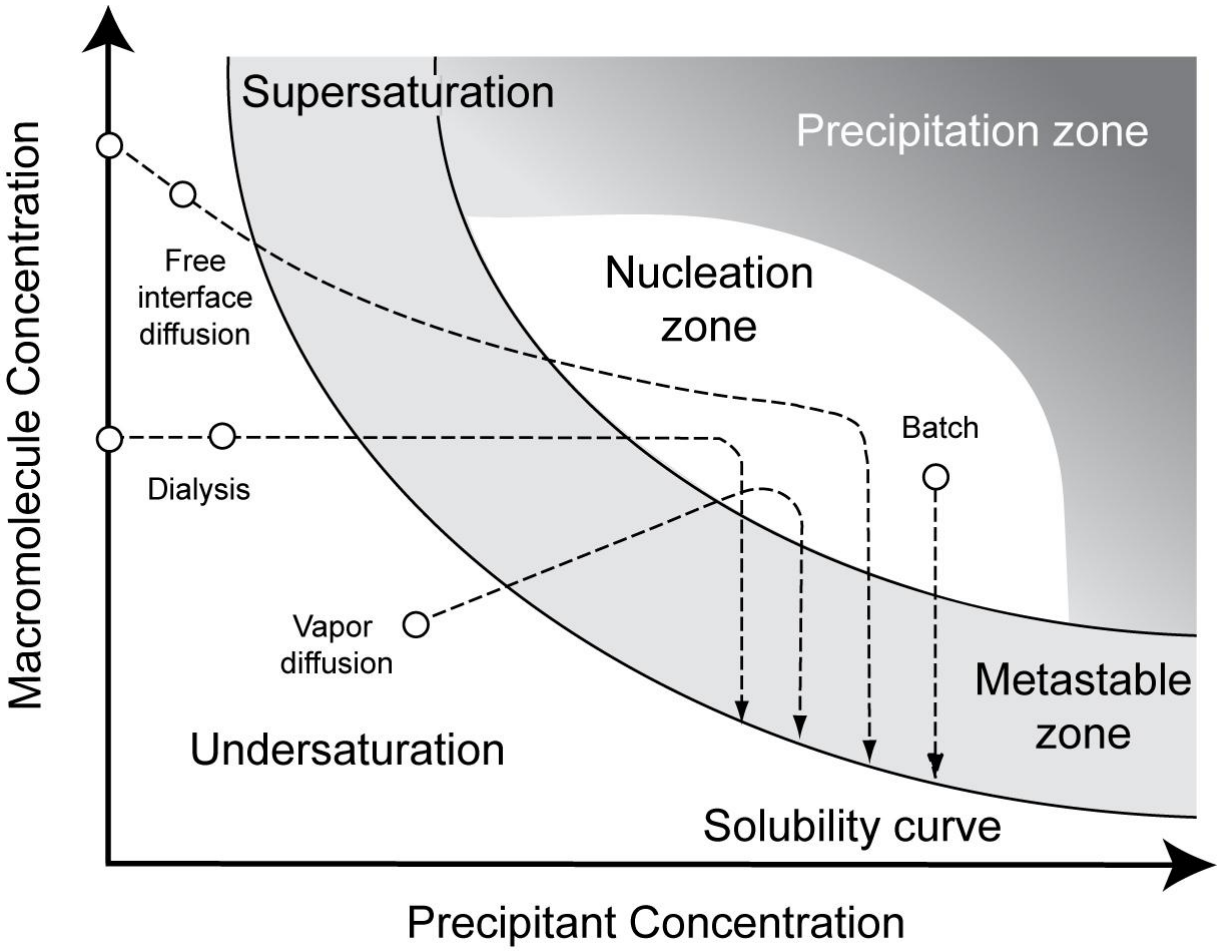
Order from Chaos

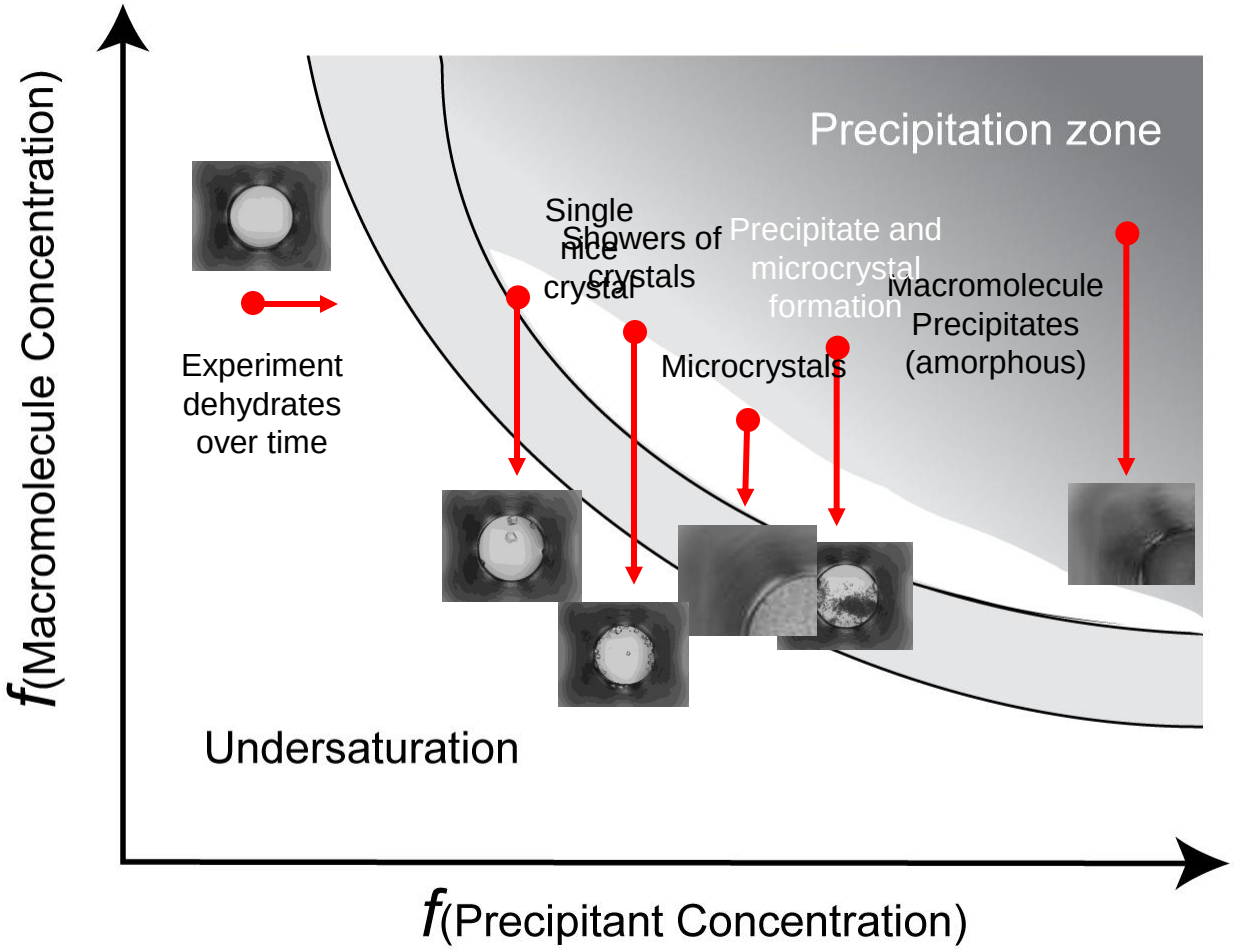
The design and interpretation of high-throughput crystallization screens to guide optimization

