

The Application and Value of New Solubility and Prediction Modeling in Early Stage Pharmaceutical Process Development

An Industry White Paper

*By Phil Norris, Marketing Manager for Pharmaceuticals, AspenTech
Chau-Chyun Chen, VP of Technology, AspenTech*

Executive Summary:

A new and easy-to-use solubility model has been developed, which represents a significant advance in the accuracy of solubility prediction over currently available models. The value benefit from applying this modelling and prediction capability to characterization of New Chemical Entities' (NCEs') solubility properties in the "Lead Optimization" stage of drug discovery is potentially very significant to the pharmaceutical industry. Several leading practitioners in the domains of NCE solubility, best-practice crystallization, and crystal polymorph stability screening, to name but a few, are already using the model to drive efficiency, reduce costs, and facilitate better decision making in their drug development programs. The financial benefits may run into millions of dollars. Leading companies are already exploring the model in other areas of interest and challenge: excipient selection, solubility across pH ranges, solubility of New Biological Entities (NBEs), and selection of solvents for optimum chromatographic performance.

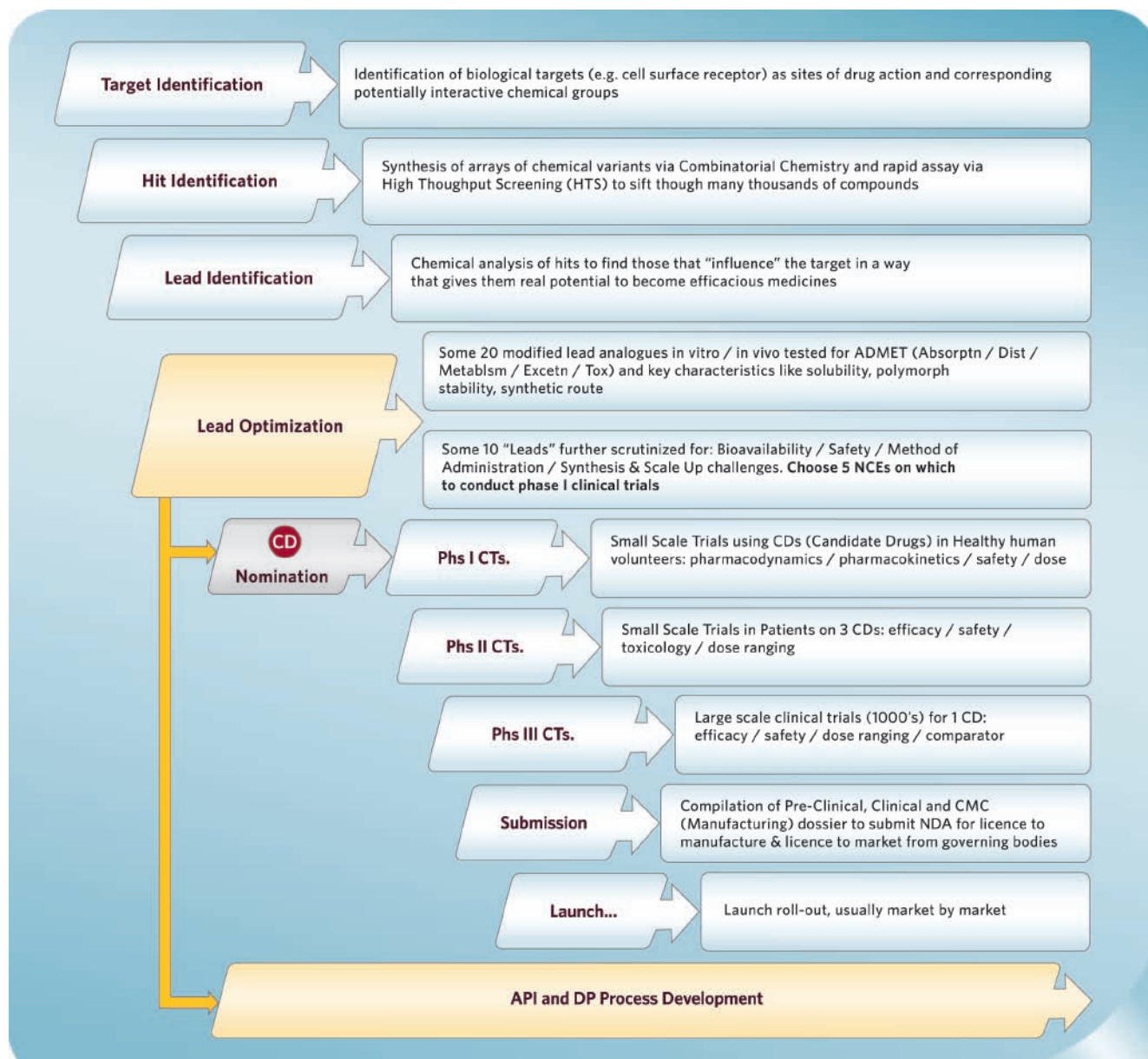


Figure 1: High-level drug development timeline, noting areas of application for solubility modelling and prediction (yellow).

The Solubility Challenge:

Many questions in drug discovery revolve around the issue of solubility. However, scant NCE material in late drug discovery means that the opportunity for experimental determination of solubility is rare, at best, and often rate limiting. This may delay the development of solubility characterization upon which valuable and far-reaching business decisions can be made, such as decisions relating to which NCE to take forward into formal development or how to process or formulate for manufacture.

Such “partially blind development” has the potential to lead to a huge research and development spend on candidate drugs that may turn out to be “unprocessable” or yield physical properties late in development (or even after launch!) that may ultimately delay or suspend market supply. Even in the recent history of drug research and development there are precedents for such events. The HIV protease inhibitor Ritonavir suffered a late emerging polymorph—ultimately a function of solubility and crystallization—that necessitated a reformulation after launch, temporarily impacting market availability.

