

—●— Profile —●—

An Interview with Kenneth C. Lewis, Ph.D., CEO, and Joseph D. Simpkins, Vice President, Software, OpAns (Optimized Analytical Solutions), Durham, NC

Ken Lewis co-founded OpAns, a contract analytical services and software company serving the preclinical therapeutic drug discovery market, in 2004 with Joseph Simpkins. Dr. Lewis' extensive background optimizing analytical chemistry for drug discovery prepared him to launch OpAns. He served in research scientist/investigator roles, as well as research advisor, for Lilly Research Labs and GlaxoWellcome, both in North Carolina's Research Triangle Park. As research advisor at Lilly, Dr. Lewis integrated analytical chemistry into lead generation by supporting medicinal chemistry, biology, compound handling, and screening. His leadership and vision led to a threefold growth in staff, while leveraging productivity gains to improve the efficiency and effectiveness of lead generation. Dr. Lewis was a postdoctoral fellow at Affymax Research Institute in Palo Alto, CA. He earned a bachelor's of science degree from Mercer University in Macon, GA and his Ph.D. in analytical chemistry from the University of North Carolina at Chapel Hill.



Kenneth C. Lewis (left) and Joseph D. Simpkins (right)

Joseph Simpkins co-founded OpAns, a contract analytical services and software company serving the preclinical therapeutic drug discovery market, in 2004 with Ken Lewis. As vice president for software, Simpkins oversees the management and development of OpAns software products. Specifically, Simpkins is responsible for the development of the analytical chemistry processing and workflow application toolset, OpAns Analytical Studio, a vendor-neutral liquid chromatography-mass spectrometry data processing, publishing, and reporting tool. Simpkins was responsible for the delivery and integration of the core processing algorithms for noise reduction, peak picking, baseline estimation, and compound determination as well as the plug-in architecture of OpAns Analytical Studio. Simpkins' background, with nearly 20 years' industry experience, includes informatics consultancy, management, and engineering at Eli Lilly and Company, GlaxoSmithKline Inc., Duke University, and Quantum Research Services. His experience includes designing and building large-scale software systems, including fermentation control, biological inventory, and desktop grid-computing. While at Lilly Research Labs in Research Triangle Park, NC, Simpkins designed and developed an automated high-throughput vendor-neutral system for retrieving, archiving, processing, and publishing liquid chromatography-mass spectrometry data to support drug discovery. While at GlaxoSmithKline, Simpkins designed and built the first web-based biological inventory system and the first large-scale desktop grid-computing system for supporting a variety of cheminformatics and computational chemistry algorithms using a common framework. He earned his Bachelor's of Science degree from Mississippi State University (Starkville, MS) in mechanical engineering and a Master's of Science degree from North Carolina State University (Raleigh, NC) in aerospace engineering.

In your experience, how would you describe recent trends in the quality (integrity/purity) of compounds being stored in chemical libraries for the purpose of medicinal chemistry and HTS applications? Based on your analyses, have you noted any general trends regarding sample purity?

Everyone impacted by *in vitro* screening is much more concerned about sample integrity than they were a decade ago, when few scientists worried about sample quality. Five years ago many in the industry thought they could analyze their compound collection, purify or discard the low quality samples, and have a pristine collection for years to come. Certainly, if a sample were not pure to start with, then it would not get better in the store. However, just because a sample is clean going into the store does not guarantee that it will remain pure and in solution. Most of the industry is now moving to a two-tiered approach in which there is some monitoring of the collection (usually of subsets, such as targeted screening sets—*e.g.*, nuclear hormone receptors, kinases, proteases, etc.) and then a higher level of analytical testing performed on the screening actives before structure–activity relationship work begins.

Is the definition of what constitutes appropriate quality assurance (QA)/quality control (QC) for chemical libraries becoming more standardized across the industry?

Yes, quality is becoming more standardized. For general collections as a whole, the focus is on determining whether samples are of reasonable purity. Purity is the primary driver, and identity is secondary. If a compound is not the correct molecular weight, and if it shows up as active in a screen, then the company will invest the time and resources to figure out the identity of the compound. If you consider, historically, why samples in collections might not have the correct molecular weight, a good percentage of the time it is because the structure of the compound was entered into the database incorrectly. Many other types of things can go wrong as well, such as use of the wrong reagent (the reagent supplied by a vendor may contain impurities).

