# Methods Testing in High-Throughput Crystallization

Sinem Ozyurt, Rich Romero, Laura Pelletier, Sonal Sojitra, Kevin Bain, Tarun Gheyi, Jason Adams, Fred Zhang, J. Michael Sauder, Spencer Emtage, Stephen K. Burley and Shane Atwell.
SGX Pharmaceuticals, 10505 Roselle Street San Diego, CA 92121

RECENT ADVANCES IN MACROMOLECULAR CRYSTALLIZATION September 23-26, 2007

## Abstract

The New York SGX Research Consortium (NYSGXRC), an industrial and academic consortium, is a Large-Scale Production Center for the NIH-funded Protein Structure Initiative (PSI). Crystallization is the rate-limiting step in the Target→Structure process for PSI proteins. Approximately 35% of purified proteins yield crystalline leads, and ~13% yield sufficient diffraction data to determine a structure. To further optimize the rate of successful structure determination, SGX continues to evaluate new crystallization techniques. Recently tested approaches include the following: protein buffer exchange, lysine methylation, seeding with protein precipitate and *in situ* subtilisin digestion. The effectiveness of these techniques ranges from 5-13%, as judged by the production of crystalline leads where none existed using the standard screening protocol. While apparently modest, a robust method for obtaining structures from an additional 5% of purified targets would represent a 40% increase in deposited structures. Herein, we describe our experimental protocols and analyze the effectiveness of each technique. The PSI project provides an excellent opportunity to test rigorously the effectiveness of crystallization techniques with large numbers of previously uncharacterized targets, as opposed to relying on anecdotal reports of individual success.

## Success Rates per Target

| | Targets | Success relative to previous step | Cumulative success rate |
|---|---|---|---|
| Selected | 6054 | | |
| Attempted | 3849 | | |
| Cloned | 3492 | 90.7% | 90.7% |
| Expressed | 2832 | 81.1% | 73.6% |
| Soluble | 2206 | 77.8% | 57.3% |
| Purified | 1230 | 55.8% | 32.0% |
| Crystal | 418 | 34.0% | 10.8% |
| Dataset | 228 | 54.5% | 5.9% |
| Structure | 166 | 72.8% | 4.3% |

July 1, 2005 through Sept 7, 2007

## Standard Initial Screening Process

1. A precipitation experiment is conducted using 20% PEG 6000 to determine if protein concentration is suitable
2. Proteins are first screened in 3 commercial, 96-well sparse matrix screens at 21°C
   - HR Index Screen (Qiagen Classics 2)
   - Qiagen Classics
   - Qiagen ComPAS (PEGS and Organics)
3. Multiple constructs are screened for each target
4. Medically relevant and Community-nominated targets (phosphatases, enolases and amidohydrolases) are screened with and without ligands or co-factors
5. Seeding of initial screens with crystal leads is employed to identify new or better crystallization conditions

## Standard Optimization Techniques

1. Optimization of precipitant concentration, salt concentration, buffer pH
2. Modification of protein buffer (especially removal of glycerol)
3. Seeding to facilitate crystallization and improve morphology (especially streak seeding)
4. Substitution of PEG (changing MW), salt or buffers
5. Variation of drop size or ratio
6. Additive or detergent screening
7. Variation of protein concentration
8. Increasing or substituting reducing agent

## Buffer Exchange: Initial Screenings

- Screening protein with and without glycerol present allows the protein to crystallize in either condition, and defines its stability and behavior
- 13 targets with unknown behaviors were tested in buffers containing glycerol (GOL+) and not containing glycerol (GOL-)
- 1 target yielded hits, leading to a structure