As the number of NCEs is whittled down to a handful at the Lead Optimization stage (see *Figure 1*), solubility characterization becomes increasingly critical. Whereas aqueous solubility is a key and high-level determinant at preceding drug discovery stages, the impact of solubility at the “Lead Optimization” stage fans out into myriad critical areas. These are usually investigated by chemists operating at the earliest stages of API (bulk) and DP (formulation) process development. These critical areas are summarized in *Figure 2*, where they are divided into those areas under current and published investigation (blue) and those areas for which solubility modelling and prediction may also add considerable value (red).

Current Areas of Investigation / Application:	Future Areas of Investigation / Application:
<ul style="list-style-type: none"> ▪ NCE solubility in a wide range of solvents and solvent mixtures ▪ Crystallization process design (optimal solvent choice), including selection of anti-solvents (for crashing NCEs out of solution) ▪ Prediction of phase separation behavior ▪ Stable crystal polymorph screening ▪ Studies to determine ideal solvent for “solvent swapping” process <ul style="list-style-type: none"> ▪ Determine ideal solvent for water removal to prevent hydrates at distillation 	<ul style="list-style-type: none"> ▪ NBE (biotech macromolecules) solubility modelling & prediction ▪ Intermediate salt screening & selection ▪ Excipient selection ▪ Co-crystal screening ▪ Determine ideal solvent for spray drying in the formulation and process design of partly insoluble drugs (i.e. use of “amorphous” form of NCE to drive up bioavailability) ▪ Solubility modelling across pH ranges ▪ Determination of optimal carrier solvent in chromatographic separations

Figure 2: Critical areas of solubility characterization in early stages of process development.

Most of the solubility prediction and modelling value opportunities typically arise at that stage of drug discovery where there is considerable overlap of activity between late-stage Lead Optimization and early-stage Process Development. This alignment of overlap is illustrated in Figure 3.