So purity is the primary goal of QA/QC, followed by identity, and then concentration. Based on price pressures, companies will usually start with a less-detailed, survey kind of analysis based on purity. Then they may use higher quality liquid chromatography (LC)-mass spectrometry (MS) analysis for screening actives.

How have methods of compound storage and retrieval changed in recent years? Do you request a history of the sample's storage conditions from your clients?

The industry has mostly coalesced to a set of best practices, which include the following: use of dimethyl sul-

foxide (DMSO) as the solvent of choice; keeping the sample dry; access of the store as infrequently as possible; and the creation of small screening sets (10,000 compounds) instead of cherry-picking from the entire collection. Colder storage is better, but frozen samples have to be thawed, and if that is not done in an extremely dry environment condensation will occur, and that can result in water uptake during sampling. Water uptake in DMSO is known to exacerbate solubility and stability issues. So, I think the jury is still out on very cold storage—in practice is it better, and is it worth the cost?

In general, we do not get a history of storage conditions from our clients. However, it is very important to know how they want us to handle the samples so that we are consistent with their internal processes. Storage temperature, warming method, mixing, sonication, centrifugation, dilution, etc., are all very important for us to replicate their sample handling process.

What types of analytical methods/technologies does OpAns use to assess compound collections? How have these evolved, and what new and emerging technologies are being tested and implemented?

Our standard analytical platform is a “hyphenated” system consisting of HPLC (high-performance LC)-diode array UV (ultraviolet)-ELSD (evaporative light-scattering detector)-CLND (chemiluminescent nitrogen detector)-MS. We use four on-line detectors. The foundation of the system is good chromatography. Without it we cannot perform accurate purity or quantity assessments. We prefer to determine purity by a weighted average of relative peak areas from each detector. However, virtually all of our customers request purity information from the UV detector. While UV relative peak area is not the most accurate assessment of purity, it is the lowest common denominator (essentially all HPLC systems have a UV detector) and is consistent with the customers' historical data. Obviously, the mass spectrometer is used for molecular weight confirmation. The CLND and ELSD are used for quantitative assessment. The CLND can be extremely accurate and is fairly universal for pharmaceutical collections. Correction factors must be applied for compounds with nitrogen-nitrogen bonds, but these are reasonably well understood and straightforward. We provide our clients a tool to pass their structures through to calculate the nitrogen count for CLND response. That way they do not have to reveal the structures to us. The ELSD is used for quantification if the compound does not contain nitrogen or if the customer only wants “ballpark” quantification.

Over the last 10 years, there have not been any disruptive technologies brought to the marketplace. It has been a time of slow, steady enhancements. That is not to say that many technologies have not claimed to be dis-

ruptive, they just have not really delivered on that claim to the market. Here are my opinions on some of them (in chronological order):

1. *MUX™ technology from Micromass/Waters (Milford, MA)*. Clearly the ability to analyze samples in parallel should drastically increase sample throughput. Many people are successfully using this technology to analyze millions of samples per year. However, after nearly a decade on the market, it is still not the industry-dominated standard because many technical details have not been solved, such as:

- The lack of redundancy. If one channel, column, sprayer, etc., goes bad, the whole system is down. There is no way to automatically detect failures and go from running eight channels to seven channels to six channels. This multiplicative nature of the probability of failures greatly increases downtime over standard single channel systems.
- Incompatibility with very fast methods. The per channel data acquisition rate on these systems is an order of magnitude or more slower than for a single channel system. That means the chromatographic peaks must be wider in time in order to collect enough points across the peak. Typically, that means they must be used with 6-min gradients instead of 2-min gradients, effectively negating much of the gain in parallel productivity.
- Cost. These systems take up less space than the number of single channel systems needed to do the same amount of work, but the cost to purchase is about the same as the number of single channel systems needed to do the same amount of work. There is not much cost savings to be had.
- Lack of other parallel detectors (e.g., CLND).
- All of these issues could be overcome with R&D investment, but that investment has not been made and the market may not be large enough to warrant it.

