An integrated high throughput workflow for pre-formulations: Polymorph and salt selection studies

ERIC D. CARLSON
PEIJUN CONG
WILLIAM H. CHANDLER Jr.
HENRY K. CHAU
THOMAS CREVIER
PETER J. DESROSIERS *
ROBERT D. DOOLEN
CHRIS FREITAG
LAUREN ATAGI HALL
THOMAS KUDLA
ROLLAND LUO
COLIN MASUI
JON ROGERS
LI SONG
ANNY TANGKILISAN
KAREN QUYEN UNG
LUPING WU

Symyx Technologies Inc.
3100 Central Expressway
Santa Clara, CA 95051, USA

* Corresponding author

published by B5 srl
Via Cesare da Sesto, 10
20123 Milano (Italy)
Tel. 0039 02 83241119
0039 02 83241358
Fax 0039 02 8376457

INTRODUCTION

Pre-formulation, the selection of the form of an active pharmaceutical ingredient (API) most suitable for formulation, is a critical step in the drug development process. Recently, there has been considerable interest in reducing both the amount of time and the quantity of API needed to complete these studies including the use of high throughput techniques (1). One of the earliest reports of the use of high throughput crystallization of pharmaceuticals describes techniques for preparing arrays of salts in microtiter plates and screening for crystallinity using microscopy (2). More recently, reports describing the application of spectroscopic techniques to screen arrays of crystals in order to identify polymorphs have appeared (3). There are also reports of interesting techniques for preparing arrays of crystals (4) including polymer libraries (5), crystal nucleating chips (6), microfluidic devices (7), and acoustic injection of fluid droplets (8). In addition, methods describing the use of software for planning and analyzing the results of high throughput solid form screening have also appeared (9). Although each of these methods increases the speed or efficiency of individual steps in the pre-formulation process, none of these provide a truly integrated solution.

Symyx has been developing high throughput techniques for material science since 1995 and has created integrated workflows in such diverse areas as homogeneous catalysis, heterogeneous catalysis, polymer formulations, pigments, electronic materials, and nano-dispersions (10). An important part of all these workflows is an integrated hardware and software system that allows the throughput at each of the key steps in a process, (*e.g.*, synthesis, analysis, data storage, data retrieval, data manipulation, and reporting, *etc.*), to be matched thereby eliminating bottlenecks. Working in collaboration with Merck, we have recently developed a high throughout workflow that allows crystallization, salt selection, and polymorph studies to be completed in a fraction of the time and using a smaller quantity of API than conventional studies (11).

LIBRARY DESIGN

An essential part of a successful high throughput workflow is the rational design of libraries that cover as wide a range of variables as possible, followed by additional experiments on those variables that are found to be critical to a particular API. For both salt selection and polymorph studies the goal is to prepare and characterization as many crystalline forms of an API as possible. Because finding crystallization conditions that result in the formation of different forms is essential to both types of studies, the initial experiments should cover a broad range of crystallization conditions, including method (*e.g.*, cooling, evaporation, precipitation, and slurry), conditions (*e.g.*, time, temperature, rate), solvent, and in the case of salt selection studies, counter-ion. Varying counter-ion in one Cartesian coordinate and crystallization solvent in the other dimension of an 8 x 12 array generates 96 unique compositions that can be daughtered to allow three different sets of crystallizations conditions to be explored simultaneously (3 x 96 crystallizations = 288 crystallizations/design). For polymorph studies compositional diversity is generated using solvent mixtures (Figure 1). Library

designs are constructed using Symyx' Library Studio® a proprietary software program that allows the creation of complicated designs in minutes and stores to a database maps that describe the composition of each element and instructions that can be read by robotic automation software to execute the dispensing, heating, stirring, cooling specified in the design.

For both polymorph and salt selection designs, the initial choice of solvents is critical. Although extensive literature exists describing methods for the classification of solvents for chromatography including the use of principal component analysis to assess properties such as solvent polarity (12), polarity is just one of the variables known to effect crystallization. To address this, a table consisting of over a hundred solvents and numerical values of fifteen of their physical properties was compiled. Standard statistical clustering techniques were then used to create a hierarchy of solvent relationships from which it was possible to divide the solvents into any number of groups. For a given crystallization design requiring *n* solvents, the solvents are clustered into *n* groups and one solvent is selected from each different group.