Edward Snell

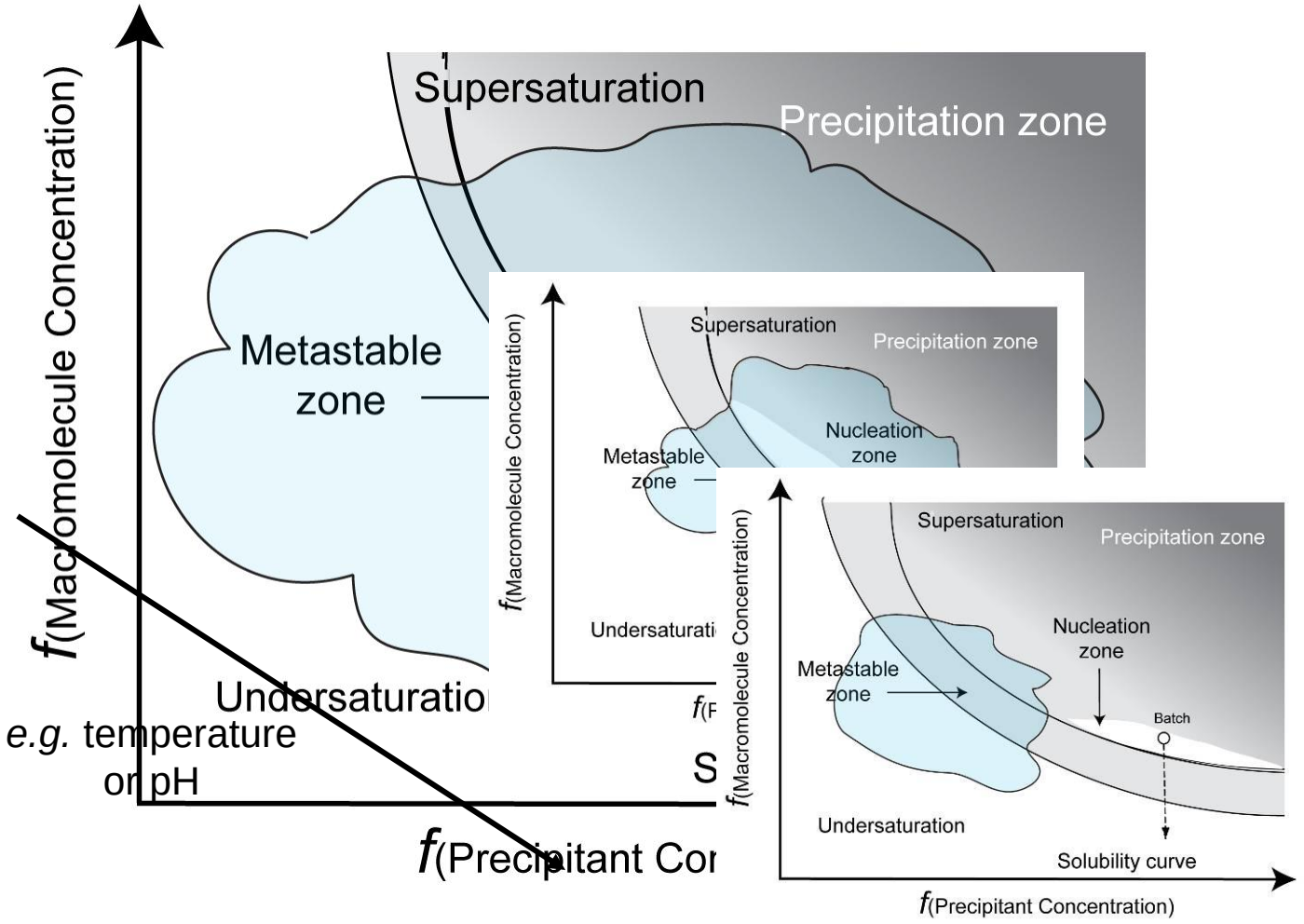
Simplified phase diagram for crystallization