2. *Nanostream (Pasadena, CA) parallel HPLC*. This is an enhanced LC platform that can do 24 channels at a time. Solutions to robustness issues have been successfully engineered. However, run times have to be much longer (five times) to get the same performance as good, fast HPLC. More importantly, there is not a good way to interface this system with a mass spectrometer.

3. *Corona™ Charged Aerosol Detector (ESA Biosciences, Chelmsford, MA)*. This detector has great theoretical potential for quantification without the need for a reference standard. However, it has suffered from the following limitations:

- The signal must be manipulated before being processed (it must be squared before it is integrated).

- The peak width is not compatible with fast chromatography.
- The linear range is not very wide.
- Overall, the CLND is a better detector for accurate high-throughput quantification without reference standards.

4. *UltraPerformance LC® (UPLC®) (Waters)*. Certainly no one is a bigger fan of higher-pressure chromatography than I. John MacNair (now at Merck) and I designed and built the first 100,000 psi HPLC system in Jim Jorgenson's lab before UPLC ever existed. I know the potential of the technology, and I know the hurdles of getting there. Clearly going to smaller particle sizes (<3 μm) and higher pressures has enabled significant reduction in gradient times and increased throughput. However, it has not been a disruptive technology because the rest of the analysis system could not keep up. Issues include the following:

- Overhead time between runs for downloading method parameters, closing files, etc., typically take 20 s on most LC-MS systems. That is a lot of wasted time on a 90-s analysis.
- Keeping the peak width narrow on the destructive detectors. Post-column band spreading becomes a significant issue when a half-second wide peak in the UV becomes 1 s wide in the mass spectrometer.
- Injection solvent effects reduce potential performance. Often library analysis is performed from DMSO, which significantly broadens peaks relative to ideal solvents (e.g., water/methanol). Small particles do not reduce this effect.

5. *CNLS D (condensation nucleation light scattering detection)*. This is a new quantitative detector that just came on the market this year. However, the technology will have difficulty with fast gradients and does not look to be applicable to this space.

How big a challenge for a company such as yours is educating your customers about what you do and what the results mean?

There is definitely a need for education. At the end of the day, most people want a single answer about their sample—is it a good sample or is it a bad sample; should I follow it up or not? We try to give them that answer. But there are many details to consider before getting to that answer. For example, all LC-MS is not the same. Many LC-MS practitioners do not think in terms of peak capacity and the ability to resolve chromatographically things that have oxidized. Often, in DMSO solutions, the sample will get oxidized—that is the most common form of degradation. Oxidation typically results in the formation of a double bond and the loss of two hydrogens. To determine the correct

purity and quantity that change should be chromatographically resolved.

When a company wants to know the purity of a sample, we have to talk about what level of chromatography they are willing to do to get that answer. The better the chromatography, the lower the purity will come back, because you will be able to see more impurities. But that will also take more time, and time means money. In our experience, a peak capacity of 60 is the minimum necessary to be able to see that you have something that has been oxidized. Peak capacity is based on taking the average 4σ peak width of an individual peak and determining how many peaks you can line up in the separation window, that is, how many things can you resolve in the separation time. Ideally, we like to operate at a peak capacity of 80–100.

Another area in which we have to educate is around the different ways to represent purity. Each type of detector will look at a sample differently and come up with a different purity number. The most accurate approach is to come up with a potency number—for example, if a powder is weighed out to create a 10 mM solution but when that solution is assayed it is quantified to be 8 mM, then the sample is only 80% pure. However, no one in discovery works off of potency. It only comes into play in active pharmaceutical ingredient manufacturing, not in early discovery.

In the end, we do not want people to accept blindly an answer of “yes” or “no,” the sample is or is not pure. We want them to dig deeper and understand what the results of the analysis mean. They need to understand how the results were derived, for example, what UV wavelength was used and why.

How big of an issue is protection of intellectual property in your relationship with customers?

It is huge; we never get structures. However, if we had access to structures we could better interpret the results.

What types of information do you provide to the customer? What are they asking for that may be atypical of what you normally provide?

