in part (figures, tables) or in entirety, see:
www.pnas.org/misc/rightperm.shtml

Reprints
To order reprints, see:
www.pnas.org/misc/reprints.shtml

Notes:
Phase knowledge enables rational screens for protein crystallization

Megan J. Anderson*, Carl L. Hansen†, and Stephen R. Quake‡§

*Department of Biochemistry and Molecular Biophysics, California Institute of Technology, MS 128-95, Pasadena, CA 91125; †Departments of Physics and Astronomy and Electrical and Computer Engineering, Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, Canada V6T 1Z4; and ‡Department of Bioengineering, Stanford University and Howard Hughes Medical Institute, 318 Campus Drive, Room E300, Stanford, CA 94305

Edited by Stephen L. Mayo, California Institute of Technology, Pasadena, CA, and approved September 21, 2006 (received for review June 23, 2006)

We show that knowledge of the phase behavior of a protein allows one to create a rational screen that increases the success rate of crystallizing challenging proteins. The strategy is based on using microfluidics to perform large numbers of protein solubility experiments across many different chemical conditions to identify reagents for crystallization experiments. Phase diagrams were generated for the identified reagents and used to design customized crystallization screens for every protein. This strategy was applied with a 75% success rate to the crystallization of 12 diverse proteins, most of which failed to crystallize when using traditional techniques. The overall diffraction success rate was 33%, about double what was achieved with conventional automation in large-scale protein structure consortia. The higher diffraction success rates are achieved by designing customized crystallization screens using the phase behavior information for each target. The identification of reagents based on an understanding of protein solubility and the use of phase diagrams in the design of individualized crystallization screens therefore promotes high crystallization rates and the production of diffraction-quality crystals.

One of the main bottlenecks in the application of protein crystallography to large-scale structural biology efforts is the production of diffraction-quality crystals. Protein crystallization relies on the identification of reagents that promote crystal formation and an understanding of protein solubility in the presence of these reagents to achieve optimal crystal growth. Traditional techniques to identify reagents for crystallization include incomplete factorial searches across chemical space and sparse-matrix screening around reagents previously shown to crystallize proteins (1, 2). Although these techniques provide a starting point for crystallization experiments, they are not designed around the biophysical properties of individual proteins and have met with limited success for more challenging crystallization targets. For example, the National Institutes of Health Protein Structure Initiative (PSI) was able to generate diffraction-quality crystals from less than one in 5 of the 10,000 purified protein targets it has attempted to date (http://targetdb.pdb.org/statistics/TargetStatistics.html).

Knowledge of protein solubility in the presence of a crystallizing reagent enables one to effectively use the reagent in a way that will promote crystallization. The traditional method to study protein solubility in response to reagents is to construct a phase diagram that measures the solubility at different protein and reagent concentrations. The protein solubility curve outlined in each phase diagram depicts the most favorable protein and reagent concentrations to use in a crystallization experiment for optimal crystal growth. Currently, phase diagrams are only generated for readily available model proteins and for proteins with known crystallization conditions to improve crystal morphology (3–12). The infrequent use of phase diagrams for crystallization experiments is due to the large sample volumes and time required to construct the diagrams. The general utilization of phase diagrams to analyze protein solubility before starting crystallization experiments would lead to a more physical and rational approach to protein crystallization.

Here, we show that a crystallization strategy based on a detailed understanding of protein solubility increases the crystallization success rate for challenging proteins. A microfluidic formulator device was used to systematically screen hundreds of reagents against protein targets at several points on the phase diagrams to identify reagents that affect the solubility of the protein. Next, each of the identified reagents was explored by constructing a complete phase diagram outlining the solubility limits of the protein in the presence of the reagent. The phase diagrams were then used to design individualized crystallization screens tailored to the solubility properties of the target protein. This device was previously demonstrated to greatly enrich crystallization conditions for xylanase, UMP kinase, and the integral membrane protein SERCA (13, 14). The customized crystallization screens were tested by using a redesigned free interface diffusion screening device. Successful crystallization conditions were transported to a larger-scale crystal growth format for diffraction analysis.