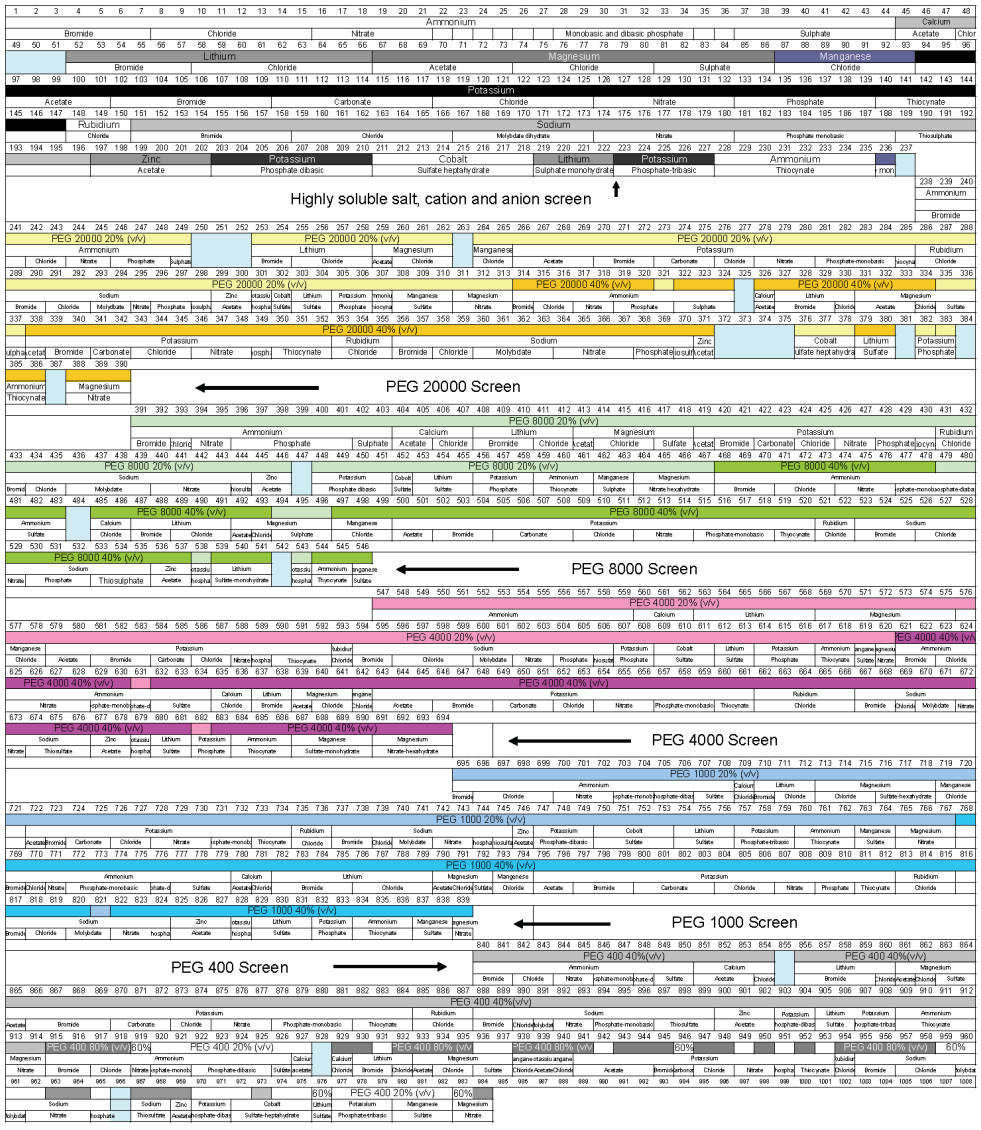


What results can we expect to see?



