doi.org/10.26434/chemrxiv.13198688.v1

Automated Solubility Screening Platform Using Computer Vision

Parisa Shiri, Veronica Lai, Tara Zepel, Daniel Griffin, Jonathan Reifman, Sean Clark, Shad Grunert, Lars Yunker, Sebastian Steiner, Henry Situ, Fan Yang, Paloma Prieto, Jason Hein

Submitted date: 06/11/2020 • Posted date: 09/11/2020

Citation information: Shiri, Parisa; Lai, Veronica; Zepel, Tara; Griffin, Daniel; Reifman, Jonathan; Clark, Sean; et al. (2020): Automated Solubility Screening Platform Using Computer Vision. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.13198688.v1

Solubility screening is an essential, routine process that is often labour intensive. Robotic platforms have been developed to automate some aspects of the manual labour involved. However, many of the existing systems rely on traditional analytic techniques such as High Performance Liquid Chromatography or HPLC, which require pre-calibration for each compound and can be prohibitively expensive. In addition, automation is not typically end-to-end, requiring user intervention to move vials, establish analytical methods for each compound and interpret the raw data. We developed a closed-loop, flexible robotics system with integrated solid and liquid dosing capabilities that relies on computer vision and iterative feedback to successfully measure caffeine solubility in multiple solvents. After initial researcher input (<2 min), the system ran autonomously, screening five different solvent systems (20-80 min each). The resulting data matched values obtained using traditional manual techniques.

Automated solubility screening platform using computer vision

P. Shiri[1], V. Lai[1], T. Zepel[1], D. Griffin[2], J. Reifman[2], S. Clark, S. Grunert[1], L. P. E. Yunker[1], S. Steiner[1], H. Situ[1], F. Yang[1], P. L. Prieto[1], J. E. Hein[1,*]

[1]Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada
[2]Amgen Inc. Cambridge, MA, 02141 USA
[*]Corresponding author. Email: jhein@chem.ubc.ca

Abstract

Solubility screening is an essential, routine process that is often labour intensive. Robotic platforms have been developed to automate some aspects of the manual labour involved. However, many of the existing systems rely on traditional analytic techniques such as High Performance Liquid Chromatography or HPLC, which require pre-calibration for each compound and can be prohibitively expensive. In addition, automation is not typically end-to-end, requiring user intervention to move vials, establish analytical methods for each compound and interpret the raw data. We developed a closed-loop, flexible robotics system with integrated solid and liquid dosing capabilities that relies on computer vision and iterative feedback to successfully measure caffeine solubility in multiple solvents. After initial researcher input (<2 min), the system ran autonomously, screening five different solvent systems (20-80 min each). The resulting data matched values obtained using traditional manual techniques.

Introduction

Solubility is defined as the amount of a compound that will dissolve in a certain amount of solvent at a given temperature, when the system is in equilibrium. The ability to reliably screen solubility is fundamental to many areas of chemical research across industries, from pharmaceutical applications[1,2], where solubility is used for process development and solvent selection, to material science, where it limits formulation and processing parameters[3–5]. Common methods, which can be broadly categorized as the addition of 'excess solid' or 'excess solvent' typically require researchers to perform incremental, manual tasks such as solid dosing that must be adjusted and repeated for each solute-solvent system[6].

In the case of excess solid, solute is added until an excess is observed. The solution is allowed to equilibrate for a long incubation time and is then sampled and analyzed for composition using a variety of analytic techniques (e.g. gravimetric methods, titration, UV-spectrometry, or most commonly High Performance Liquid Chromatography or HPLC). The excess solid method with filtration, also called 'shake-plate' or 'slurry', is known to provide the highest quality data. However, it additionally requires filtration or at-temperature centrifugation prior to analysis[7], which can be challenging and further drain researcher time, thereby considerably slowing down the workflow[8].

In the excess solvent method, a titration-like procedure is in place where step addition of solvent continues until a fully dissolved state is observed. Composition of the solution and dissolution point are then typically measured using light scattering or imaging-based technology[9]. The excess solvent method is therefore advantageous in that it may remove

the need for pre-calibrated and expensive analytical methods like HPLC, but it comes at the cost of adding time-consuming experimental actions (step-addition of solvent).

Recent advances in technology have enabled the development of automated robotic platforms[10,11] that aid in decreasing hands-on time and increasing throughput[7]. While effective when applied to solubility screening[7,12,13], most existing systems are only partially automated and still require human intervention and interpretation at many stages of the workflow. In addition, these systems continue to rely on analytical techniques that require expensive equipment and time-consuming method development as well as calibration prior to the solubility screen.

The vast majority of automated systems rely on HPLC, the most common analytical instrument applied in excess solid strategies[14–18]. While reliable and a standard method for quantifying the amount of material in a solution, HPLC has several drawbacks -- the most important for high throughput and automation being that it requires calibration. A UV chromophore must be present on each set of test materials and researcher time is spent developing the customized methods necessary for accurate analysis.