We applied this rational screening strategy to the crystallization of 12 biologically diverse and challenging proteins, 8 of which failed to crystallize when using traditional techniques. The crystallization targets range in size from 16 kDa to 360 kDa complexes and include membrane proteins, large multiprotein complexes, a protein/RNA complex, a metalloprotein complex, metabolic proteins, and an extracellular matrix protein. Identification of reagents and phase diagram generation for each target were completed by using only microliters of sample, and we found that crystallization and diffraction success rates were roughly double that obtained by the PSI when using conventional automation. The increased success rates are noteworthy given the target diversity and that 67% of the targets are from eukaryotic sources while the PSI crystallization targets to date are biased toward prokaryotic sources with only 30% of the targets from eukaryotic sources (http://targetdb.pdb.org/statistics/TargetStatistics.html).

Results and Discussion

The phase diagram-based approach to protein crystallography was used to design and implement customized crystallization strategies for 12 challenging protein targets. To begin, protein solubility was screened against an extensive set of chemical conditions to identify suitable reagents for crystallization trials (Fig. 1A). The solubility screening included a full factorial search

Author contributions: M.J.A., C.L.H., and S.R.Q. designed research; M.J.A. performed research; C.L.H. contributed new reagents/analytic tools; M.J.A., C.L.H.; and S.R.Q. analyzed data; and M.J.A., C.L.H., and S.R.Q. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Abbreviations: MME, modified monomethyl ether; PSI, Protein Structure Initiative.

To whom correspondence should be addressed. E-mail: quake@stanford.edu.

© 2006 by The National Academy of Sciences of the USA

PNAS | November 7, 2006 | vol. 103 | no. 45 | 16746–16751
of 448 reagents composed of unique precipitant, buffer, and salt combinations. For each reagent, protein phase behavior was explored by sampling protein solubility at six different protein and precipitant concentrations across the phase diagram for a total of 2,688 solubility experiments per target. The ability of each reagent to manipulate protein solubility was quantified by the amount of protein aggregation observed when the protein was introduced to the reagent. Reagents that caused protein aggregation above a threshold level were considered to be potential crystallizing reagents for that target. The solubility screening across chemical and phase space identified between 51 and 246 potential crystallizing reagents for each of the targets.

Trends in the solubility screening results indicate that protein phase behavior was predominately controlled by the precipitant.

---

**Fig. 1.** Solubility information for protein targets. (A) Solubility screening results for 12 protein targets. Pixel intensity standard deviation represents the amount of protein aggregation seen in response to each reagent. The proteins are displayed from top to bottom in the order of decreasing protein aggregation: bR D85S, bR, Cbb3, AlaRS, Fis1, P450 1–12G, SMC/ScpA, P2X, AMG, VCP/Vimp, Tf/TfR, and 19S Lid. The solubility screening results are grouped by precipitating agent and are further subdivided by buffering agent. See Materials and Methods for a list of the reagents screened. (B) Classification of identified reagents by the molecular weight of the precipitating agent. (C) Classification of identified reagents by the deviation from the pl of the protein targets. (D) Classification of identified reagents by the ionic strength of the solution. (E) Classification of the identified reagents by the anion component of the reagent. The anions acetate and tartrate are abbreviated as Ace and Tar, respectively. (F) Classification of the identified reagents by the cation component of the reagent. (G) Comparison of the target molecular weight and the identified reagents for each target. For B–G, the percentage of identified reagents is normalized by the number of reagents screened in each category.