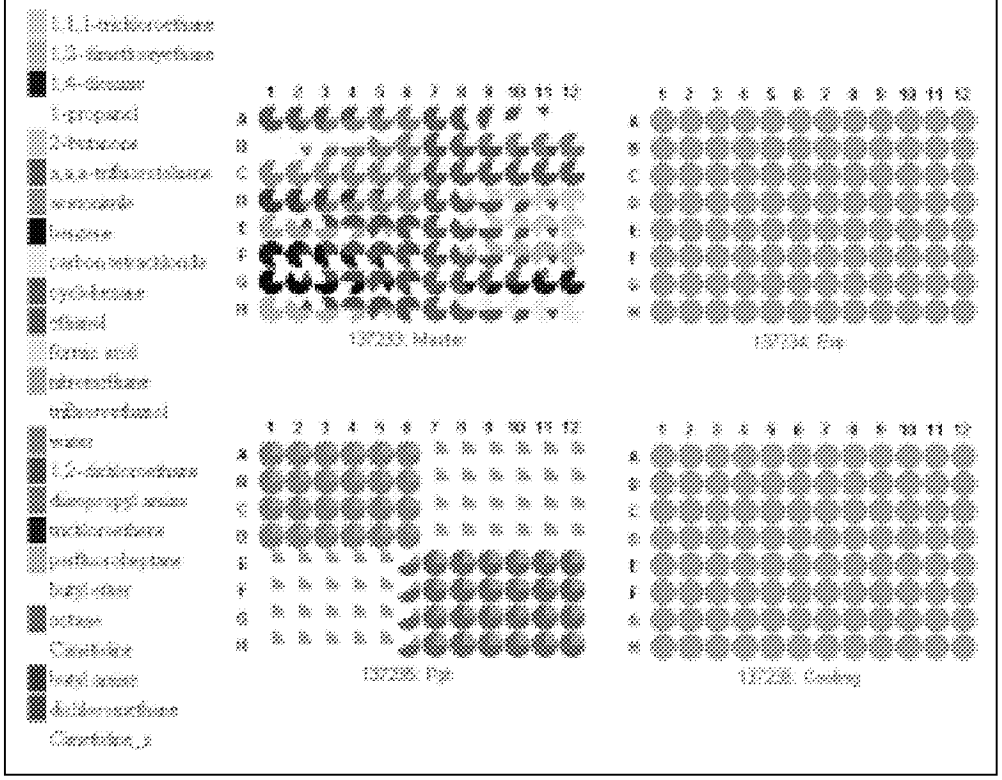

Figure 1 – A Library Studio® design used for polymorph studies contains twenty different solvents that are distributed to the master plate containing the API, equilibrated, filtered, and then daughtered to three different crystallizer assemblies allowing 96 evaporations, 96 precipitations, and 96 cooling crystallizations to be completed in a single day

CRYSTALLIZATION

Liquid handling robots suitable for dispensing solvents, heating, daughtering, and cooling arrays of solutions are readily available. There were, however, at least two significant challenges to developing a system for automated crystallizations. The first was hot filtration. All of the commercially available equipment for parallel filtration was deemed unsuitable for parallel crystallizations with multiple solvents due to a common headspace. A unique design developed for this workflow made it possible to create 96 isolated filters from a single sheet of filtration media, while at the same time eliminating cross talk in the vapor phase. The second challenge was to form crystals on a substrate that would allow all of the analyses to be performed without manual manipulation of the crystalline samples. Although optically transparent crystallization vessels compatible with birefringence and Raman measurements are commercially available, none is compatible with XRD since the depth of the wells prevented reflectance XRD measurements. The solution was to create an array of vessels by sealing a flat glass substrate to a Teflon coated block as shown in Figure 2. The crystallizer assembly functions as 96 isolated vessels that can be heated or cooled. After removal of the supernatant from the crystallizer assembly, the Teflon block can be removed leaving the crystals on a flat glass substrate.