Typical situation, multidimensional area sampled

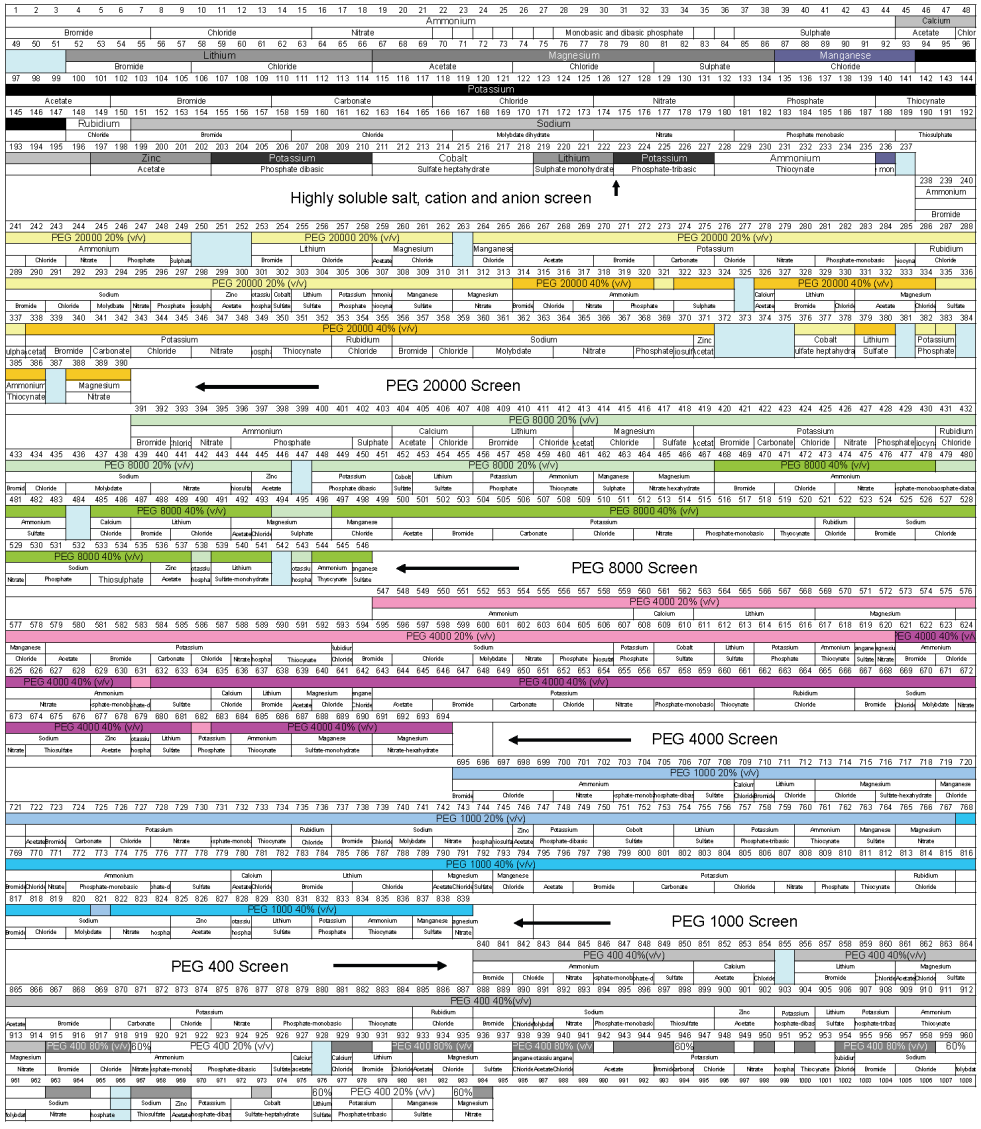




The HWI crystallization cocktail screen.

The 1536 diverse chemical cocktails (Luft et al., 2003). The 984 in-house conditions comprise an incomplete factorial sampling of 36 salts, eight buffers, and 5 different PEGs.

The remainder of 1536 cocktails are comprised of commercial screens available from Hampton Research. Specifically, in order of use; the Natrix Screen, Quick Screen, Nucleic Acid Screen, Sodium Malonate Grid, PEG/Ion, PEG 6000 Grid, Ammonium Sulfate Grid, Sodium Chloride Grid, HT Screen, Index and the SaltRx screen.

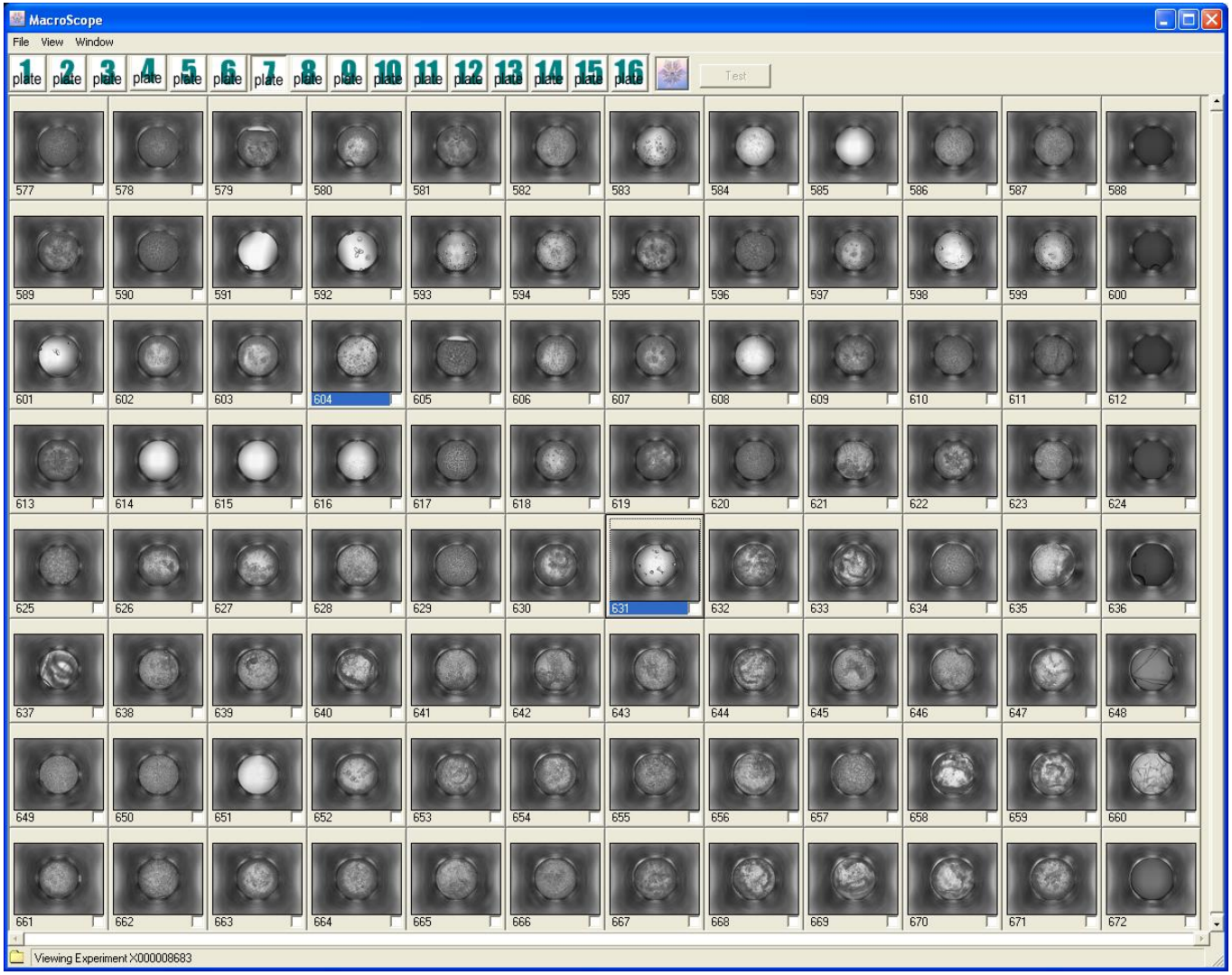


The HWI crystallization cocktail screen.