**Fig. 2.** Phase behavior characterization for targets. One complete phase diagram is shown for each target. (A) br DBSS with 0.125 M potassium acetate, 0.1 M sodium citrate (pH 6.5), and PEG 1500. (B) br with 0.125 M ammonium chloride, 0.1 M sodium acetate (pH 4.5), and PEG 1500. (C) Cbb3 with 0.075 M sodium acetate, 0.1 M sodium citrate (pH 6.5), and PEG 1500. (D) AlaRS with 0.1 M magnesium sulfate, 0.1 M sodium citrate (pH 5.5), and PEG 5000 MME. (E) Fis1 with 0.1 M ammonium acetate, 0.1 M Tris/HCl (pH 8.5), and PEG 8000. (F) P450 1–12G with 0.2 M ammonium sulfate, 0.1 M imidazole (pH 7.5), and PEG 8000. (G) SMC/ScpA with 0.05 M magnesium sulfate, 0.1 M sodium citrate (pH 5.5), and PEG 2000 MME. (H) P2X with 0.3 M sodium acetate, 0.1 M Hepes (pH 7.5), and PEG 1500. (I) AMG with 0.5 M sodium chloride, 0.1 M Tris/HCl (pH 8.5), and PEG 8000. (J) VCP/Vimp with 0.4 M ammonium acetate, 0.1 M sodium citrate (pH 6.5), and PEG 1500. (K) Tf/TfR with 0.2 M ammonium sulfate, 0.1 M Hepes (pH 7.5), and PEG 1500. (L) 19S Lid with 0.1 M ammonium chloride, 0.1 M sodium citrate (pH 5.5), and PEG 5000 MME. Each phase diagram screened protein solubility at 72 different precipitant and protein concentrations. The blue points represent concentrations where the protein is soluble, and the yellow points represent concentrations where the protein is insoluble.
The effectiveness of each of the reagent components on protein aggregation was determined by comparing the solubility results of reagents with two of the three components held constant. The precipitants used in the solubility experiments are polyethylene glycol (PEG) polymers with different chain lengths and end groups, giving each PEG polymer unique volume exclusion properties. Since the volume exclusion effect is more significant in longer polymers because of their larger radius of gyration, the longer polymers were used at a lower concentration than the shorter polymers. Polymers of four different chain lengths were examined, and for every target the number of potential reagents identified with longer polymers surpassed the number of reagents identified with shorter polymers. Additionally, a second PEG polymer with a modified monomethyl ether (MME) end group was screened at each polymer length; the MME polymers produced fewer identified reagents at each polymer length. Therefore, increasing the polymer chain length of the precipitating agent increases the protein aggregation response, and the polymer chemical composition also influences protein aggregation.

The solubility screening results also show that reagents with a pH value in the vicinity of the theoretical isoelectric point (pI) of the targets were identified as potential reagents with a slightly higher frequency than reagents that deviated from the target pI values (Fig. 1C). This increased identification of reagents is consistent with the reduced intermolecular electrostatic repulsion of proteins near their pI. The variation in the identified reagents at different pH values was moderate in comparison with the variations seen for precipitant composition. The ionic strength of the salt solutions used in the solubility screening exhibited the expected trend of increased identified reagents at higher ionic strengths (Fig. 1D). Although ionic strength had a small influence on the identified reagents, no trends were observable for the ionic composition of the reagents (Fig. 1E-F). The ions are displayed by increasing protein aggregation strength based on the Hofmeister series, and little variation in the identified reagents is seen for the cation or anion components of the reagent (15). A trend is also evident between the decreasing molecular weight of the target and an increasing number of identified reagents (Fig. 1G).