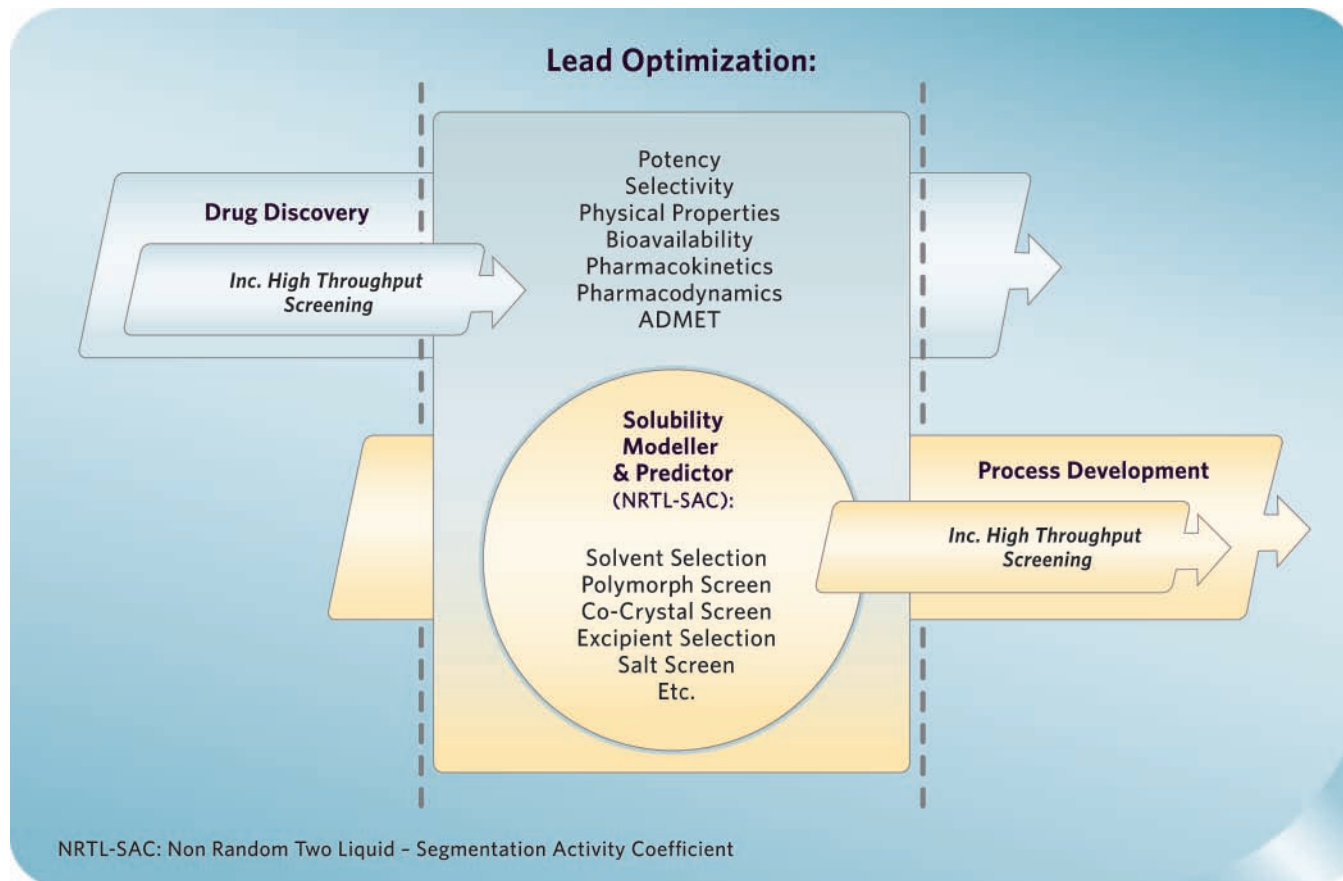


Figure 3: Solubility modelling and prediction at the lead optimization stage.