The filtration and crystallization assemblies fit into a series of heating and cooling devices on the deck of a liquid handling robot (Figure 2). Crystallization Station™, proprietary Symyx software, allows the entire crystallization procedure including dispensing the crystallization solvents to an array of vials containing the API, heating the mixtures of API and crystallization solvents, filtering the hot solutions, daughtering the hot filtrates to three different crystallizer assemblies, cooling hot solutions, removing the solvents, sampling of aliquots for solubility measurements, and creation of a log file to be fully automated. A CCD camera with a telecentric lens mounted on a carriage underneath the crystallizer assemblies allows *in situ* birefringence images of the crystals formed to be obtained before the crystallizer assemblies are opened, facilitating the detection of unstable forms. Although the entire crystallization procedure requires 16 hours, the level of automation is sufficiently high it requires less than one hour of user time to carry out several hundred crystallizations.

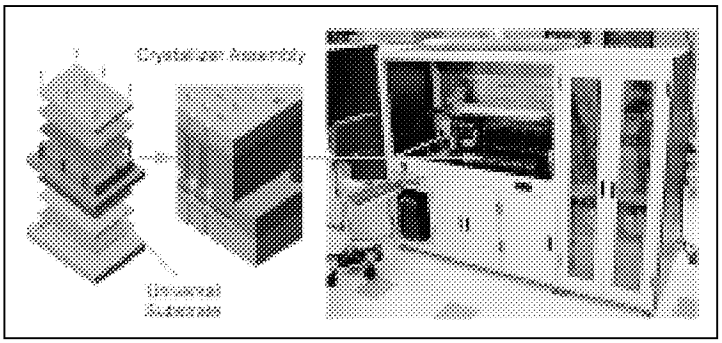

Figure 2 – The crystallizer assembly shown in expanded and assembled views provides a means of growing crystals that allows birefringence, Raman, XRD, and melting point measurements to be made without manual manipulation

SCREENING

The selection of analytical techniques is critical to an integrated high throughput workflow. In addition to requiring

the use of rapid serial or parallel analytical techniques to match the throughput of the crystallization station (288 crystallizations/day), it is essential that the techniques are capable of detecting and distinguishing different crystalline forms of the same composition. Although some pre-formulation systems might rely solely on Raman or XRD to screen for polymorphs, this was considered and deemed inadequate because no single analytical technique can be used to distinguish all unique crystalline forms. Thus, five complimentary analytical techniques were selected: solubility, birefringence, Raman, XRD, and melting point.

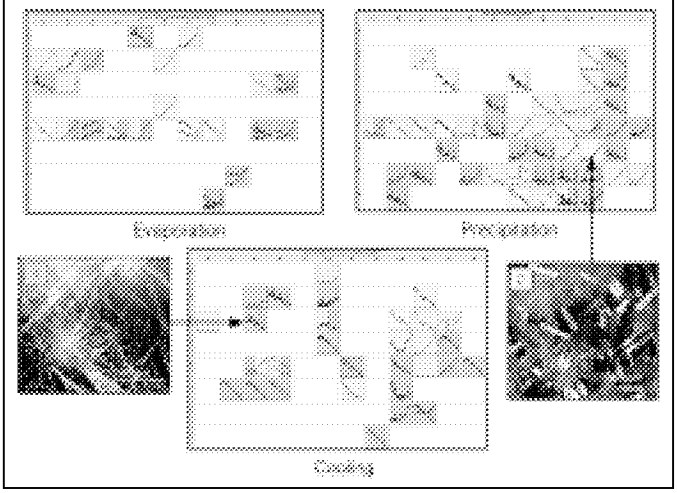

Figure 3 – Raman data for cimetidine sorted using Spectra Studio™ and displayed in library format shows three distinct groups. The starting form is shown in blue. Two other forms that were shown to be polymorphs are shown in green and yellow.

Solubility measurements are obtained by drawing aliquots from the solutions at various points during the crystallization. Removal of the solvents, dilution, and subsequent LC analysis allows the concentration and stability of the API in all of the solvents used in the design to be determined in a few hours. In practice the solutions are sampled immediately after filtration, providing the concentration of the API or its salts at high temperature, and at the end of the cooling cycle providing the concentration at low temperature. Epoch™ software controls the dilution, data acquisition, data analysis, and stores the data and acquisition parameters to the database.

