Pharm Res (2012) 29:1745–1754
DOI 10.1007/s11095-012-0730-8

RESEARCH PAPER

Solubility at the Molecular Level: Development of a Critical Aggregation Concentration (CAC) Assay for Estimating Compound Monomer Solubility

Jie Wang • Edmund Matayoshi

Received: 21 October 2011 / Accepted: 2 March 2012 / Published online: 16 March 2012

ABSTRACT

Purpose In drug discovery research the formation of soluble compound aggregates is a major cause of false positives, false negatives, and distorted values in High-Throughput Screening assays that measure either binding or functional activity. These aggregation-based artifacts lead to serious distortions in the SAR which are critical to successful lead optimization. In this work we introduce a new approach by which the "critical aggregation concentration" (CAC) is determined, thereby overcoming limitations inherent to traditional solubility methods and enabling estimation of small molecule *monomer* solubility.

Methods The theoretical and experimental basis of a new confocal Static Light Scattering plate reader assay is presented.

Results Tests conducted with model systems, commercial compounds, and Abbott library compounds show that the CAC assay can measure aqueous monomer solubilities reproducibly and reliably, achieving a sensitivity of ~0.2 μm, which is an improvement of approximately two orders of magnitude over nephelometry.

Conclusions Determination of compound monomer solubilities in a screening format is possible for the first time with the cSLS-CAC methodology. It is currently in routine use in Abbott's drug discovery program, and has enabled identification of many compound induced artifacts in binding or activity assays that are missed by traditional kinetic solubility measurements.

J. Wang • E. Matayoshi (✉)
Department of Structural Biology, Abbott Laboratories
Global Pharmaceutical R & D
100 Abbott Park Rd
Abbott Park, Illinois 60064-6114, USA
e-mail: ed.matayoshi@abbott.com

KEY WORDS aggregation · confocal · drug discovery · static light scattering · solubility

ABBREVIATIONS

API	active pharmaceutical ingredient
CAC	critical aggregation concentration
cSLS	confocal static light scattering
DLS	dynamic light scattering
SAR	structure-activity relationships
SLS	static light scattering

INTRODUCTION

Establishment of high quality structure-activity relationships (SAR) is a critical part of drug discovery research, but it is frequently compromised by distorted affinity or potency values caused by low compound solubility and aggregation (1–3). In the past few years it has become clear that most lipophilic and amphiphilic compounds, when dissolved at sufficiently high concentrations in aqueous solutions, can be present as *soluble aggregates*. The simplest example of this phenomenon is the solubilization of detergents in aqueous solutions as two physical molecular species, monomers and micelles. However, while detergent micelles are aggregates which are monodisperse in size and monomer number, most lipophilic compounds in pharamaceutical libraries will likely form a distribution of soluble aggregates that can range in size from ~1 nm to greater than 1 μm (4–6; *Matayoshi & Wang, unpublished results*). Solutions containing soluble aggregates can appear transparent to the naked eye due to their relatively small particle size and/or number concentration.

False positives or highly overestimated affinities/potencies can result from nonspecific interactions between

compound aggregates and targets ("promiscuous inhibition"), particularly when the aggregate particles formed are much larger than the protein targets themselves (6–9; *Matayoshi & Wang, unpublished results*). Conversely, since aggregation reduces the concentration of monomer available to the target, in the absence of promiscuous binding phenomena false negatives or greatly underestimated affinities/potencies can be generated if the actual monomer concentration is much less than the total compound concentration. The repercussions of aggregation will vary with total compound concentration (over a range which in assays often covers several log units). The nature and magnitude of an aggregation-induced artifact is further complicated by the specifics of the assay, e.g. whether the assay monitors a signal arising from the compound, a competitive probe, an enzymatic product, or the target itself.

Various low and high throughput methods have been developed for compound kinetic or equilibrium solubility screening (10–15, 27–29). Ideally, solubility screening would identify conditions where compound aggregation occurs, thereby defining true monomer solubility and providing guidance for setting appropriate concentration ranges to be tested, or for interpreting assay results at concentrations where aggregation occurs. Unfortunately, previous methods employed for measuring compound solubility do not distinguish soluble monomers from soluble aggregates, hence these measured "total solubilities" can greatly exceed monomer solubilities. Only compound monomer concentrations are directly relevant for evaluating the potencies and affinities used to establish SAR; aggregates at best can be considered as compound reservoirs.

The current kinetic solubility assays can be roughly divided into two categories, homogeneous and heterogeneous. In heterogeneous assays compound aggregates or precipitates are removed by filtration or centrifugation, and the compound concentration remaining in solution is measured by various detection methods such as UV–vis absorbance, mass spectrometry, or chemiluminescent nitrogen detection (CLND) (any of which may optionally be coupled to HPLC) (2,16–18). However, these procedures raise many concerns and ambiguities. Filters have very large pore sizes (practical minimum ~0.2 μm) on a molecular scale and thus fail to remove even large soluble nanoparticle-sized aggregates. In addition, compound binding to filtration membranes is common and can represent a significant fraction of the total compound present, especially at low concentrations. Centrifugation is similarly ineffective: low particle densities and/or low centrifugation speeds result in the removal of only precipitates and extremely large soluble aggregates.

Homogeneous assays such as nephelometry (11,19), flow cytometry (6,20), and dynamic light scattering (DLS) (4,5) avoid the problematic separation process by analyzing compound aggregates and precipitates in free solution directly. These assays are all based on various light scattering techniques and have the advantage of simpler sample processing than in the heterogeneous assays. Commercial nephelometers, however, have a detection limit of only ~20 μM for most compounds (14). The single commercially available flow cytometer designed for solubility analysis (BD Gentest™) reportedly can detect quantitatively only those particles larger than ~100–300 nm (6). Although DLS in principle has the greatest sensitivity amongst this group of technologies for detecting the presence of soluble aggregates at very low concentrations, this capability unfortunately cannot provide reliable particle number information if the particle size distribution is very wide. Low concentrations of large aggregates or rare foreign particles (e.g. 50 nm to ~1 μm in size) can effectively blind the best multi-component DLS analyses to the presence of even an overwhelming mass fraction of monomers (the latter being sub-nanometer in size). DLS is perhaps best employed for its unique ability to characterize particle size and distribution in a homogeneous format when aggregation is known to be present, rather than for screening compound monomer solubilities.