The Solubility Model:

To carry out solubility prediction, a scientifically sound and thermodynamically consistent mathematical model is required, and this needs to be embedded in a usable software tool (including relevant solvent property databases, associated calculation tools, display graphics, user interfaces, and user workflows.) The model described here is the NRTL-SAC (Non Random Two Liquid-Segmentation Activity Coefficient) model^{1,2} and its related model eNRTL-SAC (for salts). These models represent a significant advance in the accuracy of prediction over other currently available models. Its novel “conceptual segmentation” approach to predicting solubility makes it uniquely positioned to handle the complex NCEs and solvent mixtures that are typically used in today’s challenging pharmaceutical research and development. **The NRTL-SAC model here described is exclusively marketed and sold by AspenTech Inc. under the product name “Aspen Solubility Modeler” (or ASM).** Fundamentally, its design enables solubility predictions based on as few as four solubility experiments (on four solvents of varying hydrophobicity and polarity) using only tiny amounts of precious NCE. On running the results through the model, it will predict solubility profiles in myriad other solvents and solvent mixtures. As such, it represents a quantum step forward in the accuracy of prediction of NCE solubility. Its rigorous thermodynamic framework underpins its leading capability in predicting the NCE solubility profiles required at the earliest stages of process development, and thereafter throughout ongoing process development.

The application potential of this capability is extensive and some companies have already been active in applying the model to some of their most challenging areas, with exciting results:

- **Eli Lilly** scientists have applied NRTL-SAC to screen solvents for processing steps to maximize solubility and reduce solvent usage.³ The NRTL-SAC solubility predictions were first identified from ten data points in six solvents at four temperatures. They then developed a protocol to automatically evaluate solubility in 120 pure solvents and 122,000 binary combinations. These “virtual” experiments, which took only five CPU hours to complete, were then repeated at different temperatures and pressures. The NRTL-SAC predictions identified promising solvent candidates and conditions, which were then validated in the physical laboratory. A binary green solvent system, which has much enhanced solubility over the original solvent used in the lab, was chosen for scale-up. The study ably demonstrated the effectiveness of this solubility modelling technique. Eli Lilly has also carried out extensive work on modelling drug solubility to identify optimal solvent systems for crystallization,⁴ and is standardizing its use as part of their work flows in this area.
- Design of crystallization processes for the manufacture of API is a significant technical challenge to process research and development groups and an equally rich seam of value to pharma. **AstraZeneca** has examined the role of solubility modelling and its application within the crystallization process design framework.⁵ Through the case study on Cimetidine, NRTL-SAC has been demonstrated as a valuable aid to solubility data assessment and targeted solvent selection for crystallization process design. The model is becoming their standard way of selecting the right solvent for optimal performance in the process steps associated with API manufacture and crystallization in particular.
- **Bristol-Myers Squibb** researchers have reported a modelling strategy for optimal solvent composition selection in the design of a new API process.⁶ They have developed a modelling strategy for solvent selection and process optimization for API processes, including reaction, extraction, distillation, and crystallization. This modelling strategy helps them identify a solvent composition “sweet spot” for the design of their API processes.
- **GSK** researchers have developed an exciting methodology utilizing NRTL-SAC for high-throughput crystal form screening, with a view to understanding and characterizing the right solvent conditions to entice the most stable crystal polymorph to appear early on in process development.⁷

So What Is The Benefit?

At the highest level, and for any R&D-centric industry, the sooner improvements to new products (or the quality of decisions surrounding them) can be made, the greater the overall value potential. Value potential here is not just measured by value delivered to a product or process, but by “redundant cost avoidance” too. It is this that hits pharma’s “value sweet spot” square on, because pharma research is fundamentally much more “selection” than “instruction” in its nature, and so the biggest value impact may be felt in cost avoidance. This is where modelling and prediction come into their own, and can yield millions of dollars of accrued value in the course of subsequent R&D and throughout the lifecycle of the new drug thereafter. This is illustrated in *Figure 4*: a value plot against timeline to launch a new drug.