Birefringence measurements are used to distinguish between crystalline and amorphous material. Higher resolution images are used to determine crystal habit and size. Epoch™ software controls the acquisition and stores the images to the database. The birefringence data for a set of 288 crystallizations can be acquired in less than three hours and requires less than twenty minutes of user time.

Raman spectra are acquired with a dispersive Raman microscope to distinguish different solid forms. Although several commercial instruments equipped with XY stages are suitable for analyzing libraries of crystals, software was needed to achieve the desired goal of an integrated workflow as the acquisition and analysis of hundred of Raman using the commercially available hard and software was prohibitively time consuming. Because the Raman focuses on single crystals rather than the entire sample, Epoch™ software is used to facilitate the location of crystals and the creation of a map, *i.e.,* a list of coordinates identifying the location of the crystals of interest. Once the coordinates and the corresponding images for an entire library are stored to the database, the software acquires the Raman spectra for an entire library in a fully automated manner and stores the spectra and acquisition parameters to the database. Sorting the tens to hundreds of spectra generated each day is accomplished using Spectra Studio™, a proprietary Symyx software program, that allows the chemist to load hundreds or even thousands of Raman spectra from the database, enter a correlation factor, and then sort the spectra into groups such that members of the same group have a higher correlation than the user-defined value. To the extent that the decision of whether any two spectra are the same or different is ultimately a subjective one, the software has a number of features that allow the user to overlay spectra, merge and split groups, reassign spectra, and also to display the groups as they appear in library format (Figure 3). This software reduces the time required to sort the Raman spectra obtained from a set or sets of crystallizations from several days to a few hours.

Although Raman is exceedingly sensitive (samples as small as 5 to 10 µg are sufficient) and the Raman spectra of the unique crystalline forms of an API are usually sufficiently different to be distinguished, this is not always the case. Indeed, we have encountered examples of unique crystalline forms that give indistinguishable Raman spectra. In addition, a small but non-negligible number of compounds fluoresce, thus requiring the use of other analytical techniques. XRD is inherently sensitive to changes in crystal packing. As was noted above for Raman, XRD instruments equipped with XY stages are commercially available although new software was needed to facilitate mapping and sorting. Epoch™ software facilitates the location of crystals, create maps, controls the data acquisition, and stores the images, coordinates, area plots, and 2 theta plots to the database. The time used to acquire an XRD pattern varies with sample size, but the hardware used in this workflow allows adequate signal to noise ratios to be obtained on samples as small as 100 µg in as little as three to five minutes, allowing the data from an entire set of 288 crystallizations to be acquired in less than a day.

Spectra Studio™ is used to sort the XRD data as described above for Raman. Due to the small sample size, non-statistical distributions of crystal orientations often result in dramatic changes in intensity for crystals of the same form rendering the direct XY correlation ineffective for sorting. The software was modified to allow the XRD patterns to be displayed as line spectra and then correlated based on peak positions.

Even XRD data are insufficient to distinguish all unique crystalline forms of an API. Iso-structural solvates in particular can have indistinguishable Raman and XRD patterns. Moreover, although Raman and XRD can be used to distinguish different forms, they provide no information concerning the identity of these forms, *e.g.,* solvates, hydrates, or true polymorphs, or their relative stability. DSC, TGA, and NMR are often used to provide this additional information in conventional studies, but technical challenges associated with making these measurements rapid serial or parallel caused us to explore other techniques. Using proprietary parallel birefringence technology, a parallel melting device was constructed. Initially designed to determine the melting point by recording the temperature at which the birefringence signal dropped to zero as the crystals melted, the equipment developed can also detect changes in birefringence attributable to solid-solid transitions such as polymorphic phase transitions and/or desolvation events. Typically run from 40 to 240°C at a ramp rate of 1°C/minute, a run is complete in less than four hours allowing data from an entire set of crystallizations to be acquired in less than one day. Epoch™ software controls the acquisition, data analysis, and stores traces of intensity versus temperature to the database.