Fig. 3. Crystallization by using phase behavior information. Nine of the 12 targets were crystallized in free interface diffusion screening devices. (A) VCP/Vimp rhombohedral crystals were grown in 0.44 M sodium chloride, 0.1 M sodium citrate (pH 6.0), and 6% PEG 1500. (B) bR rod crystals were grown in 0.25 M ammonium chloride, 0.1 M sodium acetate, 0.1 M sodium citrate (pH 6.5), and 35% PEG 1500. (C) bR rod crystals were grown in 0.125 M potassium acetate, 0.1 M sodium citrate (pH 6.5), and 35% PEG 1500. (D) P450 1-12G plate crystals were grown in 0.28 M ammonium sulfate, 15% 1,3-propanediol, 0.1 M imidazole (pH 7.5), and 30% PEG 8000. (E) bR D85S rod crystals were grown in 0.125 M potassium acetate, 0.1 M sodium citrate (pH 6.5), and 35% PEG 1500. (F) SMC/ScpA hexagonal crystals were grown in 0.05 M magnesium acetate, 0.1 M sodium citrate (pH 5.5), and 60% PEG 5500 MME. (G) Fis1 crystals were grown in 0.1 M ammonium acetate, 0.1 M Tris-HCl (pH 8.5), and 20% PEG 8000. (H) 19S Lid rectangular prism crystals were grown in 0.15 M sodium acetate, 0.1 M sodium citrate (pH 6.5), and 15% PEG 1500. (I) Classification of crystal hits by the molecular weight of the precipitating agent in the reagent, the deviation of the reagent from the pI of the protein targets, the ionic strength of the reagent, and the ionic component of the reagent. The percentage of crystal hits is normalized by the number of reagents screened in each category. (Scale bars, 100 μm.)
Protein phase behavior was further characterized by the generation of complete phase diagrams for a subset of the reagents identified during the solubility screening. Empirically determined solubility boundaries allow for the selection of precipitant and protein concentrations that maximize the chance of successful crystallization by excluding areas of phase space that lie too far into the precipitation or soluble regions. The complete phase diagrams were produced by combining protein and reagent at 72 different protein and precipitating agent concentrations across the phase diagram and measuring protein aggregation at each position (Fig. 2). Positions where the protein aggregated above a threshold level were classified as insoluble, and positions where the protein did not aggregate above a threshold value were classified as soluble. The phase diagrams were used to design free interface diffusion experiments that target the solubility boundary of the protein with each reagent. For example, the phase diagram shown in Fig. 2A suggests that targeting the position corresponding to 6 mg/ml bR D85S and 20% (wt/vol) PEG 1500 would promote crystallization. A free interface diffusion experiment targeting this position requires combining 12 mg/ml bR D85S with 40% (wt/vol) PEG 1500 at a 1:1 protein to reagent ratio. As the free interface diffusion experiment proceeds, the trajectories for each solution will evolve to reach the desired position and promote protein crystallization (see figure 3 in ref. 16 for details). Customized crystallization conditions were designed by using this technique for the specific reagents identified for each of the targets.

Crystallization experiments were carried out with the individualized rational screens by using a redesigned free interface diffusion screening device (16). Each of the crystallization conditions was tested against the protein at five different mixing ratios to completely sample the solubility boundary of the associated phase diagram (see Materials and Methods for details). Free interface diffusion experiments with the rational screens resulted in a 75% crystallization success rate for the proteins, whereas previous crystallization attempts in the labs of collaborators failed to identify crystallization conditions for the majority of the samples (Fig. 3). For the successfully crystallized proteins, between 12% and 79% of the reagents screened produced crystals. The targets were bimodal in their crystallization success rates; the majority of the targets crystallized for a small subset of the reagents screened while the remaining targets crystallized for the majority of the reagents screened.

The crystallization results show that the crystal hits fall within a well-defined region of precipitant and pH (Fig. 3J). Among the precipitants used for crystallization experiments, more crystal hits occurred with larger precipitants, and no crystal hits were identified for the smallest precipitants. In contrast to the observed protein aggregation trends, more crystal hits were identified with the chemically modified MME precipitants for the larger precipitants. Although the variation in pH was larger, more crystal hits were identified from reagents near the pI value of the targets. No trend is observed in the number of crystal hits based on the ion identity of the reagent. As was observed for the protein aggregation results, the variation in crystal hits based on the pH was larger for the smallest precipitants. In contrast to the observed trends, more crystal hits were identified for the smallest precipitants. In contrast to the observed trends, more crystal hits were identified for the smallest precipitants.

Various crystal morphologies appeared in response to different reagents, including rectangular prism crystals, rhombohedral crystals, hexagonal crystals, rod crystals, plates, needles, dendrites, and spherulites. Optimization of the initial crystallization results with additional free interface diffusion experiments improved crystal size and morphology. The best crystallization conditions for each target were transported to larger experimental formats to generate crystals large enough for diffraction analysis (Fig. 4). The crystallization conditions were translated to scale-up diffraction devices (17), vapor diffusion format, or microbatch format with a 67% success rate. Successful translation to larger formats was also clearly correlated with the crystal quality attainable in the smaller format crystallization experiments.