Other automated platforms have used in situ techniques for analysis. For example, Dinter et. al.[19] have reported a solubility screening platform using a fiberoptic turbidity probe for qualitative solubility measurements, categorizing the solution as soluble, turbid or uncertain within a range. Other groups have reported using a UV plate reader for solubility measurements in both absorption and light scattering modes[8,17]. Dehring et. al. [20] developed a system that uses stepwise addition of solvent and laser-based nephelometry for high throughput measurement of aqueous solubility. However, a human researcher is still required for solid dosing and visual checks. Image-based analysis in both the excess solid and excess solvent strategies has also been used. Stukelj et. al. [21–23] have reported an image-based single-particle- analysis method to measure amorphous solubility. Reus et. al. [9] have used vision to determine clear points based on images taken by a ReactIR probe. While less time consuming and hands-on than HPLC, such methods still require calibration and the addition of a probe or controlled emission of light, which may not be accurate due to limited detection[9].

Here we present a modular, closed-loop robotic platform for automating solubility screening by combining a titration (excess solvent) method with simple computer vision**.** Unlike the majority of existing platforms, this system is HPLC-free, relying solely on images recorded by the webcam and a custom algorithm to monitor turbidity and determine when the solid material in a solution is dissolved. It runs in a fully-automated way, going all the way from setup to data after an initial user input (< 2min) without further intervention. No additional specialized knowledge or equipment is needed for analysis. Solubility is calculated to a narrow range through iterative feedback, without the need for calibration or human intervention. The platform's design is purposefully modular, enabling its easy adaptation to other workflows such as crystallization and multi-solvent systems. In addition, the system can be accessed via remote desktop and a security camera, which easily allows the user to have a live view of the experiment or perform further development on the system remotely (see SI for further detail).

To demonstrate the capabilities of our system, we used it to screen the solubility of caffeine in five different solvents: acetone, methanol, ethanol, water and acetonitrile. After

initial human input, the system ran uninterrupted for 36 hours and produced solubility values consistent with those obtained by other methods.

Results

Automation Platform Modules

Our automated platform consists of five basic modules: 1) a robotic arm, 2) liquid dosing, 3) solid dosing, 4) stirring and heating, and 5) computer vision (**Fig. 1 A**). Each module performs an essential component of solubility testing -- manipulating objects and dosing material, controlling reaction conditions, and analyzing particular solute/solvent combinations. While fully integrated and run from a master script, the modules are designed to function independently, thereby enabling easy reconfiguration or expansion to suit a given set of experimental conditions or needs. All modules are located atop a robotic deck (**Fig. 1 B**) equipped with additional hardware components and general-purpose tools that help the modules interact (refer to SI).

N9 Robotic Arm

The N9 is a customizable robotic platform with an arm equipped with a rotatable gripper and a probe, capable of 10μm repeatability and a maximum velocity of ~1 m/s. The arm consists of 3 elbows and can move in three dimensions. The gripper is used to transfer, uncap, and recap HPLC vials, and the probe is used to handle needles for liquid dosing. Different hardware modules are placed on the N9 deck and can be engaged by the arm to run experiments.

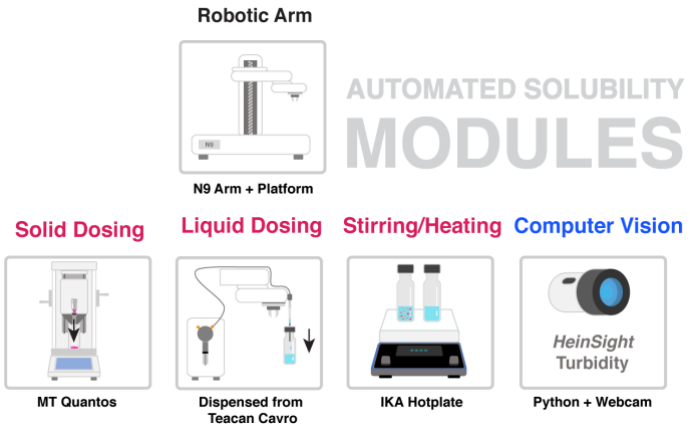
Liquid dosing

For liquid handling, a two-port Tecan Cavro syringe pump is connected to the N9 controller. One port is connected to a solvent reservoir, and the other to the N9 probe, enabling sampling to and from any vial on the deck. The solvent reservoir can contain any solvent, acetonitrile in our case, and was used to prime the tubing line prior to any liquid dosing. The syringe for the pump could hold 1 mL and accurately deliver between 0.01 mL to 0.95 mL. Solvents are located on the deck in vials capped by pre-slit septa to minimize solvent evaporation and contamination while allowing sampling with a needle on the end of the N9 probe.