POLYMORPH STUDIES

Cimetidine (1.00 g) was dissolved in methylene chloride (20 mL) and aliquots (200 µL, 10 mg cimetidine) were dispensed to an 8 x 12 array of 1 mL glass vials in an aluminum block. The solvent was removed by evaporation and the array was then transferred to the deck of the crystallization station and sealed with a gasket covered by a stainless steel cover with holes that allows liquid to be added and removed from the vials via a piercing needle. The crystallization procedure was executed using Crystallization Station™ that first prompted the user for the Library Studio® design number and the locations on the deck where the three crystallizer assemblies, two LC vial arrays, and 23 recrystallization solvents specified in the design were placed. The program then controlled the dispense of the crystallization solvents (800 µL/well) into the array of sealed vials containing cimetidine according the design shown in Figure 1 using a single tipped piercing needle. The program then activated heaters and the array of cimetidine and recrystallization solvents along with the filtration assembly, the four-tipped needle, and the crystallizer assembly used for the cooling crystallization was brought to 65°C and allowed to equilibrate for two hours. During this time the program dispensed anti-solvents (300 µL/well) to the sealed crystallizer assembly to be used for precipitations. After equilibration, the program executed a filter daughter sequence. Aliquots of the hot mixtures of cimetidine in the various recrystallization solvents (650 µL/well) were aspirated and then passed through the filters of the filtration assembly into an array of vials using the four-tipped piercing needle. The needles were washed to remove any traces of solids that may act as seeds, then used to withdraw the hot filtrates (550 µL/well) and to dispense aliquots to an open crystallizer for evaporative crystallization (200 µL/well), to the sealed crystallizer containing anti-solvents for precipitations (100 µL/well), to the sealed crystallizer heated to 65°C for cooling crystallizations (200 µL/well), and to the first array of LC sample vials (50 µL/well). The entire filter daughter sequence was complete in less than two hours. The program then executed the controlled cooling cycle specified in the Library Studio® design, 65 to 10°C over 8 hours. After equilibrating for one hour at 10°C, the four-tipped needle aspirated aliquots of supernatant from the cooling crystallizations and dispensed them to the second array of LC sample vials (50 µL/well). Finally, the remaining supernatants were aspirated from the precipitation and cooling crystallizations and dispensed to a third array of vials for recovery.

The following day, a camera mounted on a carriage underneath the crystallizer assemblies and controlled via software was used to acquire birefringence images of the 288 crystallizations, two rows at a time. The covers and gaskets were then removed from the precipitation and cooling crystallizers and twelve prong wicks were inserted across rows to prevent the last traces of solvent from evaporating on the crystals. After drying in air for four hours, the crystallizer assemblies were taken apart and the three universal substrates with crystals were removed and stored in cases for subsequent analysis.

The two arrays of LC vials were placed in a Genevac to remove the last traces of solvent, and aliquots of acetonitrile (500 µL/well) were added as a dilution solvent using the liquid handling robot of the solubility station. After sealing and shaking, the vials were transferred to the auto sampler of Agilent 1100 LC and Epoch™ was used to execute a sequence that allowed for the collection, analysis, and storage to the database of LC data for each of the samples. The first data set that was used to determine the solubility at the beginning of the crystallization (65°C) and the second set to determine the solubility at the end of the crystallization (10°C).

The three universal substrates from the crystallizations were then analyzed using birefringence, Raman, XRD, and melting point. The birefringence images were obtained the same day the crystallizer assemblies were taken apart. Epoch™ software prompted the user for the library number associated with each of the three substrates and then controlled the acquisition and stored the images to the database. The acquisition for each library was completely automated and required approximately 45 minutes/library.

While the birefringence images were being acquired on the second universal substrate, the first universal substrate was placed on the stage of a Jobin Yvon Raman spectrometer. Epoch™ software controlled movement of the stage and after locating crystals suitable for Raman acquisition, the software stored coordinates and images of the crystals to the database. The software controlled the acquisition of data using the positions stored on the database and stored Raman spectra and acquisition parameters to the database. The process was then repeated for the remaining two universal substrates. In total, 147 Raman spectra were acquired over the course of two days and required less than three hours of user time. The Raman data were then sorted using Spectra Studio™ in less than two hours and resulted in the identification of three unique types of Raman, one of which was the same as the commercial material (Figure 3).