| Clone | # hits in GOL+ (with glycerol) | # hits in GOL- (without glycerol) | Results | Comments |
|---|---|---|---|---|
| 10024h2BCt1p1 | 0 | 6 | + | 2.5A structure in GOL- |
| 10017c3BCt2p1 | 0 | 0 | no change | |
| 10068b1BCt6p1 | 0 | 0 | no change | |
| 10076b3BSt1p1 | 0 | 0 | no change | |
| 10107f2BSt4p1 | 0 | 0 | no change | |
| 10124a1BCt1p1 | 0 | 0 | no change | |
| 10168e2BCt1p1 | 0 | 0 | no change | |
| 10183a2BSt9p1 | 0 | 0 | no change | |
| 10357g4BCt7p1 | 0 | 0 | no change | |
| 10398c2BCt12p1 | 0 | 0 | no change | |
| 10400h1BCt8p1 | 0 | 0 | no change | |
| 9335a1BCt5p1 | 0 | 0 | no change | |
| 9372f2BCt6p1 | 0 | 0 | no change | |

## Buffer Exchange: Optimization

- Rescreening targets in the absence of glycerol may result in better forming crystals or new morphologies
- 17 targets that had already yielded hits in glycerol (GOL+) were re-screened without glycerol (GOL-)
- Altogether, 7 structures were obtained, 5 from removal of glycerol

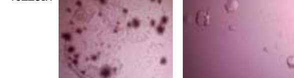
| Clone | # hits in GOL+ | # hits in GOL- | Results | Comments |
|---|---|---|---|---|
| 8663b1BCt2p1 | 1 | 8 | + | 1.9A structure from GOL-, no GOL+ crystals harvested |
| 10072b2BSt7p1 | 2 | 1 | + | no diffraction from GOL+ crystals, no GOL- crystals harvested |
| 10418f2BCt6p1 | 2 | 2 | no change | 2.8A dataset from GOL+, no datasets from GOL- |
| 10038i1BCt12p1 | 3 | 2 | - | 2.7A structure in GOL-, 2.2A dataset in GOL+ |
| 102281BCt7p1 | 1 | 1 | + | 3.7A dataset from GOL-, no datasets from GOL+ |
| 10117d1BCt6p1 | 3 | 3 | + | 2.6A dataset from GOL-, no datasets from GOL+ |
| 10201b2BSt7p1 | 3 | 1 | - | 2.8A structure from GOL+ (methylated) |
| 10015b1BCt8p1 | 4 | 9 | + | 3.0A structure from GOL+, no GOL- crystals harvested |
| 8702a4BCt2p1 | 4 | 5 | + | 1.9A structure from GOL-, no datasets from GOL+ |
| 10200d1BCt5p1 | 5 | 1 | - | 1.4A structure from GOL-, no dataset collected from GOL+ |
| 10119a1BCt2p1 | 6 | 3 | - | 2.2A dataset from GOL+, no GOL- crystals harvested |
| 10193b1BCt6p1 | 7 | 21 | + | GOL+ and GOL- crystals harvested, but no datasets |
| 10057f3BCt11p1 | 9 | 6 | - | 4.0A dataset from GOL+, no datasets from GOL- |
| 9229a1BCt9p1 | 14 | 13 | | 3.2A structure from GOL-, no datasets from GOL+ |
| 9364a1BCt8p1 | 19 | 14 | - | Crystals harvested, but no dataset collected |
| 10196a2BSt4p1 | 21 | 3 | - | 1.9A dataset from GOL+, no dataset from GOL- |
| 9356a1BCt1p1 | 40 | 19 | - | 2.8A dataset from GOL+, no GOL- crystals harvested |

## Buffer Exchange Examples

With Glycerol (GOL+) Without Glycerol (GOL-)

10200d1

10228a1



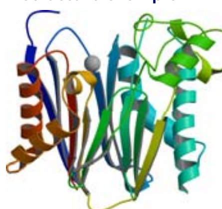
## Buffer Exchange: Conclusion

- The value of using multiple buffers in initial screening is unclear
- Buffer exchanging in optimization of existing crystals is clearly advantageous