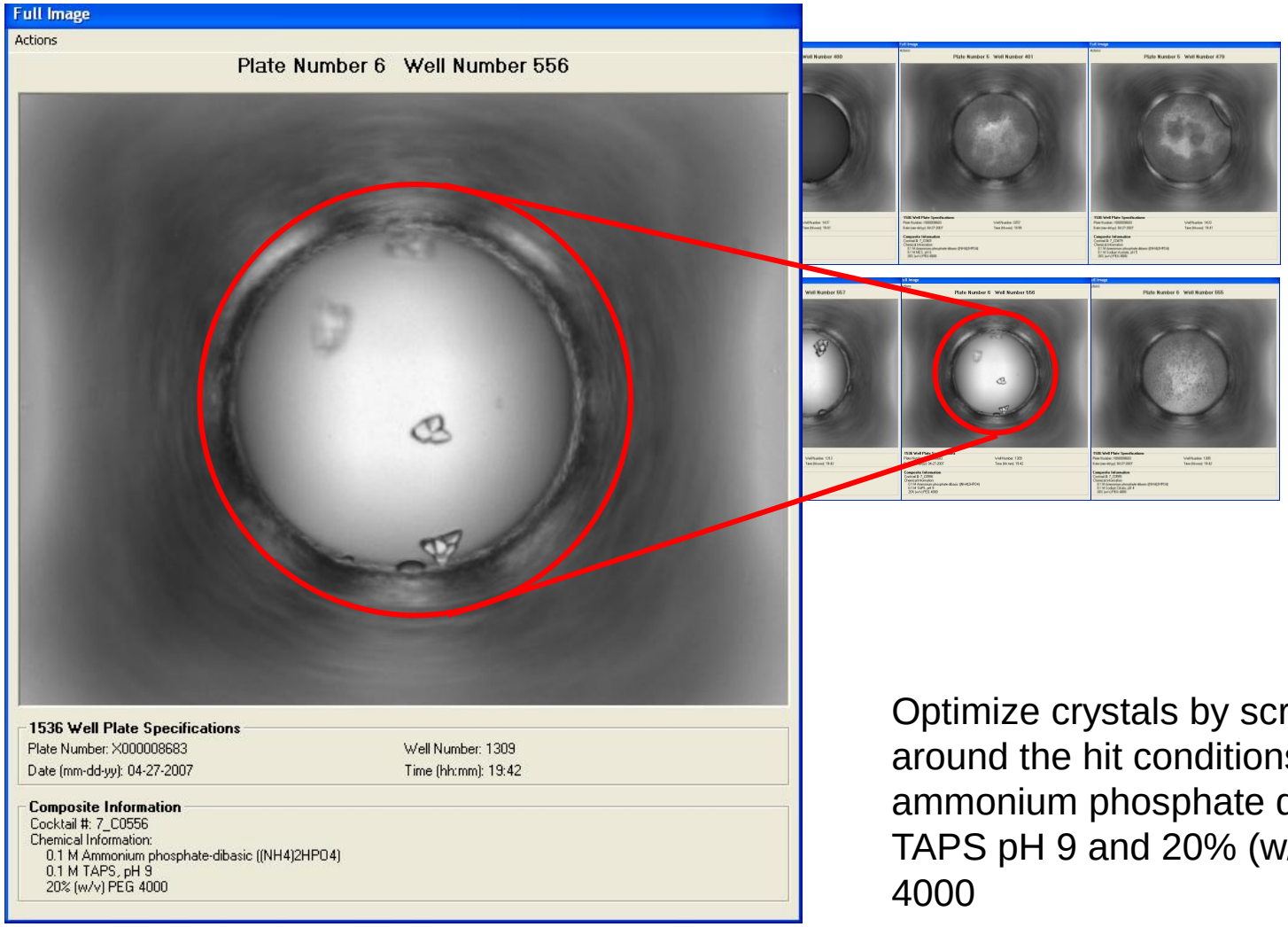
The 1536 diverse chemical cocktails (Luft et al., 2003). The 984 in-house conditions comprise an **incomplete factorial sampling** of 36 salts, eight buffers, and 5 different PEGs.

The remainder of 1536 cocktails are comprised of commercial screens available from Hampton Research. Specifically, in order of use; the Natrix Screen, Quick Screen, Nucleic Acid Screen, Sodium Malonate Grid, PEG/Ion, PEG 6000 Grid, Ammonium Sulfate Grid, Sodium Chloride Grid, HT Screen, Index and the SaltRx screen.

What do we see from the data?