Similar methods were used to obtain XRD data on a Bruker DX diffractometer and sorting of the XRD data using Spectra Studio™ confirmed the presence of three unique crystalline forms. Using the crystallization conditions from the automated crystallization experiments, samples of each of the three forms was

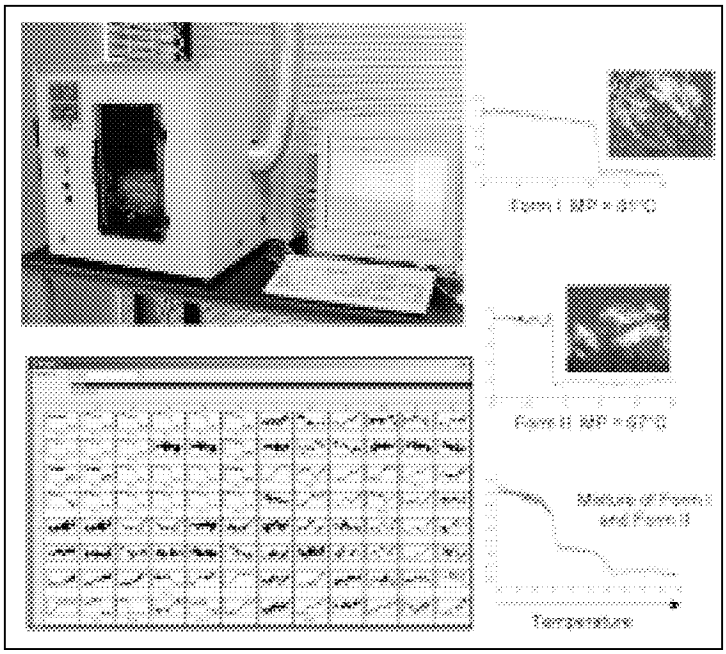

Figure 4 – The Parallel Melting Point Station™ used to analyze libraries of samples (upper left) measures changes in the intensity of a birefringence signal as a function of temperature (lower left). Traces corresponding to the samples containing pure Form I (upper right), pure Form II (middle right) and a mixture of the two forms (lower right) obtained during a polymorph study of nabumetone.

prepared on 50 mg scale and completely characterization by DSC, TGA, and NMR. Comparison of this data to the literature confirmed that the three groups identified from a single set of fully automated crystallizations and high throughput screening were true polymorphs of cimetidine (13).

When an identical set of crystallizations was carried out using nabumetone (1.00 g, 3.3 mg/crystallization) Raman and XRD data were consistent with the presence of two forms. The melting points determined using the parallel melting point device agree to within a degree of the melting points reported previously for the two polymorphs of nabumetone (14). As shown in Figure 4, the presence of mixtures can be readily detected using this technology. The formation and characterization of Form II is particularly noteworthy because it is unstable with respect to conversion to Form I and was first observed from capillary tube crystallizations.

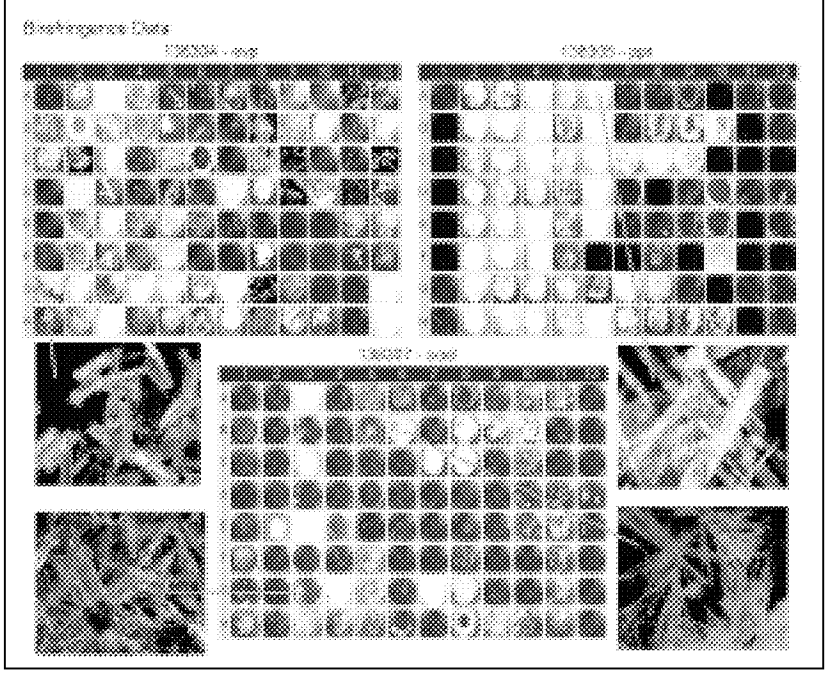

Figure 5 – The birefringence images taken during a salt selection study on naproxen illustrates the quantity and quality of crystals that may be obtained in a completely automated manner.