Diffraction experiments were performed on the crystals grown in larger formats to evaluate the quality of the crystals. Of the six crystals successfully transported to larger formats for diffraction studies, 67% of these crystals produced diffraction data (Fig. 5). This corresponds to a 33% overall success rate of starting with purified protein and finishing with diffraction-quality crystals. The membrane proteins diffracted to between 6.7 Å and 16 Å, and the metabolic protein diffracted to 3.7 Å with a centered monoclinic unit cell of dimensions a = 19.9 Å, b = 62.2 Å, c = 188.7 Å, α = 90°, β = 118.91°, and γ = 90°. A point of concern is that although this method is highly successful at generating diffraction-quality crystals, the diffraction resolution has not been high enough to solve the structures. In contrast, the PSI has been quite successful in transitioning from diffraction-quality crystals to solved structures. Reasons for this discrepancy could include the inherent crystallization difficulty of the targets we attempted, quality control on the protein production side, and a lack of target optimization by the systematic truncation of floppy subunits. Additionally, further optimization of the crystal growth conditions by using tools such as small-molecule additive screening or temperature control...
could improve the diffraction resolution of the crystals. The crystal optimization efforts were directly proportional to the limited amount of material provided for each of the targets.

Our success rates of 75% for crystallization and 33% overall rate of producing diffraction-quality crystals from purified protein are roughly double those of the PSI, which are 38% and 18%, respectively (http://targetdb.pdb.org/statistics/TargetStatistics.html). One must, however, also take into account the relatively small sample size in the present survey ($N = 12$) compared with the large number ($N \sim 10,000$) tested by the PSI, which leads to a lower statistical significance in our observed rates. We have tried to mitigate this effect by choosing protein targets that are well above average in their crystallization difficulty; the vast majority of our targets failed conventional crystallization attempts. The size of the present data set notwithstanding, we believe the evidence is strong enough to suggest that microfluidic crystallization tools should become incorporated in large-scale structural genomics efforts.

In conclusion, the rational phase diagram-based crystallization strategy presented in this article was successfully used to crystallize diverse and challenging proteins. The use of solubility information to design customized crystallization screens doubled the crystallization success rate over traditional screening techniques and increased the production of diffraction-quality crystals. Microfluidic devices such as these consume small amounts of protein, are inexpensive, and are amenable to use in high-throughput crystallization efforts.

Materials and Methods

Protein Preparation. Protein samples were provided by the following collaborators: 98-kDa *Aquifex aeolicus* alanyl-tRNA synthetase (AlaRS) at 15 mg/ml was from Manal Swairjo (The Scripps Research Institute); 22-kDa human amelogenin (AMG) at 2 mg/ml was from Stefan Habelitz (University of California, San Francisco); 27-kDa *Halobacterium* bacteriorhodopsin (bR) and bacteriorhodopsin mutant D85S (bRD85S) at 18 mg/ml were from Marc Facciotti and Lee Hood (Institute for Systems Biology); 126-kDa *Rhodobacter sphaeroides* cytochrome cbb3 (Cbb3) at 10 mg/ml was from Michael Merckel (University of Helsinki); 16-kDa *Saccharomyces cerevisiae* Fis1 at 60 mg/ml was from Takumi Koshiba and David Chan (California Institute of Technology); 360-kDa *Saccharomyces cerevisiae* proteasome 19S Lid particle (19S Lid) at 25 mg/ml was from Xavier Ambroggio, Douglas Rees, and Raymond Deshaies (California Institute of Technology); 45-kDa *Mus musculus* P2X at 6 mg/ml was from Pavel Strop and Axel Brunger (Stanford University); 54-kDa heme domain of *Bacillus subtilis* cytochrome P450 mutant 1–12G (P450 1–12G) at 30 mg/ml was from Matthew Peters, Peter Meinhold, and Frances Arnold (California Institute of Technology); 123-kDa *Aquifex aeolicus* SMC/ScpA complex at 22 mg/ml was from Scott Gradia and James Berger (University of California, Berkeley); 72-kDa human transferrin/transferrin receptor mutant complex (Tf/TIR) at 27 mg/ml was from Anthony Gliammetti and Pamela Bjorkman (California Institute of Technology); and 254-kDa *Rattus norvegicus* VCP/77/Vmp complex (VCP/Vmp) at 10 mg/ml was from Byron DeLaBarre and Axel Brunger (Stanford University).