The limitations of traditional assays affect not only the accuracy of solubility measurements, but also run the risk of providing misleading information. The procedures of filtration or centrifugation under different conditions effectively remove different fractions (sample-dependent) of the aggregate particle distribution from the final analysis, potentially leading to widely varying solubility values. In the homogeneous assays a relative insensitivity to monomers and small aggregates can lead to either large overestimation or underestimation of the solubility, depending on how the different fractions (monomers *vs.* aggregates) are effectively weighted in the analysis. Conclusions drawn from observations of the apparent equilibrium time dependence of the compound dissolution process may therefore be highly method dependent, *i.e.* due to inherently differing sensitivities to monomer *vs.* aggregate fractions, and/or which fraction is actually being monitored.

The mechanism of compound aggregation is probably compound dependent. For some amphipathic molecules (such as surfactants), molecules may rapidly equilibrate as stable micelle-like structures under certain conditions, whereas for most compounds in a typical compound library, aggregates may represent the early steps of nucleation and/or precipitation. In drug discovery research compounds are typically pre-dissolved in DMSO, followed by dispersion into the assay buffers. For compounds with low aqueous solubilities, the addition of a concentrated DMSO stock solution to the buffer often generates supersaturated solutions. Compound aggregates form immediately, and can eventually grow into still larger amorphous (or crystalline) particles which sometimes precipitate out of solution when the solution reaches true equilibrium. The precipitation

process can be slow and many compound aggregates may be stable for a very long time. Assays are typically run shortly after addition of compounds, hence do not have time to reach true equilibrium and most likely contain compound aggregates during the assay measurement.

In this paper, we report on the development of a confocal static light scattering (cSLS) based homogeneous assay for compound monomer solubility screening in a plate format. Although utilization of cSLS has been reported for CMC determinations (31) or label-free imaging (30), to the best of our knowledge cSLS has not been previously exploited for systematically screening compound solubility in a high throughput fashion. Our assay is based on measuring the compound "critical aggregation concentration" (CAC), to deduce *monomer* solubility. The CAC represents a threshold above which compound aggregation occurs; hence, it is an experimental estimation of the maximum obtainable concentration of monomers in the given aqueous solution. As discussed above, homogeneous solubility assays based on light scattering in various formats have been used for many years, but the combination of high sensitivity, high throughput, and the capability of discriminating monomers from aggregates is unprecedented. In addition, the reliability of light scattering based assays has not been previously fully investigated. In this paper we present both theoretical and experimental methods for evaluating the reliability of a light scattering-based system. Our results can serve as a theoretical basis for further improvements to light scattering based assays. We demonstrate that our cSLS based method can detect compound CACs down to sub-micromolar concentrations, *i.e.* achieving a sensitivity which is approximately two orders of magnitude lower than a typical nephelometric assay.

MATERIALS AND METHODS

Materials

Flunarizine dihydrochloride**,** miconazole nitrate, tamoxifen, n-Dodecy-β-D-maltopyranoside (DDM) and DMSO (#02245KC) were purchased from Sigma Aldrich and used as received. Standard polystyrene nanoparticles were purchased from Bangs Laboratories Inc. Phosphate buffered saline (PBS) was from Invitrogen (#14190) and PEG-400 from Hampton Research (HR2-603). Water used in all assays was Milli-Q (Millipore) purified.

Sample Preparation

Compound stock solutions were prepared in DMSO at a concentration of 10 mM. The solubilities of Abbott research compounds were tested in 40 mM sodium phosphate buffer (pH 7.4). A series of 2X dilutions was typically made for each compound (e.g. from 200 μM to 0.1 μM) across rows in 96-well plates (Corning UV plate, #3679), providing 12 data points per compound for the CAC analysis. For very soluble compounds additional concentrations were also prepared (e.g. 2X dilutions from 500 μM to 15.6 μM). It should be noted that virtually any buffer can be used for the CAC assay; buffer "cleanliness" (free of particulates and micro air bubbles) is essential for providing a low and stable background scattering intensity.

Light Scattering Measurements

Confocal static light scattering measurements were carried out with a commercial light scattering plate reader (DynaPro™, Wyatt Technology Corporation) equipped with an 830 nm diode laser. Although this plate reader was designed primarily for dynamic light scattering (DLS) measurements, it can also simultaneously record SLS intensities. The near-infrared scattering wavelength is optimal for the present application because very few pharmaceutical library compounds have absorption in this wavelength range, virtually eliminating interference by colored or fluorescent compounds, and common assay solution excipients. The dramatic reduction of background signal (primarily due to reflections from well surfaces and the sample meniscus) conferred by the confocal optics, which we exploit to obtain unprecedented sensitivity for the present SLS-based assay, is similar to that documented theoretically and experimentally for imaging by scanning confocal fluorescence microscopy, and for single molecule fluorescence correlation spectroscopy applications (34). In the DynaPro scattered light is detected at an angle of 157°, and the physical confocal volume is ≈ 67 pL (M. Larkin, Wyatt Technology Corporation, personal communication). This instrument can automatically attenuate both laser power as well as the scattered light intensity during the measurement, and therefore is well-suited for testing solutions where scattered intensities can range over several orders of magnitude. As will become clear from the discussion herein, extreme linearity of the measured SLS intensities across this very wide amplitude range (as would be required if molecular weight determinations were the purpose of the assay) is not necessary for a CAC determination. All CAC measurements in this paper were performed at 25°C (temperature is controlled with an absolute accuracy of 1°C within the DynaPro instrument). A 96-well plate (containing 8 compounds if dilutions are made across rows as described above) can be read typically in 30–60 min.