SALT SELECTION STUDIES

Dispensing solutions of either acids or bases across rows or columns of an array of vials containing API followed by reaction and removal of solvents, generates arrays of crude salts. Recrystallization and analysis using the same procedure described above for the polymorph studies makes it possible to rapidly identify crystalline salt forms. For example, the reaction of naproxen (10.0 mg/well) with stoichiometric amounts of seven different bases dispensed across rows and crystallization from twelve different solvents dispensed down columns using the procedure described above yields the crystals shown in Figure 5. Raman and XRD screening completed within days of the crystallization resulted in the identification of eight crystalline salt forms of naproxen including both hydrated and anhydrous forms of the sodium salt and two polymorphs of the tromethamine salt. A similar set of experiments carried out using ephedrine (5 mg/well) and eleven acids distributed by column and eight crystallization solvents distributed by row resulted in the identification of seven crystalline salts of ephedrine.

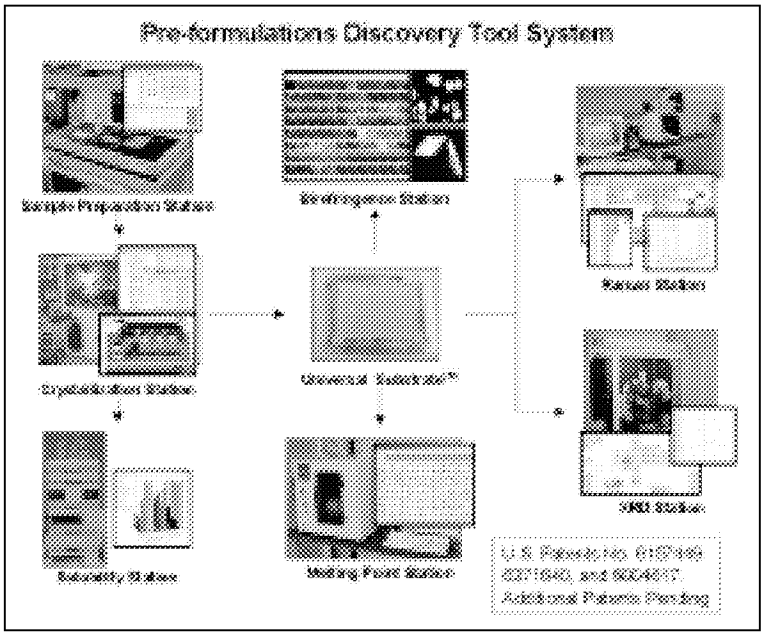

Figure 6 – The Pre-formulations Disocovery Tool™ System consists of seven stations linked by the Renaissance software suite and constitutes an integrated system capable of screening more than 50 API/year for salts and polymorphs.

That three polymorphs of cimetidine, the two polymorphs of nabumetone including an unstable form, eight salts of naproxen and seven salts of ephedrine were identified from four sets of crystallizations each using a gram or less of API illustrates the usefulness of the high throughput approach. The true value of the integrated system, however, only becomes apparent when it is realized that although each of the studies described above required a week, because each piece of hardware, including the crystallizer was used just one day during a given set of experiments, studies may be run concurrently allowing five such studies to be complete every week.

PRE-FORMULATIONS DISCOVERY TOOL

The hardware and software components along with proprietary searching (Polyview™), browsing (Renaissance Web Browser™), and reporting software (Polymorph Reporting Tool™), comprise an integrated system for pre-formulations referred to as the Pre-formulations Discovery Tool (Figure 6). Working in collaboration with ten of the world's leading pharmaceutical companies, including Merck and Eli Lilly who have since purchased this system, a team of six chemists has performed over 75,000 crystallizations encompassing crystallization, salt selection, and polymorph studies on more than 70 pharmaceutically active

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.