Solid Dispensing

In a solubility screening workflow, it is important to accurately dose a specified amount of solid. An on-deck solid dosing system extends the capabilities of an automated solubility screening platform. However, it is not common in automation procedures. For example, Alsenz et al. [16] developed a partially automated system and dissolved the solids in the beginning of the experiment to avoid this usually labour intensive step. In this platform, we make use of Quantos powder dosing system by Mettler Toledo. Aside from dosing, Quantos can also be used as a balance throughout the workflow for gravimetric feedback. This direct measurement of mass means that everything dispensed, both solids and liquids, can be accurately measured so that we can gravimetrically calculate solubility. For the solubility screens in this paper the dosing head of the Quantos was filled with caffeine.

(A)



(B)

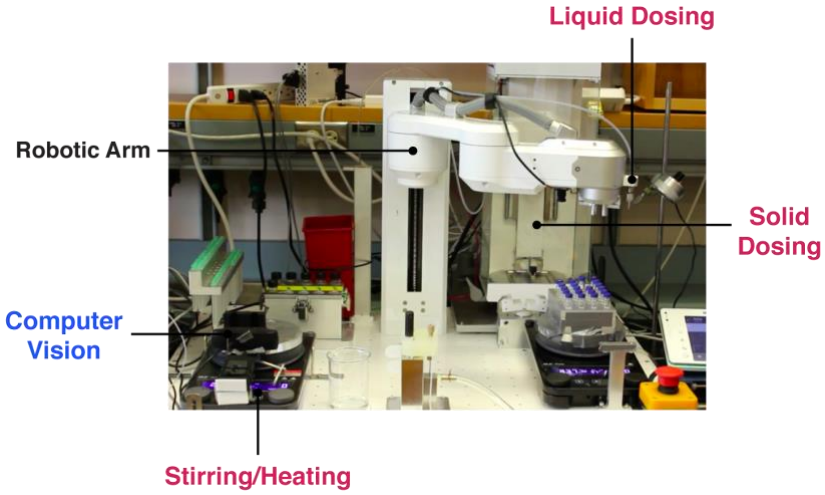


Fig. 1. Automated solubility screening modules. (A), Modules on the platform including the specific equipment used, **(B),** A front view of module positioning on the robotic deck.

Stirring and Heating

An IKA RCT digital stir plate is used to stir vials containing solvent and solute mixtures. The computer is connected to the stir plate via USB, and stir plate control is available via the IKA Python package developed by our group. The plate is also capable of heating but all the tests performed in this paper occurred at room temperature. To prevent the vials from moving on the plate as they are stirred, vials are placed in a vial holder called the "vision station". The holder is designed with an open front to enable visual monitoring of the vial.

Computer vision

The computer vision component of the platform relies on HeinSight, an open-source Python package previously developed by our group to automate common laboratory tasks usually based on visual observation[24]. It is responsible for a key component in our automated workflow: monitoring and measuring turbidity. The turbidity module, specifically designed for the purpose of solubility screening, analyzes the turbidity of a

solution from recorded images without the need to build a reference library or additional calibration. Images captured by a basic webcam are processed using a custom algorithm that measures the average brightness of a given solute/solvent combination. Most image-based turbidity detection methods rely on a standard comparison library that must be hand-built using laboratory instruments[25]. However, because turbidity is an optical property characterized by how the suspended and dispersed solids scatter light[26], and most organic systems are composed of a clear solvent and a white solid, we were able to use this far simpler method to obtain a proxy value. The value output by our algorithm is then referenced against an image of pure solvent and a decision is made as to whether the solution is fully dissolved or not (**Fig. 2**).