Principle of Critical Aggregation Concentration (CAC) Measurement

In the CAC assay presented here, SLS intensities are measured at a series of different compound concentrations in

aqueous solution. For Rayleigh scattering, where particle size is much smaller than the laser wavelength, the scattered intensity increases linearly with increasing particle number, but has a 6th power dependence on particle size for a given particle number concentration. This steep increase in intensity with increasing particle size underlies the extreme sensitivity of SLS for detecting the presence of very low concentrations of soluble aggregates in a solution of compound monomers (the latter being sub-nanometer in size).

We define the CAC as the concentration of compound at which molecular aggregates first begin forming; this transition point is readily observed as an abrupt (nonlinear) jump in SLS intensity with increasing compound concentration. Our observations indicate that compound aggregation in general is a heterogeneous process in which a distribution of particle sizes is generated. For this reason we do not analyze the CAC data in terms of a two state model commonly assumed for the special case of micelle formation by amphiphilic molecules like detergents. For purposes of high throughput screening, we consider it sufficient to estimate CAC by the highest concentration point exhibiting SLS intensity indistinguishable from the solvent background (taking into consideration S:N and S:B specific for the given sample by inspecting fluctuations between adjacent data points). Since the reproducibility of replicate readings or replicate samples in the assay is very high, the accuracy of the CAC value determined in the high throughput screening mode as described is essentially defined by the dilution factor between samples (2-fold for the data herein). If using the assay at a lower throughput mode of operation for analytical applications, the accuracy can always be improved by reducing the dilution factor between concentration points. Figure 1 shows a typical set of experimental data collected from an Abbott research compound. The transition point (CAC) is treated as an estimate of the monomer solubility of the compound (for further discussion regarding this interpretation, see below). Thus in general, the CAC value will be less than or equal to the "solubility" obtained by a traditional heterogeneous or homogeneous format assay.

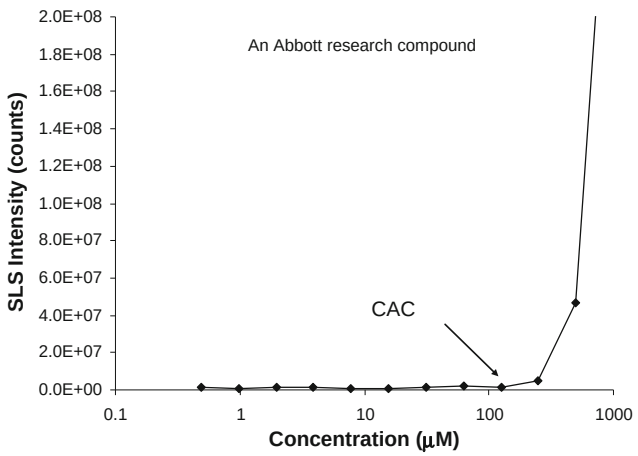

Fig. 1 Typical SLS experimental data collected from an Abbott research compound. See discussion in text.

We have adopted the term "critical aggregation concentration" in order to emphasize the distinction in meaning from the conventionally determined "solubility". As discussed above, conventional solubility methodology yields a result that represents the sum of the soluble monomer and soluble aggregate fractions. The CAC terminology has been used previously in the physical chemistry literature, and more recently for drug aggregation (6). However, in the latter paper, determination of the CAC was limited by the detection sensitivity of the flow cytometer, which as discussed above, is accurate only for extremely large (greater than ~100–300 nm) aggregates. This potentially leads to substantial overestimation of the monomer solubility. SLS using a conventional optical format on single samples has been applied to characterize the aggregation of several amphiphilic drugs (21,22). In the following section, we discuss in detail the sensitivity and reliability issues for the *confocal* SLS based approach.

Method for Analyzing the Reliability of the Light Scattering System

In order to validate this assay, we need to analyze the reliability of the light scattering system. We start with the Zimm treatment for Rayleigh light scattering. For a solution with a solute (weight) concentration C, we have the relations (23,32,33):

$$R_\theta = KM_wCP(\theta)[1 - 2A_2P(\theta)MC] \quad (1)$$

with $K = 4\pi^2 n_0^2 (dn/dC)^2/(N_A\lambda_0^4)$ and

$$R_\theta = i_\theta f/I_0 \quad (2)$$

where R_θ is the excess Rayleigh ratio, λ is the wavelength and I_0 the intensity of the incident light, and f is the absolute calibration value for the geometry as well as other parameters of the system (33). θ is the scattering angle, n is the refractive index of the particle and n_0 is the refractive index of the medium. i_θ (= i(solution) - i(solvent)) is the excess scattering intensity. $P(\theta)$ is the normalized the intensity distribution function as described in reference (33). A_2 is the second virial coefficient and M_w is the molar mass. For a confocal light scattering system, the total scattered intensity should be integrated over the angular and positional distributions of incident and scattered light (30). For our system the numerical apertures of both incident and scattered lights are very small, and Eq. (1) can be used as an approximation. This equation can be simplified if we treat $P(\theta)$ as a constant for all particles of a fixed system and the solute

concentration is relatively low. If we further assume a constant value of dn/dC for all organic molecules, under the Rayleigh scattering approximation we have a simplified expression for our system:

$$i_{system} = A_{system} C M_w \tag{3}$$

Here, A_{system} reflects the sensitivity of a light scattering system and can be treated as a constant for a fixed system. Under this condition, scattered intensity is proportional to the molar mass M_w for a given concentration C. We need to point out here that for purposes of evaluating a high throughput screening system (as intended in this paper), a certain degree of error is tolerable and our approximate treatment is sufficient. For highly quantitative measurements, a more accurate model should be used. In order to distinguish signal from the solvent background there is a minimum scattering intensity i_{min} needed. This minimum light scattering intensity corresponds to a minimum concentration $C_{\min\_scattering}$ (detection limit) as shown in Eq. (4):

$$C_{\min\_scattering} = \frac{i_{\min}}{A_{system} M_w} \tag{4}$$

Assuming the same density μ for all compound aggregates, the mass of a particle can be replaced by the particle diameter d (under spherical approximation):

$$C_{\min\_scattering} = \frac{6 i_{\min}}{A_{system} \pi \mu d^3} \tag{5}$$

Equation (5) can be further simplified to

$$C_{\min\_scattering} = \frac{A_0}{d^3} \tag{6}$$

Here, A_0 can be treated to good approximation as a constant, which includes all the constant parameters in Eq. (5). Equation (6) provides a convenient way to evaluate the detection limit of a light scattering system. If we know the $C_{\min\_scattering}$ for a well characterized particle (such as a protein), the detection limits for different size of particles can be deduced.

Besides the sensitivity of the cSLS system, the effective detection volume will also affect the reliability of this assay. Confocal light scattering systems have limited detection volumes, and if the particle number concentration is too low, the light scattering intensity measured from a very small volume at any instant might not be representative of the average intensity of the solution. This limitation could dramatically affect the reliability of a light scattering system. For a concentration C, the particle concentration N_p can be calculated by

$$N_p = \frac{C N_A}{M_w} = \frac{6 C N_A}{\pi \mu d^3} \tag{7}$$

where N_A is the Avogadro constant. In order to obtain a reliable measurement the particle number should be high enough such that on average at least one particle is present in the effective detection volume V_d. The minimum concentration $C_{\min\_number}$ can be estimated by

$$C_{\min\_number} = \frac{\pi u d^3}{6 V_d N_A} = B_0 d^3 \tag{8}$$

Here B_0 can be treated as a constant for a fixed system since V_d changes slowly in comparison with d^3. It should be noted that due to particle diffusion, the effective detection volume differs from the physical confocal volume. The effective detection volume depends on both particle size, solution viscosity, and acquisition time. The effective detection volume increases with decreasing particle size because smaller particles undergo faster Brownian diffusion, *i.e.* enter and exit the confocal observation volume more frequently. Likewise, a longer acquisition time corresponds to a larger effective detection volume because the probability of any particle diffusing into the confocal observation region is increased.

According to Eq. (6), the detection limit $C_{\min\_scattering}$ decreases dramatically with increasing particle size, while according to Eq. (8), the detection limit $C_{\min\_number}$ increases dramatically with increasing particle size. Therefore, there is a detectability region for a given confocal SLS system determined by both sensitivity and effective detection volume.

RESULTS AND DISCUSSION

Detection Limit Test Using Model Systems

We tested the detection limits of this assay by using several model systems. Aqueous solutions of PEG 400 ($M_w \sim 400$ D) were used to simulate the situation where small compounds are monomeric and do not form aggregates; bovine serum albumin (BSA, $M_w \sim 66$ kD, $d \sim 7$ nm) was used to simulate the situation for a small compound aggregate (such as a micelle) in solution; and 23 nm, 53 nm, 130 nm, 290 nm, 490 nm and 2 μm standard polystyrene particles were used to simulate situations where compounds form large aggregates.

SLS intensities collected from PEG 400 water solutions at different concentrations are shown in Fig. 2. The SLS intensity begins to increase above the solvent background scattering when the concentration is greater than ~10 mg/ml. DLS measurements confirmed that the latter solutions are dominated by PEG monomers (hydrodynamic $d \sim 1$ nm). From this test we conclude that in our system if small compounds do not form aggregates, the SLS intensity due to monomers will be undetectable above solvent scattering when compound concentrations are much less than ~10 mg/ml.

SLS intensities measured from BSA solutions at different concentrations are shown in Fig. 3. The detection limit for BSA is about 60 μg/ml under our experimental condition.

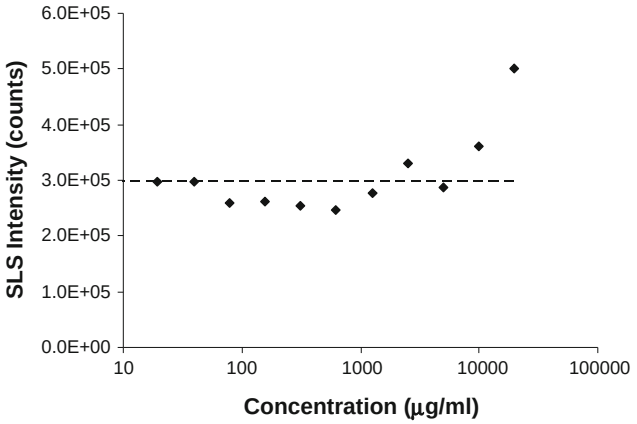

Fig. 2 SLS intensities (*solid dots*) collected from solutions of PEG 400 in water at different concentrations. Dashed line: water background.

This concentration is somewhat larger than the detection limit specified for DLS by the manufacturer of this instrument. This can probably be attributed to an effectively greater tolerance for noise by the DLS analysis.