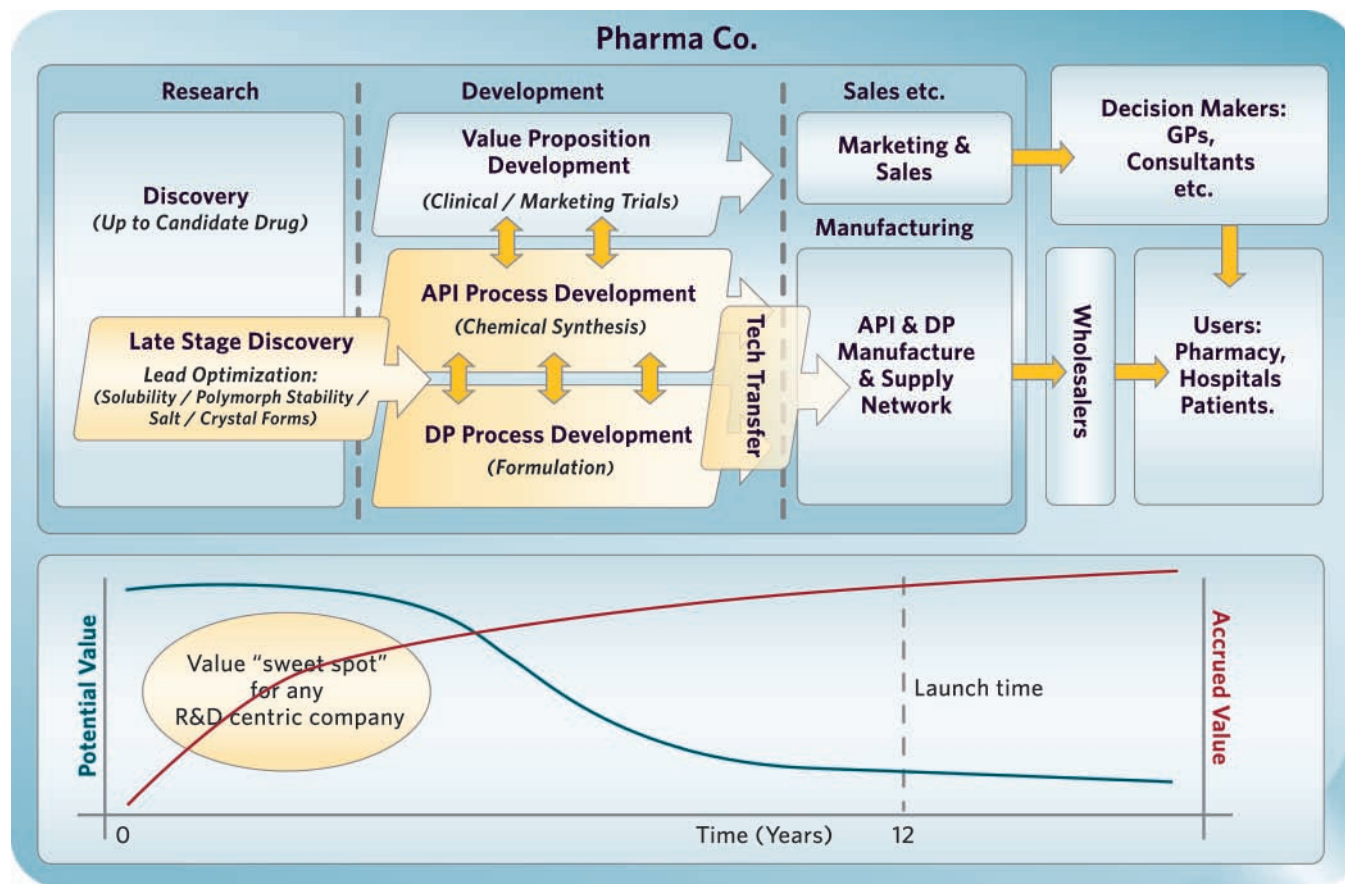


Figure 4: Value potential and benefit plot against timeline for a new drug, noting areas of application for solubility modelling and prediction (yellow).