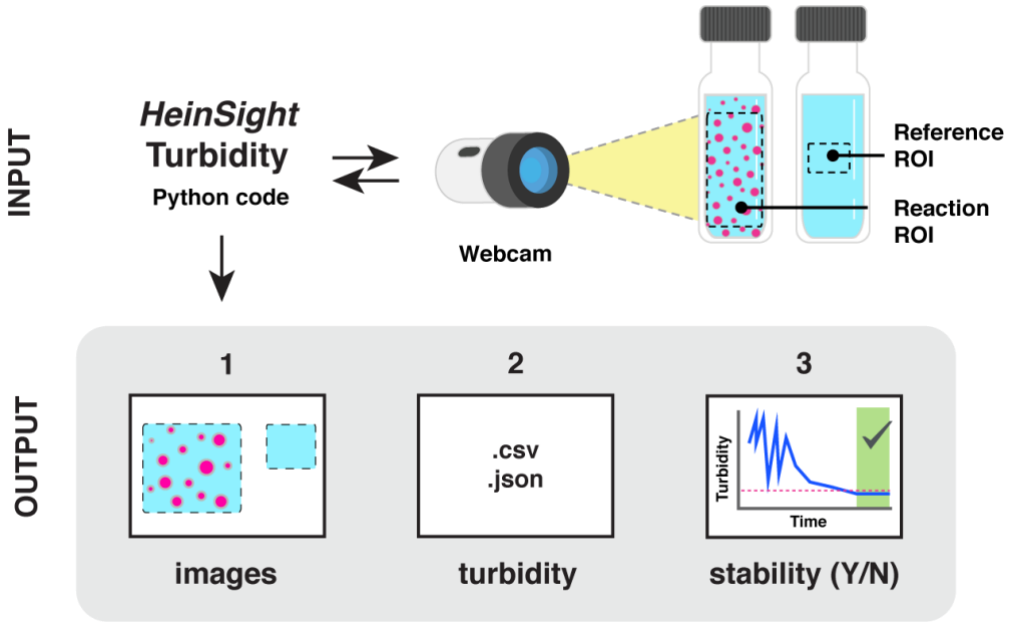


Fig. 2. Overview of HeinSight component of the computer vision module. The HeinSight turbidity code reads in images captured by a webcam and analyzes the pixels within a preselected region of interest to arrive at a proxy turbidity value. This value is then used to determine if a solution is 1) stable and 2) fully dissolved. Recorded images, turbidity data, and graphs of turbidity over time are saved as the output. In the turbidity graph the reference lines are graphed as a dashed red line, and when stability is found, the background area is shaded green.

All turbidity monitoring takes place at a dedicated vision station consisting of a webcam, IKA stir plate, and custom vial holder (for more detail on the algorithm, control logic, and specific equipment used, see the Camera and Computer Vision Methods section of Materials and Methods). A stand-alone subgroup of scripts is responsible for regular vision checks to ensure that the robot arm is handling needles and vials as expected.

Automated Workflow

The modules are integrated in a closed-loop workflow that is designed to mimic how a human researcher would manually conduct experiments. Our automated solubility screen workflow was run via a Python script and each module was controlled using our custom code packages. There are three stages to the workflow: **prepare, monitor, and process (Fig 3)**. In the preparation stage the researcher sets the experiment parameters and the

region of interest (ROI) to monitor turbidity for the rest of the workflow. After that, the N9 arm prepares a vial with the solid to dissolve and an initial volume of solvent and places it in the vision station. In the monitoring stage, the arm periodically adds a set volume of solvent into the vials while the webcam visually monitors the solution turbidity until it either looks fully dissolved or the maximum safe volume in the vial has been reached. Finally, in the processing stage the vial is weighed and cleaned up, data are processed to determine the solubility of the solute and solvent combination, and experiment parameters are automatically selected for the next experiment.