We also tested the detection limits of standard polystyrene particles suspended in water. The experimental data is shown in Fig. 4. Table I summarizes the results from all of the tests for SLS detection limits in our system. The detection limits for particles of different size are also plotted in Fig. 5 as solid symbols. In Fig. 5, when particles are small (such as 1, 7, 23, 53, and 130 nm particles) the measured detection limits could be fit well by Eq. 6 (solid line). For larger particles (such as 290, 490, and 2000 nm particles), the measured detection limits are far away from the curve fit by Eq. 6, but can be fit well by Eq. 8 (dashed line). The crossing point of the two curves is at ~0.004 μg/ml, ~176 nm. This is an important result for this SLS system: if particles are smaller than 176 nm, the detection limit of this system is limited by the sensitivity of the system, while if particles are larger than 176 nm, the detection limit is limited by the effective detection volume. The effective

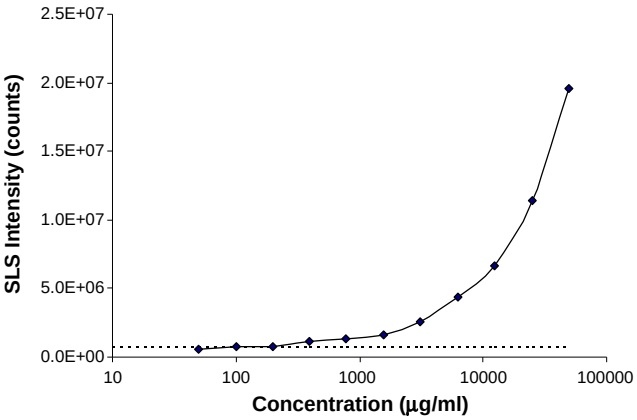

Fig. 3 SLS intensities of BSA (in PBS, pH 7.4) collected at different concentrations. Dashed line: buffer background.

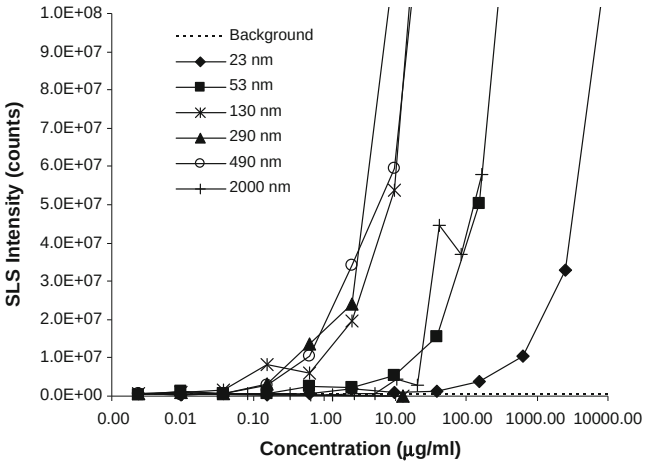

Fig. 4 SLS intensities collected from standard polystyrene nanoparticles suspended in water at different concentrations. Dashed line: water background.

detection volume of this system deduced from the fitting curve is ~1 nL, which is ~15 times larger than the physical confocal volume (67 pL). The intersection of the two curves also represents the absolute detection limit of the system. If the compound concentration is lower than this value, the likelihood of detecting aggregates from this solution would be extremely low. This point also shows that the assay has maximal sensitivity (with respect to concentration) when compounds form ~176 nm particles. For a compound of M_w ~300 D, this corresponds to an absolute detection limit of ~13 nM if the compound forms only 176 nm sized aggregates.

Test of Some Commercial Compounds

We further validated this system by testing some well characterized compounds. SLS intensities collected from n-Dodecy-β-D-maltopyranoside (DDM) water solutions at different concentrations are shown in Fig. 6. The "critical micelle concentration" (CMC) for amphiphilic molecules such as detergents might be described as a specialized case of the CAC, because the aggregate species formed are monodisperse and small (less than ~8 nm in size), in contrast to the broad, heterogeneous soluble particles (as large as microns in size) generated by aggregation of most nonpolar organic compounds. DDM has reported water solubility of ~0.1 M and a CMC ~ 0.17 mM (24,25). There is an SLS intensity transition point at ~50 μg/ml (0.1 mM) in Fig. 4, which is close to the reported CMC. We also observed by DLS that DDM forms d ~ 7 nm micelles when the solution concentration was above 50 μg/ml.

SLS intensities collected from Tamoxifen, Miconazole nitrate and Flunarizine dihydrochloride solutions at different concentrations are shown in Fig. 7. We used the same experimental conditions described in reference (14).

Table I Summary of Tested Detection Limits for Different Sizes of Particles

Particle size d (nm)	1[a]	7	23	53	130	290	490	2000
Detection limit (μg/ml)	10000	60	0.6	0.15	0.005	0.01	0.07	5

[a] The size of PEG400 monomer was estimated by using spherical approximation

The CAC of tamoxifen (~15 μM) measured by our assay is close to the reported kinetic solubility. However, for Miconazole nitrate and Flunarizine dihydrochloride, the CAC values (~5 μM and ~3 μM, respectively) are much smaller than the kinetic solubility values measured by nephelometry and very close to the thermodynamic solubility values measured by filtration/HPLC, but higher than the reported kinetic solubility measured by filtration/HPLC (14). We believe that the reported low kinetic solubility values obtained by filtration/HPLC could be affected by filter binding or experiment errors (e.g. the kinetic solubilities of these two compounds measured in reference (14) should be higher than their thermodynamic solubilities).

The aggregation above the CAC is observed to be clearly reversible, since samples are prepared by serial two-fold dilutions from the highest concentration tested.

Test of Selected Abbott Research Compounds

Twenty-nine compounds were selected from the Abbott compound library, with predicted solubilities ranging from very soluble to very insoluble. The kinetic solubilities of these 29 compounds in 20 mM sodium phosphate buffer, pH 7.0, were measured by NMR, ultrafiltration/UV–vis, centrifugation/CLND, and nephelometry. A comparison of these results with CACs from the new SLS-based assay is summarized in Table II.