Given the recent development of the NRTL-SAC model and the long timelines for product development required for a new drug, value benefit can only be a qualitative estimation, based on current applications and anticipated capabilities. Notwithstanding the inevitably “estimated” nature of value benefit, any potential benefit should also be seen against the backdrop of ever-increasing R&D costs.

In the ‘80s and ‘90s, relative R&D spending represented approximately 15 to 17 percent of revenue for the average drug company. Today, that average is approaching 20 percent and for some companies may exceed that level. Estimates have placed the cost of bringing an NCE successfully to market to be anywhere from \$700 million to \$1.2 billion over the course of 9 to 12 years of R&D. Some companies estimate that getting as far as completion of Lead Optimization requires a spend of some \$300 million over the first four to five years of research. With the discovery and development of high-value medicines becoming more and more difficult and NDA annual submissions on the decline, the time is fast approaching where dramatic operational efficiency improvements in R&D will be as much a central plank for competitive advantage in pharma as it already is for manufacturing operations.

With this in mind, the application of solubility modelling and prediction should add value in four major ways across any pharma R&D organization:

- 1. Efficiency Improvement:** By driving up the efficiency with which NCE solubility can be fully characterized and all the potentially advantageous effects that this can confer in the Lead Optimization space. Literally hundreds of “experimental hours” could be reduced to just a few through the use of modelling and prediction software. This value may manifest in decreased costs through headcount reduction or an increase in the throughput rate of NCEs in late discovery/early development. The latter is the likelier benefit route for companies with healthy pipelines of NCEs.
- 2. Risk Management/Better Decision Making:** By exploiting the predictive power of the model to drive more informed and earlier decisions relating to selecting the candidate drug to best progress with respect to its “processability” downstream in API and DP manufacture. This ultimately enables a more informed and better investment focus. Delaying or dropping candidate drugs exhibiting very significant process challenges could save time, money, and resources, or direct attention to solving “knock-out” issues first, before devoting more investment. Additionally, this should augment an “eyes open” approach to portfolio management of NCEs in early development with respect to their risk profile for manufacturability.
- 3. Speed to Market Launch/Continuity of Supply Post Launch:** By enabling aspects of process development activity (often delayed owing to insufficient NCE material) to proceed earlier. This can be achieved by using the modelling and predictive power of the software to sidestep this common cause of delay by using prediction to replace what would otherwise be experimentally derived process design data. This may translate into earlier clinical trials and possibly faster time to market. Further value may manifest by avoiding or reducing the emergence of unforeseen disasters downstream that may severely compromise launch times or continuity of supply after launch. (Late emerging crystal polymorphs—in which solubility characteristics of the active drug are permanently changed, and can impact dose form stability and even bioavailability—are a good example.)
- 4. API and DP Manufacturing Process Performance and Cost Profiles:** By enabling the informed design of many aspects of the API and DP manufacturing processes, such that the final developed process is better characterized, optimized, greener, and higher in yield, thereby reducing cost of goods from the outset of launch. Areas of application that could yield significant improvement are listed in *Figure 2*.

Looking Ahead:

In Lipinski’s thought-leading article on computational approaches to solubility in drug discovery,⁸ he states that, “*the knowledge of the thermodynamic solubility of drug candidates is of paramount importance in assisting the discovery, as well as the development, of new drug entities at later stages.*” Going forward there will undoubtedly be numerous opportunities and efforts to open up other areas of application for solubility prediction with NRTL-SAC. One example of particular interest under investigation is solubility modelling and prediction of biologically-derived or engineered macromolecules, such as monoclonal antibodies and genetically engineered proteins. This could be particularly exciting, as it appears that the segmentation nature of the NRTL-SAC model lends itself well to more and more complex chemical/biochemical entities. Another area of interest is solubility prediction in body fluid cavities of highly complex make up and therefore potentially, of bioavailability. This may well be within the model’s grasp too, and is particularly valuable to discovery chemists who may have only a few micrograms of NCE available, but may need solubility profiles in multiple fluid cavity types.