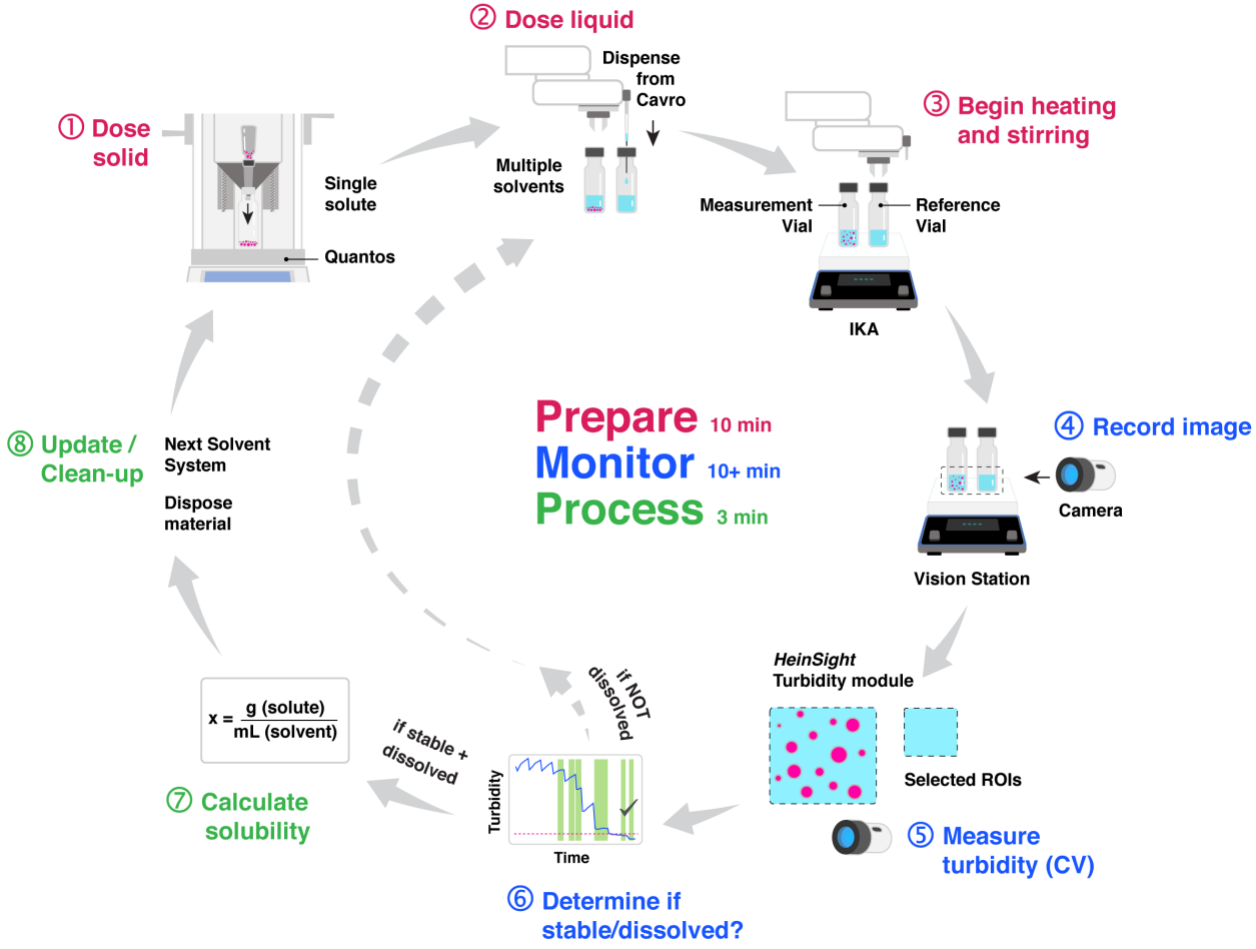


Fig. 3. Automated workflow overview. (1) A fresh HPLC vial is uncapped and transferred to the Quantos where a pre-specified amount of solid is dispensed. The vial is then recapped and transferred to the vision station. (2) A pre-specified initial volume of solvent is dispensed into the vial with the solid and stir bar. The system starts stirring in (3). In (4), (5) and (6), turbidity monitoring takes place: computer vision is used to obtain a proxy measure for turbidity based on average brightness. The value is then passed into a control algorithm that determines when the solution is stable and fully dissolved. The value is said to be 'stable' when turbidity is no longer changing and 'dissolved' when turbidity matches that of pure solvent. If the algorithm determines turbidity is stable but that the solution is not yet fully dissolved, the workflow loops back to (2) and more solvent is added. If it determines the solution is fully dissolved, the workflow proceeds to (7) where solubility is calculated based on solid mass and final solvent volume. In step (8), the used vial is returned to the tray and the system updates initial values in preparation for the next experiment.

To illustrate the applicability of this workflow, we used the N9 automated platform to determine the solubility of caffeine in a variety of common organic solvents (acetonitrile, water, acetone, ethanol, and methanol). The experiments were tested with various initial solid values to challenge the control algorithm. **Fig. 4** shows the turbidity of caffeine in acetone over time. The sharp decreases in turbidity correspond to points when additional solvent was added because the solution was determined to be stable. If turbidity has not stabilized and a dissolved state has not been reached after a user set period (in this case, 15 minutes), the next addition of solvent is triggered regardless. This provides a level of flexibility for further studying solubility behaviour by allowing more dissolution time for slow dissolving material.

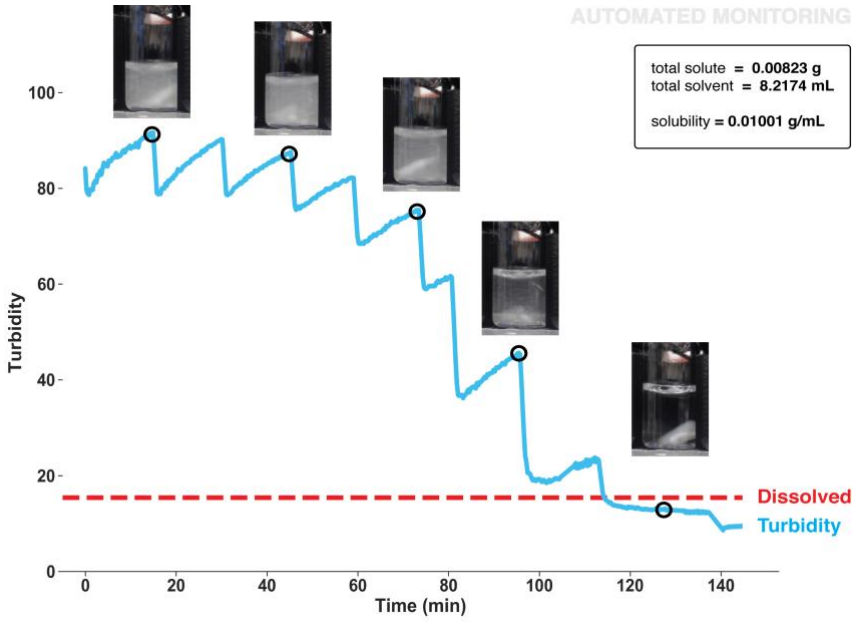


Fig. 4. **Turbidity of caffeine in acetone over time**. Real-time turbidity values are shown in blue; the dotted red line indicates the dissolved reference (turbidity of pure solvent). The stepwise behavior in this experiment can be explained by the incremental addition of solvent. Following each addition, large particles are further broken down by stirring, which visually manifests as increased turbidity. This trend was observed for some solvents at certain initial volumes; different solvents and different initial volumes produced different trend lines. The control algorithm is programmed to trigger a short stir-burst as turbidity nears the dissolved reference to ensure there is no solid left undissolved. The final step around 100 min can be explained by this increase in rate. Once the turbidity value reaches a stable point near or consistently below the dissolved reference, monitoring ceases and the workflow continues to the processing portion of the workflow. Occasionally, the turbidity value will read below the dissolved reference due to fluctuations in lighting (for more information see SI).