For most compounds, the CACs agreed well with the kinetic solubilities measured by the traditional methods, although there were also some exceptions. Nephelometry and NMR lacked sensitivity when compound concentrations were below 32 μM and 20 μM, respectively, while the CAC assay in its current configuration can measure solubilities down to ~0.2 μM concentrations. Although nephelometry is also based on light scattering, we attribute the ca. two orders of magnitude better sensitivity achieved by our SLS-based method to the confocal optics, which greatly reduces stray scattering and enables an enormous improvement in the signal to background ratio.

The kinetic solubility values of compounds 12 and 13 measured by the traditional centrifugation/CLND assay were quite large compared with the CACs. DLS analysis of these two compound solutions indicated that these two compounds form soluble aggregates in buffer with d ~ 200 nm and 300 nm, respectively (data not shown). The discrepancy with the CAC value is likely due to the failure of low speed centrifugation to efficiently remove the low-density nanoparticles, thus yielding large apparent solubilities. In principle ultrafiltration with low MW cutoff membranes (e.g. 3 kD) should be able to remove even small (d < 10 nm) compound aggregates, but the operational cost is much higher than the SLS based assay, and unpredictable non-specific binding of compounds to filters remains a major concern (e.g. compound 9). Ultrafiltration is only a separation method, and the detection procedure employed to complete the assay could introduce yet additional uncertainties.

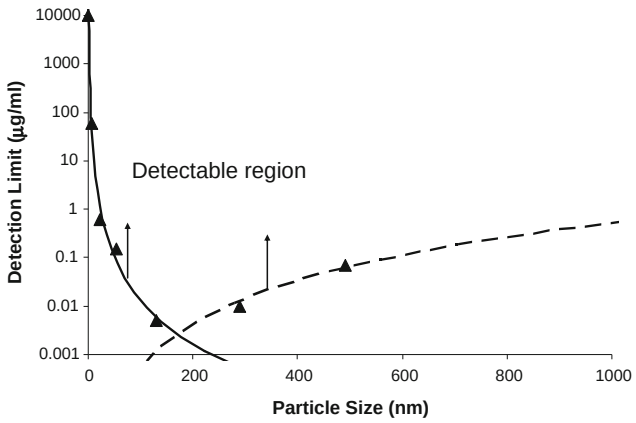

Fig. 5 Detection region for a typical experimental condition (acquisition time ~5 s). Solid line: fitted with Eq. 6; dashed line: fitted with Eq. 8; solid dots: experiment data.

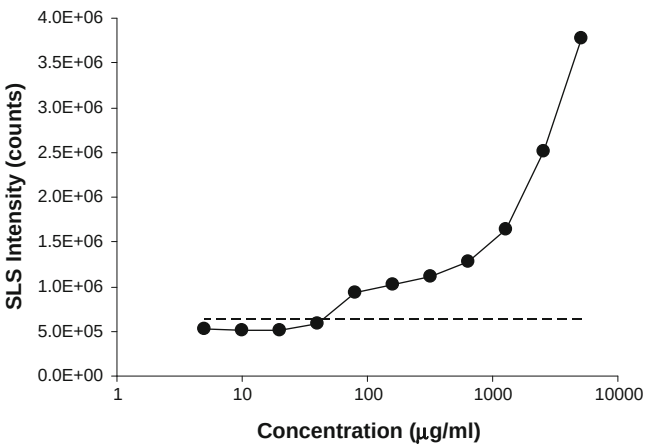

Fig. 6 SLS intensities (*solid dots*) collected from solutions of DDM in water at different concentrations. Dashed line: water background.

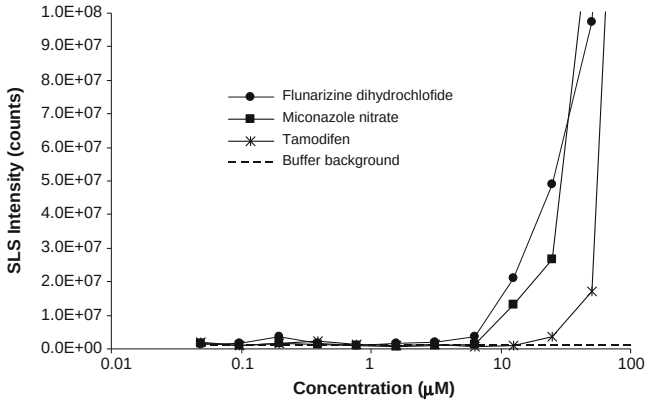

Fig. 7 SLS intensities of three commercial compounds collected at different concentrations. Dashed line: water background.

DISCUSSION

The CAC assay is suitable for compounds with relatively high purity and known concentrations. Fortunately, compound libraries in pharmaceutical companies generally meet these requirements. However, our assay can also tolerate a certain degree of distortion introduced by aggregates caused by chemical impurities. As mentioned above, the SLS measurement simultaneously performs a DLS analysis, hence particle size is also recorded. The total mass represented by these aggregates can be calculated, and the latter can be checked for consistency with the detection region (Fig. 5) of the cSLS-CAC method.