Solubility Screening and Phase Diagram Generation. Solubility experiments were performed by using a microfluidic formulator device developed by Hansen et al. (2004). Device fabrication, automation, experimental setup, and data acquisition were previously described (13). The formulator device was used to create 448 unique reagents for solubility screening by systematically combining 14 salt solutions, 4 buffering agents, and 8 precipitating agents. The salt solutions used were 4 M ammonium acetate, 3 M sodium acetate, 5 M ammonium chloride, 5 M sodium chloride, 2 M potassium sodium tartrate, 4 M ammonium sulfate, 7 M sodium nitrate, 2 M magnesium chloride, 2 M sodium phosphate, and 3 M potassium acetate. The buffering agents used were 1 M sodium phosphate (pH 4.5), 1 M sodium citrate (pH 6.5), 1 M Tris-HCl (pH 8.5), and 1 M 3-(cyclohexylamino)-1-propanesulfonic acid (pH 10.5). The precipitating agents used were 100% (wt/vol) PEG 300, 100% (wt/vol) PEG 350 MME, 100% (wt/vol) PEG 550 MME, 50% (wt/vol) PEG 1500, 50% (wt/vol) PEG 2000, 50% (wt/vol) PEG 3350, 50% (wt/vol) PEG 5000 MME, and 50% (wt/vol) PEG 8000. Salt solutions and precipitating agents were used near the maximal solubility and viscosity levels of each species. Each of the reagents was tested against the protein targets at six different protein and precipitant concentrations across the boundary of the phase diagram while holding the salt solution and buffering agent at a fixed level. The salt solutions were used at final concentration of 2.5% of the stock solutions (0.05–0.175 M; see above), and the buffering agents were used at a final concentration of 0.1 M. The six points tested were (i) 80% precipitating agent and 2.5% protein, (ii) 67.5% precipitating agent and 15% protein, (iii) 52.5% precipitating agent and 30% protein, (iv) 37.5% precipitating agent and 45% protein, (v) 22.5% precipitating agent and 60% protein, and (vi) 7.5% precipitating agent and 75% protein. A total of 2,688 solubility experiments were performed for each protein target. Protein aggregation was quantified by imaging the protein in solution with each reagent using a charge-coupled device camera, calculating the pixel intensity standard deviation, and subtracting the background pixel intensity standard deviation of the reagent in the absence of protein. Reagents that caused the protein to aggregate above 3 pixel intensity standard deviation

---

Fig. 5. Diffraction images of larger-format crystals. (A) P450 1–12G crystals were harvested by using 25% ethylene glycol as a cryoprotectant and diffracted to 3.7 Å. The unit cell is centered monoclinic with dimensions $a = 219.9$ Å, $b = 62.2$ Å, $c = 188.7$ Å, $\alpha = 90^\circ$, $\beta = 118.91^\circ$, and $\gamma = 90^\circ$. (B) Diffraction data for Cbb3 crystals were collected at room temperature by using scale-up diffraction devices, and the crystals diffracted to 14.5 Å. (C) Diffraction data for bR D85S crystals were collected through scale-up diffraction devices by using 20% glycerol as a cryoprotectant, and the crystals diffracted to 6.7 Å. (D) bR crystals grown in scale-up diffraction devices were harvested by using 15% PEG 1500 as a cryoprotectant and diffracted to 16 Å.
units for at least one of the six points tested were classified as potential crystallization reagents. For 3.3% of the images, the pixel intensity standard deviation value was negative because of errors in the formulator device, and for these measurements the pixel intensity standard deviation was set to zero.

Complete phase diagrams were generated for a subset of the crystallizing reagents identified with the solubility screening. For each phase diagram, the reagent was combined with the protein at 72 different precipitant and protein concentrations. The protein concentration was tested between 5% and 80%, and the precipitant concentration was tested between 2.5% and 77.5%. Again, buffering agent and salt solution concentrations were held constant throughout the phase diagram at the concentrations listed above for the solubility screening. Protein aggregation was measured at each point as described above. Positions on the phase diagram that caused the protein to aggregate above 3 pixel intensity standard deviation units were classified as insoluble, and positions below 3 were classified as soluble.