We have yet to find a customer that can take advantage of all the information we can provide. We have to cherry-pick the information to meet their limited data handling capacity. Let me explain further.

While managers want to know about the overall quality of the compound library or screening actives, we fundamentally believe that scientists really want to know about the status of specific samples. The compound management scientist wants to know, “do I keep this sample or throw it out?” The medicinal chemist wants to know,

“do I follow up this compound?” These are specific questions about specific samples in the much larger collection of thousands of samples. Knowing that 80% of the collection passes the purity criteria does not answer those questions for the scientists. Some people address this issue by rerunning the samples in question using a better analysis method. That certainly works, but requires a second submission and lengthens cycle time. Furthermore, it does not address an equally common scenario such as the following: a 30,000-member kinase screening set is evaluated by LC-MS, and the results are returned. The medicinal chemists working in the kinase area will search that dataset and look for the compounds they synthesized. If their samples passed the purity criteria they will be happy. However, if their samples did not pass the criteria they will vocally question the validity of the results, and in some cases will mount a campaign to discredit the whole characterization effort. In this case, the data and your interpretation of them must be impeccable. The 80/20 rule does not work here. In fact, 99/1 is probably not good enough.

To address this, we built our own data analysis package called Analytical Studio. With it, we can take the raw data from any major MS vendor and process it extensively to determine whether the compound is present, what is its purity, and what is the quantity. Furthermore, we can quickly review the automated results and drill down into the raw data to ask questions and make the right call on a sample. Once the data have been reviewed and blessed, we are able to deliver the results in any way the customer desires.

From our perspective, we have all of the raw data, all of the processed data, all of the results, all of the meta data, and we can serve it up in almost any conceivable way. But virtually all of our customers are limited to receiving PDF and Excel files. It is sad because our customers lose the ability to dig into the data to answer their questions, and this would not be difficult for them to do if they could simply access the data that we can provide.

How can companies best protect against sample degradation over time?

Underlying this answer is the fact that dry powder stores are not compatible with modern screening operations and that newer dry film techniques are not mature enough to be industry standards.

In my opinion, the best balance between cost, throughput, and quality is as follows:

- No system is 100% foolproof. Some samples will always degrade, so don't kid yourself into thinking that by implementing an elaborate system up front that degradation will not occur. It will.
- Dissolve in dry DMSO.

- Make every effort to keep samples sealed and work in a dry environment.
- Storage temperature is difficult to assess. Certainly, long-term stability is increased by storage at colder temperatures. However, solubility can be negatively affected by freeze-thaw cycles. The decision between room temperature, 4°C, or -20°C is a balance of length of storage, number of times the sample is accessed, and available funds.
- Limit the number of times a sample is accessed (preferably <10) before it is consumed or discarded.
- Implement a compound characterization strategy. At a minimum, all samples being followed up on after 50% inhibitory concentration determination should be characterized.
- Track aliquot storage conditions. Once submitted, a single lot of a compound is usually divided into many different containers. Often, there is no distinction between these containers. When a sample is requested, the scientist does not know if it came from a vial that has been at room temperature for a year and accessed 100 times or if it was the first sample from a -20°C inert store.
- No process is perfect. The key is to be “in control of your process.” Being in control will avoid costly assumptions and mistakes.

What types of QC methods should companies use to monitor the quality of their chemical libraries? What are the advantages and limitations of each?

There are spectroscopic and acoustical methods for determining things like water content in DMSO or precipitation. These techniques are useful for high-throughput process monitoring in real time to understand if certain handling practices are negatively affecting the samples. However, they are of limited value for determining the actual quality of any given sample.

The industry standard analysis is LC-MS with other detectors such as ELSD or CLND. These can answer the real quality and quantity questions, but they are not real-time measurements, and they are expensive (in a relative sense).

Flow nuclear magnetic resonance is sometimes used to evaluate samples, especially for things like stability studies. The time scale is on the order of LC-MS, but the data are of generally low quality because of the efforts necessary to get around the protonated solvents and the lack of sensitivity. Furthermore, data interpretation is time consuming because the automated tools are not very good at interpreting proton spectra (about all that can be acquired in the time frame allowed at the typical concentration range).

Regarding the analysis of HTS libraries, what level of throughput can be reached with the current technology, and what analytical information can be obtained?