The microscopic (pL) detection volume created by the confocal optics is key to attaining outstanding signal:

Table II Solubility and CAC Tests of 29 Selected Abbott Research Compounds

Compound number	NMR Solubility (uM)	Nephelometry (uM)	Centrifugation plus CLND (uM)	Ultra-filtration plus UV–vis (μM)	CAC by Static Light Scattering (μM)
1	≥ 320	>320	163.2	259	>320
2	≥ 320	>320	255	259	>320
3	≥ 320	>320	191.2	>200	>320
4	≥ 320	>320	199.9	267	240
5	≥ 320	>320	192.6	110	125
6	≥ 320	32< × <320	120.1	150	150
7	≥ 320	32< × <320	85.9	136	10
8	40	32< × <320	41.3	92	125
9	80	>32	29.7	<5	50
10	160	32< × <320	164.5	180	>320
11	≥ 320	>32	32.9	35	2
12	<20	32< × <320	135.8	<5	<0.4
13	<20	32< × <320	126.8	<5	<1
14	<20	<32	<3	<5	2
15	<20	32< × <320	65.6	<5	1.6
16	<20	<32	13.5	<5	<0.1
17	<20	<32	<1.5	<5	6
18	<20	<32	7.1	17	<0.6
19	<20	<32	20.3	<5	12.5
20	<20	<32	10	<5	2
21	<20	<32	<1.5	<5	2
22	<20	<32	3.3	<5	<0.8
23	<20	<32	16.7	<5	<0.2
24	<20	<32	<6.0	<5	16
25	<20	<32	22.2	<5	<0.2
26	<20	>320	67.8	<5	6
27	<20	<32	<1.2	<5	<0.4
28	<20	<32	10.2	<5	5
29	<20	<32	3.2	<5	3

1. For solubilities listed as ">" the indicated value, the latter represents the maximum concentration analyzed

2. When solubilities are listed as "<" the indicated value, reliable measurements were not obtainable at lower concentrations due either to assay sensitivity, or signal:noise specific to the given sample

background performance for SLS analyses at very low compound concentrations by dramatically reducing the background interference (reflection and scattering) from the well walls and the liquid/air interface. However, it potentially creates the problem of not detecting macroscopic compound precipitate settled at the bottom of the well; the latter might represent a significant fraction of the total compound mass in the sample. In such a case the confocal observation region located in the solution above the precipitate could yield artifactually low scattering, leading to overestimation of the CAC. We have found we can overcome this problem by imaging the entire 96-well plate using a home-built setup (*Wang & Matayoshi, manuscript in preparation*); the bottom of each well is thereby imaged, and precipitated particles as small as a few microns in size are easily detected. The additional imaging step takes less than a minute, and greatly improves the reliability of the CAC assay.

The reproducibility of the cSLS assay itself, from a high throughput screening perspective, is very high: for a typical compound, the data variation between replicate readings when operating in high throughput mode yields a CAC accurate to within a factor of ~2 fold (*i.e.* under the described 2X dilution protocol) over concentrations ranging from sub-μM to near mM. That is, reproducibility of the assessed CAC for a given preparation is well within the concentration difference between points when dilution factors of 2X or greater are chosen. The high reproducibility of the assay in fact enables us to routinely study the kinetics of compound solubility. However, we should emphasize that the *reliability* of the absolute CAC measured is far more dependent on the purity and accuracy of the concentrations of the compound stock solutions. Deviations from the assumed nominal 10 mM stock concentrations due to factors such as poor solubility in neat DMSO, inaccurate weighing of micro quantitities, or low chemical purity, will clearly lead to distorted CACs since absolute concentrations of the compound of interest are not directly determined. Furthermore, the presence of contaminants which differ in solubility and/or contribute significant scattering could generate an inaccurate result.

In this paper theoretical and experimental evaluation of the sensitivity of the cSLS system as a function of particle size was treated at length, and due to existing S:B limitations, direct discrimination of monomers from the smallest aggregates (e.g. dimers) is clearly not yet possible for compounds with CACs in the low μM to sub-μM range. With further innovative instrument modifications, however, S:B levels can certainly be greatly improved towards reducing this technical limitation. Our experience has been that compound aggregation always escalates rapidly over a small concentration range (*i.e.* steep CAC transition), hence the CAC assay at present minimally provides an estimate approaching true monomer solubility. The primary goal of the high throughput mode of the assay is to identify and interpret potentially distorted binding or activity assay results caused by compound aggregation; this problem represents one of the major challenges today in drug discovery research. Beyond the 29 compounds in Table II more than 2000 Abbott research compounds have been evaluated, and based on this experience, we believe the cSLS-CAC assay represents a significant advance over traditional kinetic solubility measurements.

CONCLUSIONS

Many failures in drug discovery and development are caused by an inadequate understanding of the physical properties of the solubilized API in aqueous solution (26). In this paper we have shown how small molecule solubility can be analyzed in a more meaningful way. We have demonstrated that a cSLS assay can be used to accurately measure CACs for large numbers of compounds in a plate reader format. This method achieves sensitivities at submicromolar compound concentrations, enabling for the first time a reliable estimate of compound *monomer* solubility. This represents a significant improvement upon the reliability and sensitivity of current HT homogeneous format assays such as nephelometry, flow cytometry, and DLS. It is simpler and faster than current separation based heterogeneous solubility assays, which furthermore do not discriminate soluble monomers from soluble aggregates. The throughput of the cSLS-CAC assay at the moment is moderate, but if desired could be improved by automation of the plate preparation steps. The information provided by CAC analyses dramatically reduces susceptibility to artifacts caused by compound aggregation in various binding and functional assays, thus greatly impacting SAR quality in drug discovery research.

ACKNOWLEDGMENTS & DISCLOSURES

We thank numerous Abbott GPRD Discovery colleagues for help in measuring compound solubilities by other methods, and Drs. Jonathan Greer and Vincent Stoll for support and discussions.

REFERENCES

1. Di L, Kerns EH. Biological assay challenges from compound solubility: strategies for bioassay optimization. Drug Discovery Today. 2006;11:446–51.
2. Alsenz J, Kansy M. High throughput solubility measurement in drug discovery and development. Adv Drug Del Rev. 2007;59:546–67.