**Crystallization Experiments and Device Fabrication.** Initial crystallization experiments were performed by using a modified version of the microfluidic free interface diffusion screening device (16). The screening device design was modified to perform five free interface diffusion experiments per reagent instead of three. The protein:reagent mixing ratios used were 1:4, 1:2, 1:1, 2:1, and 4:1, with a total reaction volume of 25 nl per free interface diffusion experiment. One screening device was used to test 48 reagents for a total of 240 parallel free interface diffusion experiments. The previous design used a hybrid glass/elastomer device and was modified to consist solely of elastomer to facilitate faster fabrication. The devices were fabricated by using multilayer soft lithography (18). Two negative master molds were fabricated for the control structures and for the flow structures with photore sist by using conventional lithography techniques. The 20-μm-thick control structures were fabricated by spinning photoresist (MicroChem SU-8 2025) onto a silicon wafer at 3,000 rpm for 75 s and patterned by using a negative high-resolution transparency mask printed at 101.5% of the desired device size. For the flow structures, the 13-μm-thick reaction channels were fabricated by spinning photoresist (Shipley SPR 220-7) onto a second silicon wafer at 1,300 rpm for 70 s and patterned by using a positive high-resolution transparency mask. The reaction channels were rounded by annealing at 180°C for 1 h. Next, the 45-μm-thick reaction chambers were fabricated on top of the reaction channels by spinning photoresist (MicroChem SU-8 8:50) onto the second silicon wafer at 2,000 rpm for 45 s and patterned by using a negative high-resolution transparency mask.

Elastomeric devices were created by pouring silicone elastomer (General Electric RTV 615; 5 parts A:1 part B) onto the flow structures to 1-cm thickness and by spinning elastomer (20 parts A:1 part B) at 2,000 rpm for 60 s onto the control structures for a thickness of 30 μm. The flow layer and the control layer were partially cured at 80°C for 60 min and 40 min, respectively, before the flow layer was peeled off of the flow mold and aligned to the control layer. After the aligned device was cured for an additional 1.5 h, the device was removed from the control mold and the control and flow ports were punched with 20- and 14-gauge punching tools (Technical Innovations), respectively. A blank layer of elastomer (20 parts A:1 part B) was spun at 2,000 rpm for 60 s on a blank silicone wafer and partially cured for 30 min at 80°C. The aligned device was bonded onto the blank layer overnight at 80°C, and the device was completed by removal from the blank silicone wafer and placement on a glass microscope slide (Corning).

The crystallization experimental setup is described in ref. 16. Crystal hits were verified by using IZIT protein stain, the crush test, or diffraction analysis. Crystal hits were transported to microfluidic scale-up diffraction devices, vapor diffusion format, or microbatch format for diffraction experiments. Device fabrication and experimental setup with the scale-up diffraction device is described in ref. 17. Scale-up diffraction devices were incubated at room temperature up to 2 weeks, and cryoprotection was introduced by diffusion 24 h before flash-freezing the crystals in the devices. For vapor diffusion experiments, 1 μl of protein was combined with 1 μl of reagent and suspended over 100 μl of reagent and incubated at room temperature up to 2 weeks. For microbatch experiments, a layer of paraffin oil was placed over the microbatch wells, and 1 μl of protein was combined with 1 μl of reagent under the oil and incubated at room temperature up to 2 weeks. For vapor diffusion and microbatch experiments, the crystals were looped from the drops and through cryoprotectant before flash-freezing in liquid nitrogen.

**Diffraction Studies.** Diffraction data for Cbb3 were collected at station 11.1 of the Stanford Linear Accelerator Center (Stanford University), at an incident wavelength of 1.0 Å with a 15-s exposure and 1° oscillation. Diffraction data for bR and bR D85S were collected at station 8.3.1 of the Advanced Light Source (Lawrence Berkeley National Laboratory), at an incident wavelength of 1.0 Å with a 20-s exposure and 1° oscillation. Diffraction data for P450 1–12G were collected on an R-axis IV (Lawrence Berkeley National Laboratory), at an incident wavelength of 1.0 Å with a 30-min exposure and 1° oscillation. Diffraction data for P450 1–12G were collected on an R-axis IV (California Institute of Technology) at an incident wavelength of 1.5 Å with a 30-min exposure and 1° oscillation.

We thank the laboratories of Francis Arnold, David Chan, Douglas Rees, Raymond Deshaies, Pamela Bjorkman, Axel Brunger, James Berger, Stefan Habelitz, Michael Merekel, Lee Hood, and Manal Swirjo for generously providing the materials used in this work. This work was supported by the National Institute of Health Director’s Pioneer Award.