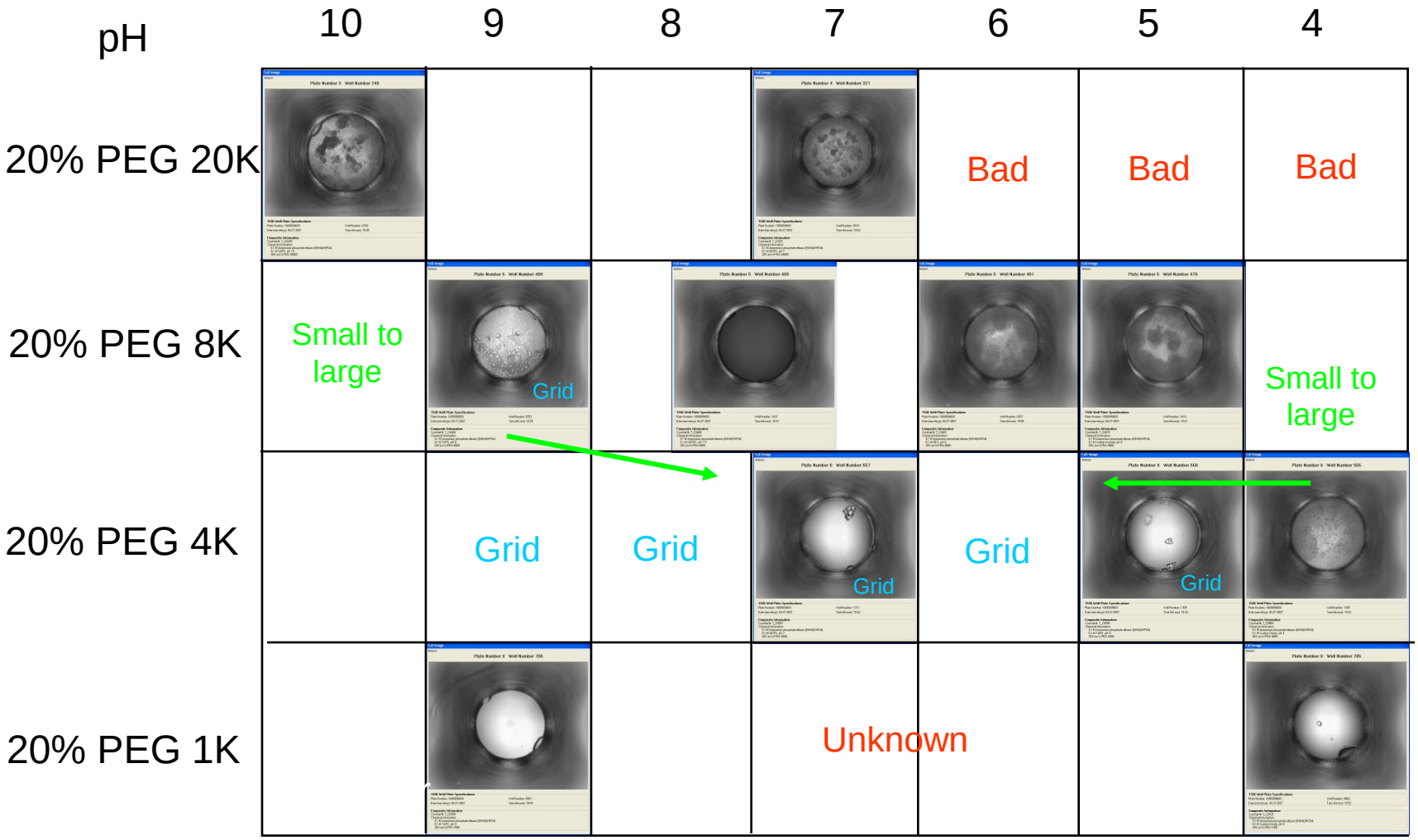


What do we actually see?



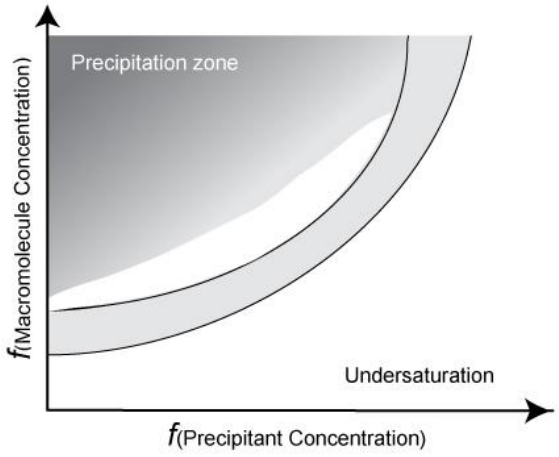
Optimize crystals by screening around the hit conditions, *i.e.* 0.1 M ammonium phosphate dibasic, 0.1 TAPS pH 9 and 20% (w/v) PEG 4000