The NRTL-SAC model is already demonstrating its value to the pharmaceutical industry. It is certain to become one of the key tools in the prediction of solubility in early process development and thereafter in the development workflow. As such, it should drive efficiency, speed, informed risk management, and cost reduction into the increasingly complex process of new drug discovery, research and development.

References:

1. Tung, Tabora, Variankaval and Baken (Merck) and Chen (AspenTech Inc.): "Prediction of Pharmaceutical Solubility via NRTL-SAC and COSMO-SAC." *Journal of Pharmaceutical Sciences*, Web Publication, September 4, 2007.
2. Gani (Tech. Univ of Denmark), Jiminez-Gonzalez (GSK), Kate (Akzo-Nobel), Crafts, Jones and Powell (AstraZeneca), Atherton (Bristel), and Cordiner (Syngenta): "A Modern Approach to Solvent Selection." *Chemical Engineering*, March 2006.
3. Chen (AspenTech Inc.) and P.B. Kokitkar (Eli Lilly): "Modelling Drug Molecule Solubility with the NRTL Segment Activity Coefficient Model." Presented at the 14th Larson Workshop of Association of Crystallization Technology, Princeton, N.J., October 8-11, 2006.
4. Kokitkar and Plocharczkk (Eli Lilly), and Chen (AspenTech Inc.): "Modelling Drug Solubility to Identify Optimal Solvent Systems for Crystallization." American Chemical Society, Organic Process Research and Development: Web publication, March 4, 2008.
5. Crafts (AstraZeneca): "The Role of Solubility Modelling and Crystallization in the Design of Active Pharmaceutical Ingredients," Pages 23-85, Chapter 2 in *Chemical Product Design: Toward a Perspective through Case Studies*, ed. Ka M. Ng, R. Gani and K. Dam-Johansen, Elsevier, 2007.
6. Hsieh (Merck), et al: "A Modelling Strategy for Optimal Solvent Composition Selection in the Design of a New API Process." Presented at the AIChE annual meeting, San Francisco, Calif, November 12-17, 2006
7. Carino and Igo (GSK): "Application of Solubility Prediction Models in High Throughput Crystal Form Screening." Presented at the AspenTech Pharmaceutical Seminar for Process Development, May 16, 2007.
8. Lipinski C.A. et al, "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings," *Adv. Drug Delivery Reviews*, 23, 3-25 (1997).



||||| About AspenTech

AspenTech is a leading provider of award-winning process optimization software and services. AspenTech's integrated aspenONE™ solutions enable manufacturers to reduce costs, increase capacity, and optimize operational performance end-to-end throughout the engineering, plant operations, and supply chain management processes.

For more information, visit www.aspentech.com.

Worldwide Headquarters

Aspen Technology, Inc.
200 Wheeler Road
Burlington, MA 01803
phone: +1-781-221-6400
fax: +1-781-221-6410
info@aspentech.com

EMEA Headquarters

AspenTech Ltd.
C1, Reading Int'l Business Park
Basingstoke Road
Reading UK
RG2 6DT
phone: +44-(0)-1189-226400
fax: +44-(0)-1189-226401
ATE_info@aspentech.com

APAC Headquarters

AspenTech - Shanghai
3rd Floor, North Wing
Zhe Da Wang Xin Building
2966 Jin Ke Road
Zhangjiang High-Tech Zone
Pudong, Shanghai
201203, China
phone: +86-21-5137-5000
fax: +86-21-5137-5100
apac_marketing@aspentech.com