## Buffer exchange Structure example

PDB ID 2q01(9229a1): glucuronate isomerase from Caulobacter crescentus

## Buffer exchange Structure example



PDB ID 2p8e (8702a4): the serine/threonine phosphatase domain of human PPM1B

## Lysine Methylation

- Lysine methylation has been shown to improve the crystallization of proteins (e.g. Structure (2006) 14:1617)
- Similar to the buffer exchange experiments, we selected 25 proteins that had never yielded crystals and 19 proteins that had crystallized but not yielded a structure, for lysine methylation
- Purified proteins were lysine methylated then re-purified using a size exclusion column and re-screened in 3 initial screens

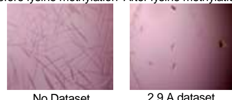
## Lysine methylation results

| | Proteins without hits | Proteins with previous hits |
|---|---|---|
| Total proteins | 25 | 19 |
| Hits with methylation | 5 | 10 |
| Structures | 0 | 2 |

## Lysine Methylation Examples

Before lysine methylation After lysine methylation

10069e2



No Dataset 2.9 A dataset

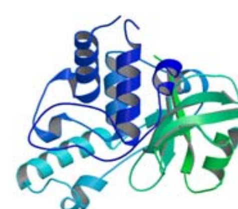
10201b2

No Dataset 2.8 A structure

## Lysine Methylation: Conclusion

- It is a useful technique for obtaining better diffracting crystals.

## Lysine Methylation Structure example



PDB ID 2p1g (10224c1): putative xylanase from Bacteroides fragilis

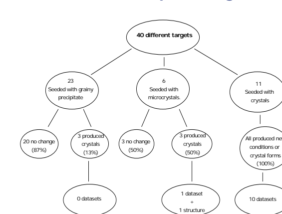
## Best Drop Seeding

- Microcrystalline precipitate can be used as a seed stock to get better quality crystals and diffraction
- February-March 2007: 40 targets were selected for seeding experiment
- A duplicate Classics 2, 96-well initials screen was set up for each target
- At Day 2, the best scored well of the duplicate tray was used to seed the remaining 95 wells

  -Best scored wells ranged from medium precipitate to actual crystals

  -Wells were streak seeded using a crystal manipulation tool
- The seeded and unseeded initial trays were compared at Day 7

## Best Drop Seeding

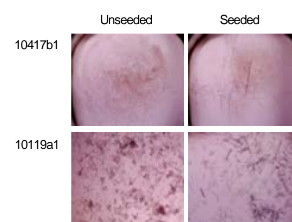



## Best Drop Seeding Examples

Unseeded Seeded

10417b1

10119a1



## Best Drop Seeding: Conclusion

- Useful technique for producing new leads
- Optimization of leads can be problematic
- Applicable for high value projects

## In Situ Proteolysis

- In situ proteolysis have been shown to improve crystallization (e.g. J Struct Biol (2003) 142:88-97)
- Screened 8 proteins with three different concentrations of Subtilisin A (2 μg/ml, 7 μg/ml and 12 μg/ml)

  Results

  5/8: did not produce any crystals

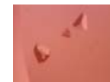
  3/8: produced crystals with Subtilisin A.

  - 1 produced sufficient quality data to solve the structure
  - Success rate: 1/8 (12.5%)
- Other PSI Centers (e.g., MCSG and NESGC) have had success with this technique as well

## In Situ Proteolysis Examples

10226a2

No Subtisilin A added



2 μg/ml Subtilisin A added

## Conclusions

- For initial screening, buffer exchange and best drop seeding are unproven techniques but are still of interest
- For optimization of existing crystals, buffer exchange yielded the highest success rate of the techniques that were tested

## Acknowledgements

- Special thanks to:

  Rina Fong, Nicole Glenn, Vandana Sridhar, Laura Stevens, Weidong Xu

  Work described herein was conducted as part of the PSI-2 project and funded by the NIH NIGMS

  Thanks to SGX Facilities, Bioinformatics, Molecular Biology, Fermentation, Protein Biochemistry,Mass Spectrometry, Crystallization and SGX-CAT groups

  Thanks to NYSGXRC collaborators
  - Steven Almo lab and Jeffrey B. Bonanno
  - S. Swaminathan lab

  Jeffrey B. Bonanno

  Yury Patskovsky

  Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.