The control algorithm includes additional checks that account for similar experimental complexities. For example, once the turbidity value is within a specified range from that of the dissolved reference, the stir rate is increased for five minutes to ensure full dissolution of the solid. Turbidity monitoring continues and the algorithm then re-checks that a dissolved state has truly been reached. The platform ran for almost 2 days without any error or interruption. **Table 1** contains the mean solubility values for 4 measurements done by our automated platform compared to solubility values found by two other

methods, namely the slurry method (HPLC analysis) and titration. The calculated standard error of the mean shows great precision and repeatability of the measurements done by the automated platform.

Table 1. Comparison of solubility values obtained by manual and automated methods

Solvent	Slurry (mg/ml)	Titration 1(mg/ml)	Titration 2(mg/ml)	N9 avg. (mg/ml) ± standard error of the mean
Acetonitrile	21.49	17.6	18.4	19.18 ± 0.811
Water	29.47	17	18.5	18.21 ± 0.404
Acetone	10.52	8.7	7	9.98 ± 0.017
Ethanol	4.04	5	4.5	5.37 ± 0.059
Methanol	9.14	8.9	6.6	8.36 ± 0.81

Duplication

To further showcase the modularity and broad applicability of the developed workflow, we replaced the N9 with a 6-axis robotic arm. The Kinova Gen3 arm has 7 degrees of freedom and is equipped with a Robotiq 2F-85 gripper. **Fig. 5** shows the system set-up with the new arm installed. Because the Kinova Gen3 is supported by a Python package, we were able to apply the master script previously used by N9 to run the solubility workflow to the new arm with minimal adaptation. Required changes included different code related to arm movement as well as a different liquid handling system. A regripping station was added to switch between horizontal and vertical gripping as needed. For liquid handling, reservoirs of solvents were connected to a selection valve and a pump, which was connected to a sampler. The sampler could be picked up by the gripper on the arm and sample or inject. For further detail on the sampler please refer to SI.

Discussion

Here we report on an automated solubility screening platform that uses our HeinSight computer vision system and a feedback algorithm to implement a titration ‘excess-solvent’ workflow. As proof of concept, we used the platform to find the solubility of caffeine in various solvents; however, the system can be used to screen the solubility of any solid with any other solvents. The purposefully modular design allows for easy adaptation and expansion to other robotic platforms or solubility study workflows. For the next iteration of the platform, we have developed a method to automate the exchange of the dosing head on the Quantos, thereby enabling the solubility screen to investigate multiple solute systems within a single workflow. The workflow could likewise be expanded for use in automated crystallization and binary solvent analysis. Furthermore, The modularity of this platform allows flexible automation via any available robotic arm with minimal adaptation. In summary, the modular robotic platform we developed can perform automated solubility screens with minimal material that might be expensive or in short supply such as when screening for pharmaceutical applications. We expect that the

platform can facilitate finding answers to solubility questions in industries dealing with novel materials that would otherwise be very costly.