Single channel LC-MS systems can analyze between 8,000 and 20,000 samples per month depending on the chromatographic run time. LC-MUX systems typically analyze between 4,000 and 6,000 samples per channel per month depending on whether it is a four- or eight-channel system and the run time.

Methods for determining water content or precipitation are fast, like plate readers, enabling real-time measurements. However, the information is fairly limited.

Ideally, for HTS, complementary to the biological activity of the sample, it would be useful to confirm the compound ID and measure the purity and concentration of each sample screened. Given that HTS throughput ranges from 10,000 to >500,000 samples per day, what advances in technology would be needed to couple analytical analysis to HTS?

This will require a disruptive technology. It is extremely challenging to assay for ID, purity, and concentration on 500,000 unique samples per day. The platform will have to be a massively parallel system that is probably spectroscopically based. In the meantime, the first attempt to design such a system will probably involve some sort of library search approach. A spectroscopic measurement for each sample would be calibrated using LC-MS or similar data. That measurement could then be done on the screening wells and related back to the reference spectrum. In that way, sample changes could be tracked, even if they could not be quantified.

What have been the main challenges in QC of very large sample collections? What are the typical sizes of the libraries being sent to you to QC? Do you employ serial sample analysis or parallel sample analysis to decrease the turnaround time in analyzing samples? What type of systems do you use?

Because we perform both simultaneous qualitative and quantitative analysis there are many more details to control than for just qualitative assays. After 10 years of effort we have developed robust processes for acquiring and interpreting the data. Significant effort is still required to shepherd hundreds of thousands of samples through the analysis process; but it is not a daunting task. We continue to improve the process in the areas of increased throughput and intelligent data interpretation.

The main issues today are business related: a consistent sample stream and price. Customers are reluctant to commit to monthly throughput, but they want fast turnaround time at a low cost. Clearly the most efficient systems (and therefore those operating at the lowest cost) are ones that run at capacity all of the time, with no wasted downtime. However, that means capacity is scheduled out months in advance, and there is not much

wiggle room for accommodating surprises. We offer very compelling rates for companies that lock into a consistent level of work each month. However, most companies choose the flexibility over the lower price.

A second issue is the overall cost of the analysis. The cost per well for our analysis is on the order of that for high-content (imaging) screening. For the money that it will cost for us to analyze 100,000 samples they could go and buy a single quad LC-MS system. Of course there is a lot more to getting the work done than just buying the instrument, but only experienced customers appreciate all of these hidden costs.

No one has yet submitted 250,000 samples on us for analysis at one time. (Although we would like to have that happen.) Most of our customers divide their collection up into smaller screening cassettes (10,000–30,000 samples), and we receive those cassettes one at a time as they are created. We also have customers that send us their actives from screening, so those samples come about every 2 weeks and are sets of 100 to 10,000 samples.

We use Agilent (Santa Clara, CA) LC-MS systems—single channel LC systems. This system is necessary because we operate the four detectors and cannot do that with MUX systems. Additionally, we have found that the redundancy of single channel systems combined with the ability to optimize each system narrows the throughput gap with an MUX system. For our higher-throughput method we run 16,000–20,000 samples per machine per month. For our longer gradient, higher quality method we run 8,000–10,000 samples per month per machine.

On the software side, what are the critical factors to consider in the collection, storage, and analysis of these QC data?

First, the raw data must be reduced to the correct answers. One of the best tools in the industry is the Waters OpenLynx system. However, OpenLynx is not sophisticated or flexible enough for us. We want our results to be correct for every sample. Therefore, we built our own system. Most of the people who analyze hundreds of thousands of samples have built internal systems to further process the OpenLynx RPT files. The fact is that you must have a data reduction strategy that gets the answers that you need from the data. Second, you must capture the human interpretation of the sample. That is where the real value exists, but most commercial browsers don't provide the capability to capture the human interpretation. And third, there must be a way to store the results where they are easily accessible.

Are the software tools and hardware currently available sufficient for these types of analyses?