If we plot the results in chemical space the road becomes clear

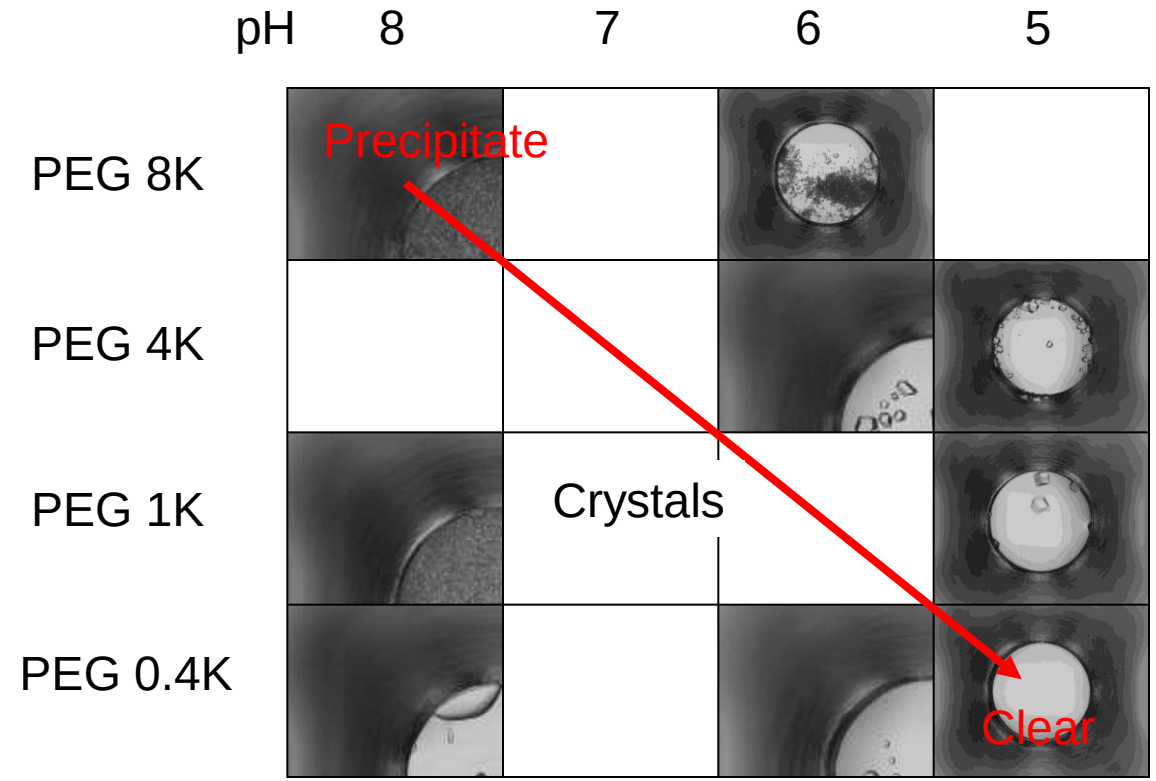


Chemical space provides a vector for optimization

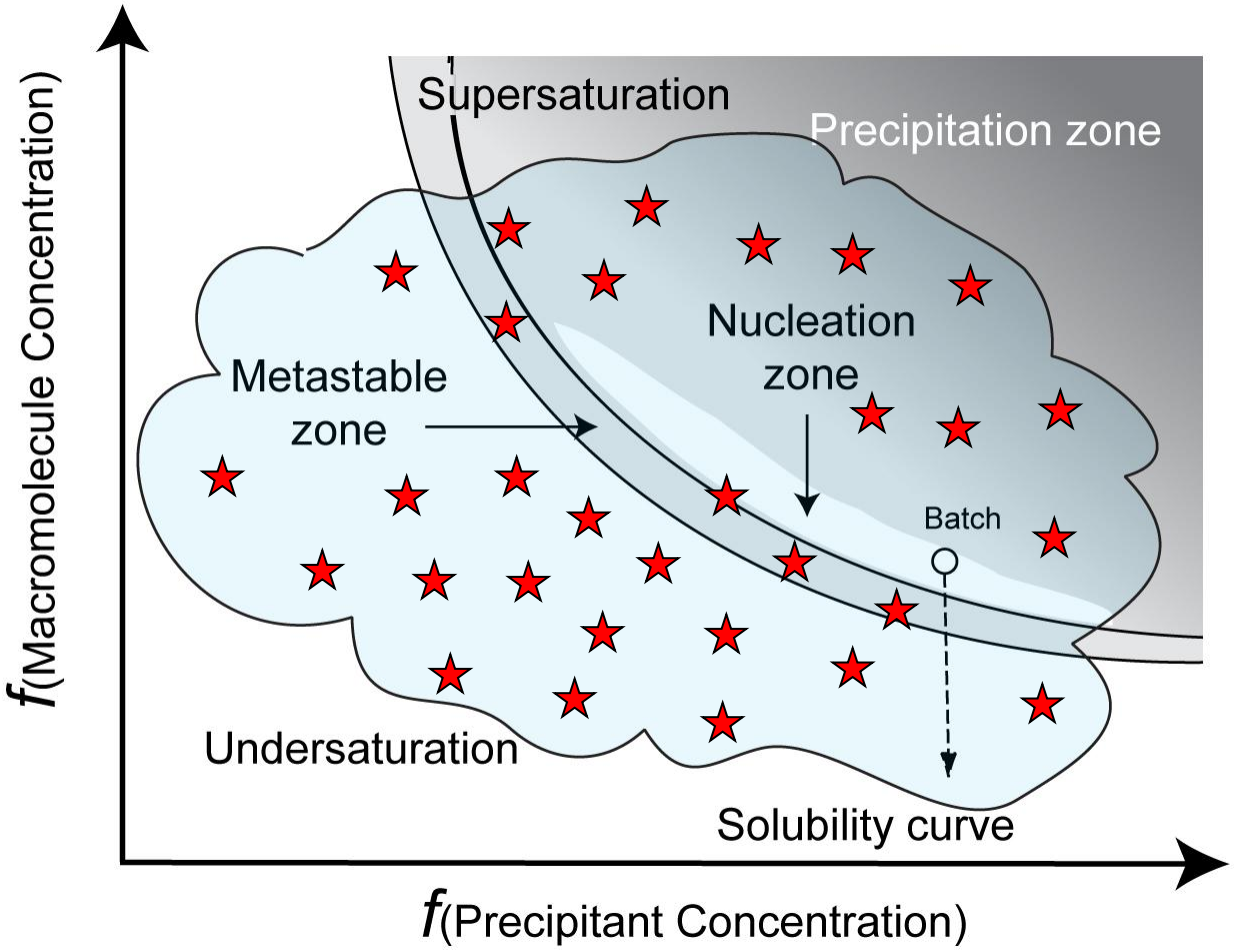
In this case the path from precipitate through crystals to clear is obvious. The phase diagram is reversed. Also clear are the number of chemical conditions that have not been sampled.



Ubiquitin, 40% PEG, 0.1M zinc acetate



It also illustrates the space we do not sample



We only sample discrete points within the sampling space

Numbers – the quantity of data

As of the New Year:

Over 850 laboratories sending samples (multiple samples from [illegible]rs)
Over 9,500 different macromolecules to date
Over 14.5 million images
Over 3000 years of computing time spent analyzing 1% of