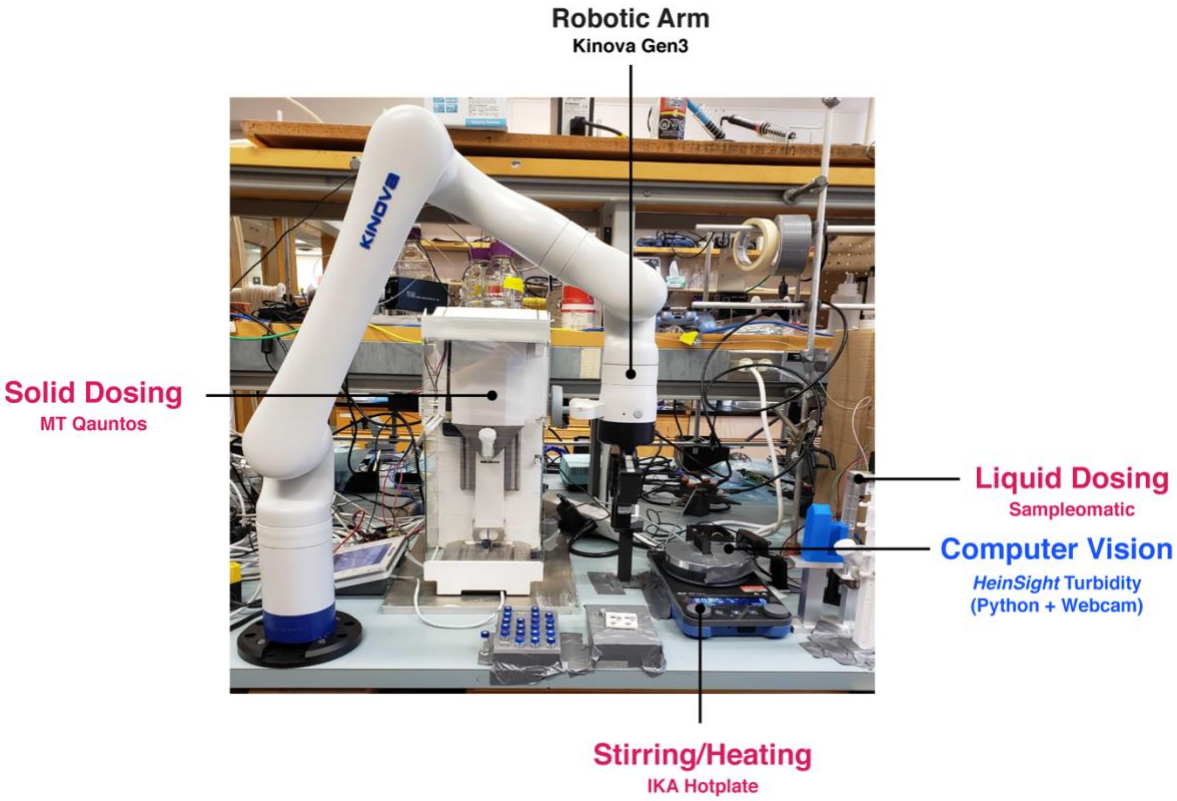


Fig. 5. Automated solubility screening modules, Kinova platform.

Materials and Methods

Materials

Caffeine was purchased from Sigma Aldrich and was used without further purifications. Acetonitrile (≥99.9%), isopropanol (≥99.9%), acetone (≥99.9%), ethanol (≥99.9%), and methanol (≥99.9%) were purchased from Sigma Aldrich and used without further purifications.

Consumable Materials of the platform

Before running the automated system, the deck must be charged with desired solvents, clean empty HPLC vials with stir bars inside them and the solid of interest inside the Quantos powder dispensing head. The choice of HPLC vials are due to their convenient small size, matching the robot gripper size as well as reducing material waste. On the computer, there are CSV files which track the needles left, clean vials and solvent information such as name, density, index, etc. The user has to ensure these files are up to date prior to run. The consumable products include: BD PrecisionGlide needles (21G, 0.8mm x 25mm) , clean 2-HPLC vials (Life Science Labs) charged with stir bars (VP Scientific). Solvents and solid were refilled as well upon running out.

Cameras/Vision Station

Any commercially available USB webcam can be used and controlled by the camera module located in our HeinSight Python package. For this particular application, 2-Logitech C920s were used - one for the vision station and one placed in the lab to provide a birds-eye view of the deck. The vision station consisted of an IKA RCT digital stir plate, a webcam, and a custom 3D printed vial holder.

Robotic Methods

This platform consists of a multipurpose robotic arm that can handle HPLC vials and needles. Liquid handling was performed using a Tecan Cavro syringe pump with a 1 mL syringe. Solid dosing was enabled using a Quantos, a Mettler Toledo balance. An IKA RCT digital stir plate was used to stir the mixtures and a webcam to capture images for computer vision. Other modules on the deck include: needle rack, waste bin and beaker, vial tray, vision station, and HPLC cap park station.

Experiment parameters are set in the script, including the solvents to use, initial mass of solute and initial volume of solvent, step addition volume of solvent, and the stir rate. The N9 solubility screening script is run, and immediately a reference vial of solvent is prepared. As the vial only contains solvent, it is used as a reference for what a completely dissolved solution looks like. The researcher is prompted to select a region in the image to normalize all turbidity measurements, and the region to monitor turbidity. The remainder of the workflow runs as below.

Step 1: solid dose

The arm takes a clean vial from the tray to the gripping station and uncaps the vial. The cap is temporarily parked on a pillar. The arm then transfers the vial to Quantos, which consists of a dosing head mounted on a z stage that can move close to the opening of the vial. There are different types of dosing heads to hold solids with different physical properties. A single dosing head contains a single solid, but the dosing head can be swapped out as needed to dose the solid of interest. Once the vial is placed inside Quantos the balance tares and starts dosing. Dosing capabilities of Quantos fall within 0.1 mg and 8 g, depending on the dosing head used.

After dosing the Quantos z stage drives upwards, allowing the arm to enter the Quantos chamber and bring the vial back to the gripping station for recapping. Recapping uses torque feedback from the arm to tighten the cap by a single degree until the torque associated with a properly tightened cap is reached. For details on Quantos accessories and uncapping and recapping refer to the SI.