The two main areas in which we see room for improvement are component analysis—determining how many different chemical species are in a sample and using MS to resolve the components that have not been resolved with chromatography—and interpretation of the components in terms of their structures relative to the target compound. We need to do a better job of determining how the components of a sample are related to each other.

As we described, data handling is a major issue for all of our customers. None of our customers has what we would consider a state-of-the-art data handling system that would allow them to capitalize on the data we are giving them.

We spoke of the limitations in throughput having to do with quality versus speed. First, we try to get the quality of the analysis as high as possible. But we eventually run up against bottlenecks in throughput and performance to produce that quality data, and that is where we have tried to apply our analytical tools. We want to make the interpretation of the data as easy and reliable as possible.

Furthermore, in the full picture of automating and delivering the results, we have to deal with the fact that every customer wants the results delivered in a certain way. We have tried to build the exact solutions they need. Unfortunately, though, we can deliver a lot more useful information than the customer usually can handle from the informatics side. We can provide much richer information than they can access and interpret at this point.

Quality is where we have seen the biggest change in the industry over the last decade. People used to want an approximation of the quality of their libraries. Now they want information on each individual sample.

What is the “diversity” of the libraries being sent to you—primarily Lipinski Rule of 5, or do samples include natural products or other nontraditional drug molecules? What factors complicate the analysis of diverse sample types?

We do not know whether the Lipinski Rules apply because we do not get structures, so we cannot assess that type of information. The diversity of the libraries we see represents the diversity you would find in a typical large pharma collection. They do not contain many natural products. We mostly see heterocycles between 200 and 400 MW; most contain nitrogen, and most are basic. We also analyze fragment libraries, from 90 to 200 MW. Those are a bit more challenging to work with. We do get some peptide libraries—less than 2,000 MW. Solubility issues can complicate analysis. Samples sometimes precipitate in the microtiter plates. That is a function of freeze-thaw cycling, and most customers send us frozen samples.