3. Feng BY, Simeonov A, Jadhav A, Babaoglu K, Inglese J, Shoichet BK, Austin CP. A high-throughput screen for aggregation-based inhibition in a large compound library. J Med Chem. 2007;50:2385–90.
4. Seidler J, McGovern SL, Doman TN, Shoichet BK. Identification and prediction of promiscuous aggregating inhibitors among known drugs. J Med Chem. 2003;46:4477–86.
5. Frenkel YV, Clark AD, Das K, Wang YH, Lewi PJ, Janssen PAJ, Arnold E. Concentration and pH dependent aggregation of hydrophobic drug molecules and relevance to oral bioavailability. J Med Chem. 2005;48:1974–83.
6. Coan KED, Shoichet BK. Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. J Am Chem Soc. 2008;130:9606–12.
7. McGovern SL, Caselli E, Grigorieff N, Schoichet BK. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. J Med Chem. 2002;45:1712–22.
8. Rishton GM. Nonleadlikeness and leadlikeness in biochemical screening. Drug Discovery Today. 2003;8:86–96.
9. Goode DR, Totten RK, Heeres JT, Hergenrother PJ. Identification of promiscuous small molecule activators in high-throughput enzyme activation screens. J Med Chem. 2008;51:2346–9.
10. Pan L, Ho Q, Tsutsui K, Takahashi L. Comparison of chromatographic and spectroscopic methods used to rank compounds for aqueous solubility. J Pharm Sci. 2001;90:521–9.
11. Dehring KA, Workman HL, Miller KD, Mandagere A, Poole SK. Automated robotic liquid handling/laser-based nephelometry system for high throughput measurement of kinetic aqueous solubility. J Pharm Biomed Anal. 2004;36:447–56.
12. Tan H, Semin D, Wacker M, Cheetham J. An automated screening assay for determination of aqueous equilibrium solubility enabling SPR study during drug lead optimization. J Lab Autom. 2005;10:364–73.
13. Sugano K, Kato T, Suzuki K, Keiko K, Sujaku T, Mano T. High throughput solubility measurement with automated polarized light microscopy analysis. J Pharm Sci. 2006;95:2115–22.
14. Hoelke B, Gieringer S, Arit M, Saal C. Comparison of nephelometric, UV-spectroscopic, and HPLC methods for high-throughput determination of aqueous drug solubility in microtiter plates. Anal Chem. 2009;81:3165–72.
15. Heikkila T, Karjalainen M, Ojala K, Partola K, Lammert F, Augustijns P, Urtti A, Yliperttula M, Peltonen L, Hirvonen J. Equilibrium drug solubility measurements in 96-well plates reveal similar drug solubilities in phosphate buffer pH 6.8 and human intestinal fluid. Int J Pharm. 2011;405:132–6.
16. Kerns EH, Di L, Carter GT. In vitro solubility assays in drug discovery. Curr Drug Metab. 2008;9:879–85.
17. Yamashita T, Dohta Y, Nakamura T, Fukami T. High-speed solubility screening assay using ultra-performance liquid chromatography/mass spectrometry in drug discovery. J Chromatogr A. 2008;1182:72–6.
18. Bhattachar SN, Wesley JA, Seadeek C. Evaluation of the chemiluminescent nitrogen detector for solubility determinations to support drug discovery. J Pharm Biomed Anal. 2006;41:152–7.
19. Bevan CD, Lloyd RS. A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. Anal Chem. 2000;72:1781–7.
20. Goodwin JJ. Rationale and benefit of using high throughput solubility screens in drug discovery. Drug Discovery Today. 2006;3:67–71.
21. Attwood D, Fletcher P. Total intensity and quasielastic light scattering studies on the association of amphiphilic drugs in aqueous electrolyte solutions. J Coll Int Sci. 1987;115:104–9.
22. Taboada P, Attwood D, Ruso JM, Garcia M, Mosquera V. Static and dynamic light scattering study on the association of some antidepressants in aqueous electrolyte solution. Phys Chem Chem Phys. 2000;2:5175–9.
23. Pecora R. Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy. New York: Plenum Press; 1985.
24. VanAken T, Foxall-VanAken S, Castleman S, Ferguson-Miller S. Alkyl glycoside detergents: synthesis and applications to the study of membrane proteins. Methods Enzymol. 1986;125:27–35.
25. Affymetrix, Inc. Anatrace 2011–2012 product catalog. Available at: http://media.affymetrix.com/support/technical/anatrace/docs/2011-12_Anatrace_products_catalog.pdf
26. Gardner CR, Walsh CT, Almarsson O. Drugs as materials: valuing physical form in drug discovery. Nat Rev Drug Discov. 2004;3:926–34.
27. Brouwers J, Brewster ME, Augustijns P. Supersaturating drug delivery system: The answer to solubility-limited oral bioavailability? J Pharm Sci. 2009;98(8):2549–71.
28. Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Aqueous solubility of crystalline and amorphous drugs:Challenges in measurement. Pharm Dev Tech. 2011;16(3):187–200.
29. Baird JA, Taylor LS, Evaluation of amorphous solid dispersion properties using thermal analysis techniques. Adv Drug Deli Rev. 2011.
30. Fang H, Qiu L, Vitkin E, Zaman MM, Andersson C, Salahuddin S, Kimerer LM, Cipolloni PB, Modell MD, Turner BS, Keates SE, Bigio I, Itzkan I, Freedman SD, Bansil R, Hanlon EB, Perelman LT. Confocal light absorption and scattering spectroscopic microscopy. Appl Opt. 2007;46(10):1760–9.
31. Gao Y, Li LB, Zhai G. Preparation and characterization of Pluronic/TPGS mixed micelles for solubilization of camptothecin. Colloids Surf B. 2008;64:194–9.
32. Zimm BH. The scattering of light and the radial distribution function of high polymer solutions. J Chem Phys. 1948;16(12):1093–9.
33. Wyatt P. Light scattering and the absolute characterization of macromolecules. Anal Chimi Acta. 1993;272(193):1–40.
34. Sandison DR, Webb WW. Background rejection and signal-to-noise optimization in confocal and alternative fluorescence microscopes. Appl Opt. 1994;33(4):603–15.