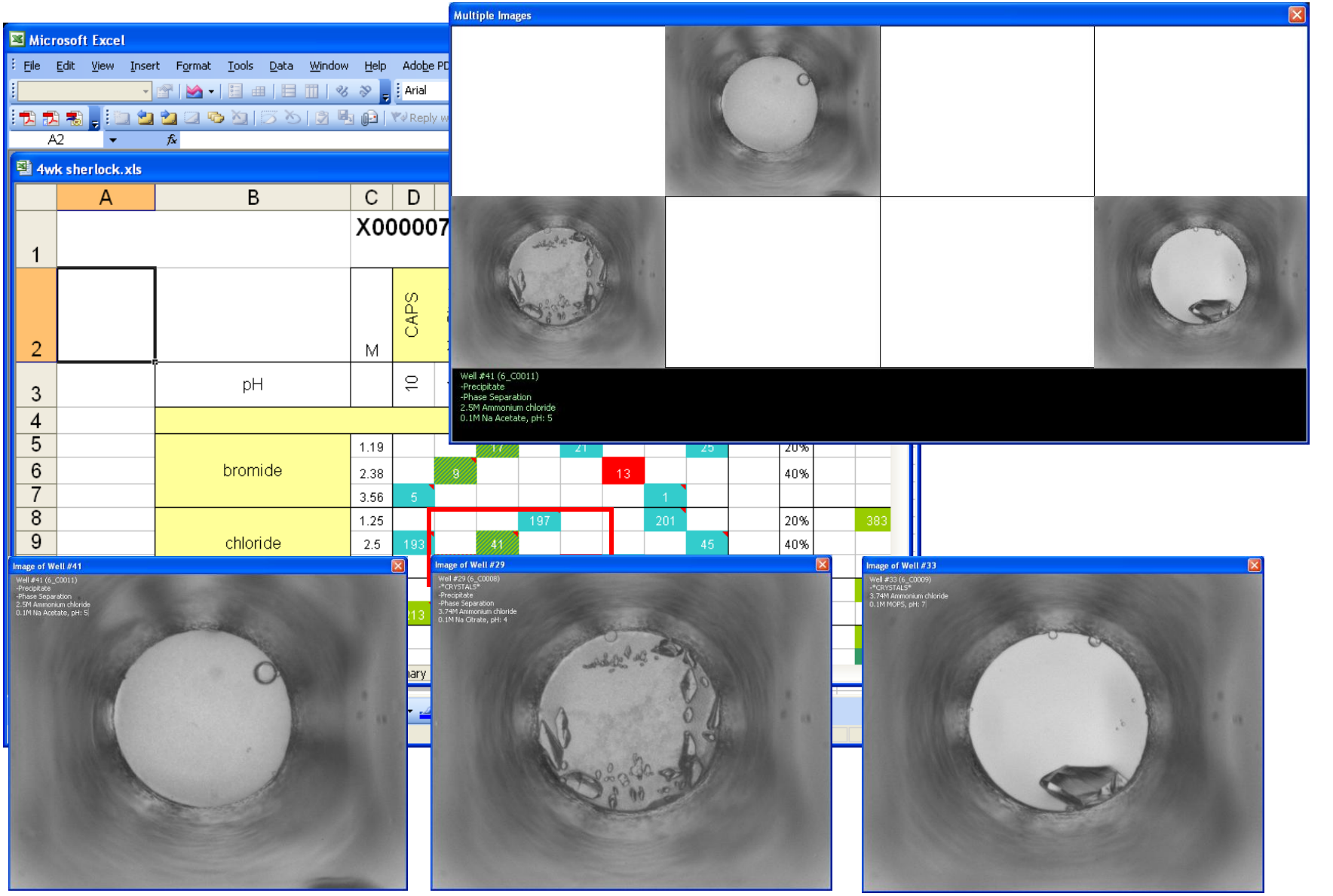
AutoSherlock™

A difficult task to easily visualize results.

"Because scientists would rather spend less time organizing their data, and more time learning from it."

Develop automated procedures.

Decreasing pH leads to crystallization. A large area of space along the crystallization pathway remains un-sampled. There are clear areas to pursue optimization.


Microsoft Excel
File Edit View Insert Format Tools Data Window Help Adobe PD
Arial
A2
4wk sherlock.xls
A B C D
X000007
M
CAPS
pH
10
bromide
chloride
1.19
2.38
3.56
1.25
2.5
Multiple Images
Well #41 (6_C0011)
-Precipitate
-Phase Separation
2.5M Ammonium chloride
0.1M Na Acetate, pH: 5
Image of Well #41
Image of Well #29
Well #29 (6_C0008)
-*CRYSTALS*
-Precipitate
-Phase Separation
3.74M Ammonium chloride
0.1M Na Citrate, pH: 4
Image of Well #33
Well #33 (6_C0009)
-*CRYSTALS*
3.74M Ammonium chloride
0.1M MOPS, pH: 7

Sherlock and Watson.

"We approached the case, you remember, with an absolutely blank mind, which is always an advantage. We had formed no theories. We were simply there to observe and to draw inferences from our observations"

Sherlock Holmes to Dr. Watson

Two pieces of related software under development;

- Sherlock to look at the individual 'crime', *i.e.* examine results from a single macromolecule
- Watson to tell the complete story, *i.e.* look at trends from many experiments.

Summary.

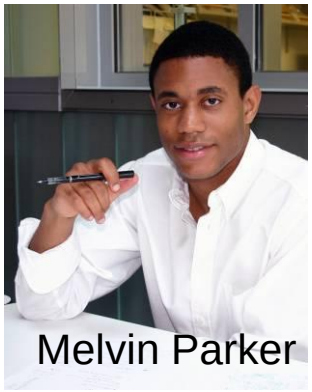
- No experiment should be considered in isolation.
- In crystallization screening when you have a sparse matrix, incomplete factorial or any other designed sampling of chemical space the results build up a picture of the crystallization landscape.
- An experiment with no crystallization hits that which generates both precipitate and clear conditions is promising when those conditions are separated by an un-sampled chemically sensible direction.
- You should know what crystallization conditions you examined but more importantly how those relate to those that were not sampled.

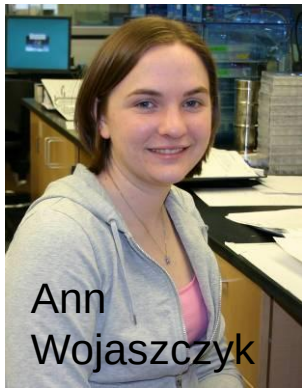
Acknowledgements

Ray Nagel for programming support

Joe Luft for discussions on the 1536 screen and analysis.

Melvin for the initial programming during his summer student project.

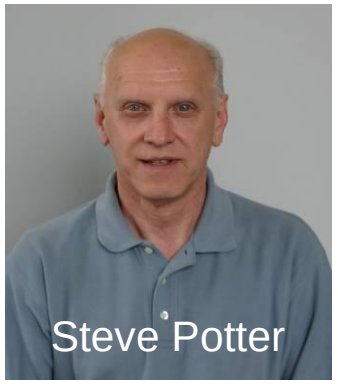




Ann, Miriem, Jen and Elizabeth for useful discussions, helping with the design of the chemical space, scoring images and testing the initial versions of the program.

Steve for multiple macromolecule data for testing and George for supporting the research.

Not to forget the rest of the High-Throughput Lab, Tina, Angela and Ellie for putting up with me. Thanks also to Dean Myles, Hugh O'Neal and Flora Meilleur for samples.



Crystallization screening service available to the general community, 600μl at 10mg/ml

http://www.chtsb.org/ or see Joe Luft here.