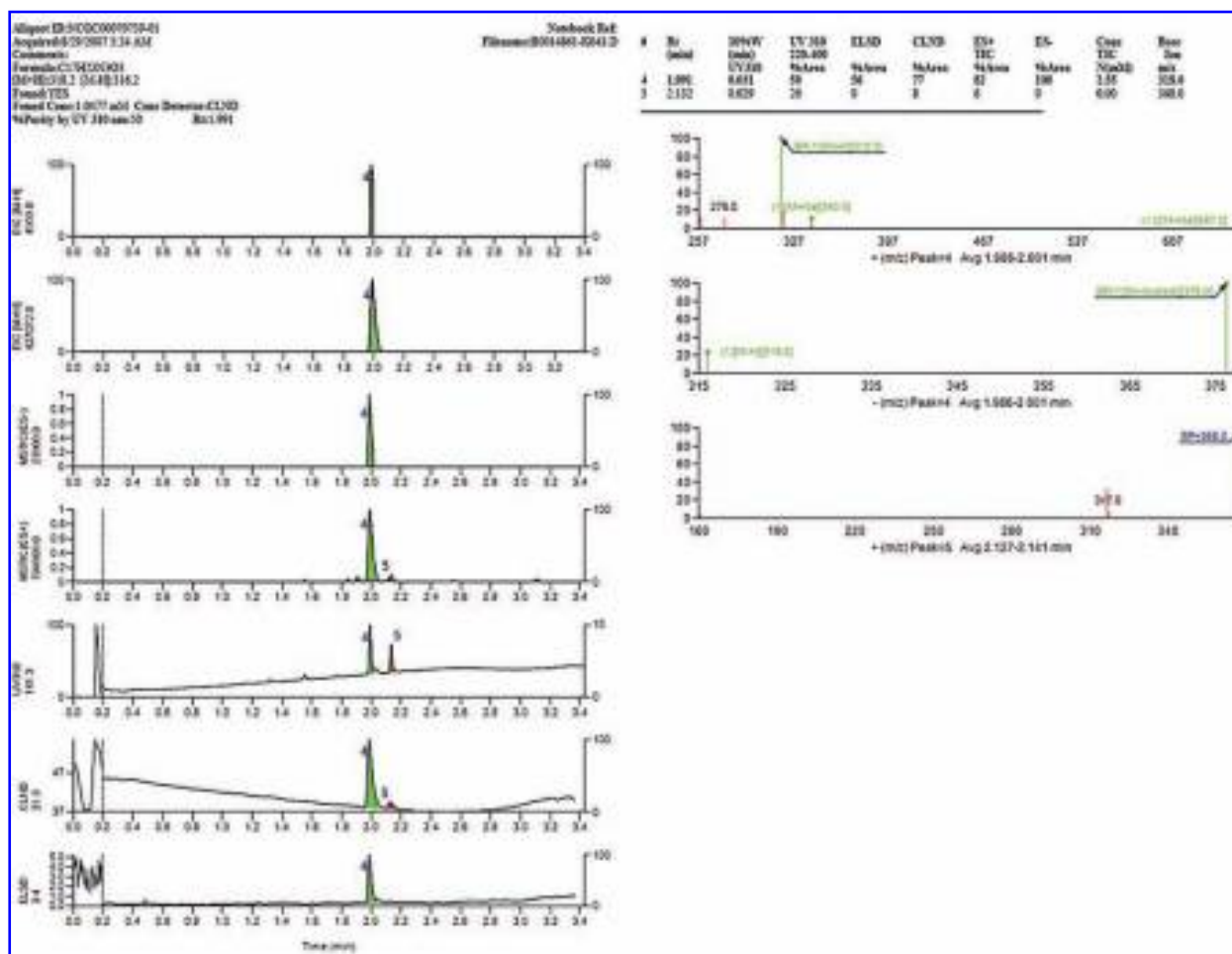


FIG. 1. An example of a typical analytical report OpAns provides to its clients.

What is an example of a typical analytical report you provide to your clients? Do you use methods to estimate sample concentration without relying on a standard curve of the compound? When no impurities are visible on LC, how do you detect sample degradation of 50%? Do you check for sample solubility if you are receiving “solutions” for testing? If so, how?

Yes, there is a typical report (see Fig. 1). We use two quantitative detectors to measure concentration—we use CLND if the compound contains nitrogen, and if it does not, we use ELSD. ELSD is not as accurate though. In a high-throughput environment—tens of thousands run per month—using CLND we typically report concentration accuracy plus or minus 15% without reference standards. In general library terms, ELSD is generally accurate for determining concentrations within a factor of 2.

To assess sample degradation, you can go back to potency, but that assumes the compound has stayed in solution and has not precipitated. Precipitation does not always imply degradation. With these four detectors it

would be rare for a sample not to yield a measurement on at least one, making it unlikely to miss impurities.

The question of solubility goes back to the issue of precipitation. People send us samples in two formats. If the samples are already diluted to the concentration we need for analysis we do not usually check for precipitation. We also receive samples in a 384-well plate format with 1 μ l of 10 mM compound in DMSO in each well. However, we cannot analyze 1 μ l of solution, so we dilute the sample using whatever solvent the customer wants us to use. If we use aqueous buffer rather than DMSO we will certainly look for precipitation, but the assessment is based only on a visual inspection.

Taking a look back, what was the original thinking behind the establishment of OpAns? How have the goals and activities of the company evolved? What is the make-up of your customer base?

We started the company because Eli Lilly closed the facility where we were working. We are pharma refugees, and we are not unique in that experience. We believed

pharma companies would increasingly outsource discovery. That is the same type of work we had done internally, and we felt we could be more focused and provide the services that pharma was looking to outsource in order to reduce their fixed costs.

As it turned out, there is very little of that general outsourcing being done in the United States. Most of the outsourcing of pharma discovery is being shifted to complete service organizations in Asia. We have had success working with individual pharmaceutical companies that want to outsource certain specific operations, such as solubility assessment, physical property measurements, or compound characterization, for example. But that is not

enough on which to build a business. We have found that the main opportunity exists in the area of providing Investigational New Drug-enabling analytical services to small pharmaceutical companies—venture-backed, start-up companies that have a small number of compounds and do not have much infrastructure. They need a whole array of services, and the large contract research organizations (CROs) tend not to be as interested in this type of work. We understand what these companies need to get their compounds to the point of an Investigational New Drug filing and are providing those services.

—Interview by Vicki Glaser