Step 2: initial solvent dose

Then the arm transfers the vial to the vision station and doses the initial volume of solvent into the vial. The syringe pump enables liquids to be drawn and dispensed precisely when there is a needle on the end of the probe. The arm can pick up a capped needle from the needle tray and uncap it using an uncapping tooth; for details on the uncapping tooth refer to the SI. To eliminate air bubbles in the line used to draw liquids, the line must first be primed by a backing solvent. Prior to drawing solvent into the lines for dosing a 50 µL air gap is drawn, part of which is injected into the solvent bottle the solvent is stored in to compensate for the volume to be drawn. This is a widely used method in medical procedures to avoid drawing air instead of solvent into the lines. The remaining air is kept to create distance between the drawn solvent and back solvent used to prime the line. After drawing in liquid a second air gap is drawn to prevent solvent from dripping out of the needle during arm movements until the solvent is dispensed.

Step 3: computer vision screening and step addition of solvent

While solution turbidity in the vial is monitored, the needle remains on the arm probe and is used to dose fixed volumes of solvent as directed by the computer vision module. These

additions of solvent comprise the step addition of solvent used to fully dissolve the solid in the vial with volumes ranging from 10 to 50 µL.

Step 4 : transferring back vial and report generation

Once the computer vision module identifies that the solid has fully dissolved in the vial or if the safe volume of the vial is reached, the arm then discards the needle and takes the vial to Quantos to weigh the vial. As the mass of just the vial and solid was previously measured, the mass and thus volume of solvent in the vial can be calculated directly. With the mass of solid and solvent known solubility can be gravimetrically determined. This also means that discrepancies between the volume that the script commanded the pump to dispense and the actual volumes dispensed are accounted for when calculating solubility. All the data gathered from the experiment is saved in a report file and key data are sent through Slack as real time updates.

Step 5: determining the next experiment

For the next experiment, the experiment parameters are discussed in the closed loop control logic section of the SI. The duration of each experiment varies depending on when the dissolved state or the safe solvent volume in the vial is reached. Since the maximum duration between each step addition of solvent was set to 10 minutes, the safe volume of the vial was set to 1.2 mL, and the initial solvent volume added was 0.4 mL, the longest an experiment could take was approximately 90 minutes including arm movement time.

Computer Vision Methods

Turbidity Monitoring

The HeinSight turbidity module is fully integrated into the automated workflow and includes the computer algorithm that determines control decisions based on visual input. **Fig. 5** shows how the steps involved in turbidity monitoring fit within the larger automated workflow.

Step 1: Select the regions of interest

Once the script is run, the user is prompted to select two regions of interest (ROI). The first, the reaction ROI, defines the area within which the algorithm will monitor turbidity and capture the dissolved reference. The second ROI, the normalization ROI, is used to normalize brightness for the turbidity value measurement as described in step 2 below (**Fig. 6**). Completion of the ROI step prompts the robot to fill a vial with pure solvent, which is transferred to the vision station and measured to obtain the dissolved reference value.

Step 2: Capture image and measure turbidity

Images are captured by the webcam and analyzed using a custom turbidity algorithm that proceeds in two steps: 1) pre-process and 2) calculate average brightness in the pre-selected ROI. To combat issues that may arise due to environmental changes in lighting, the average brightness in the ROI is normalized using a stable region in the image. While turbidity values for each image are recorded, the value used to determine stability (step 3) is an average across multiple consecutive images. This further reduces noise that may arise due to lighting changes or artifacts in the image.

Step 3: Determine if stable

A control algorithm checks for trends in turbidity values over time and determines if the solution is stable based on statistical features (mean, mode, standard deviation, and range) of the most recent turbidity measurements. If the solution is determined to be stable and has not yet

reached the dissolved reference or maximum volume, more solvent is added and monitoring continues.

Step 4: Determine if dissolved

If the solution is determined to be stable, the control algorithm checks for a fully dissolved state by comparing the current turbidity value to the dissolved reference. If "dissolved," the algorithm triggers a stir burst to agitate any solute that may have settled at the bottom of the vial. Once turbidity has stabilized at the reference value or if the five most recent minutes of turbidity data are all below the reference value then the solution is classified as dissolved. Details of the stir burst and statistical requirements for determining stable and dissolved states can be found in the SI.

Data

A non-trivial advantage of using computer vision as our analytic technique is the ease of creating a visual record, which can be reviewed to help identify and explain trends or anomalies in the data. The HeinSight turbidity module automatically saves images from each run and converts them into a graph of turbidity vs time (**Fig. 4**). Turbidity values, ROI parameters, and the initial mass and volume for each solute/solvent combination that were set in the code, as well as final solubility determined are saved as CSV and JSON files that are automatically generated at the start of each run and continuously updated as more data is gathered.

References and Notes

(1) Coltescu, A.-R.; Butnariu, M.; Sarac, I. The Importance of Solubility for New Drug Molecules. *Biomedical & pharmacology journal* **2020**, *13* (2), 577–583. https://doi.org/10.13005/bpj/1920.

(2) Hansen, C. M. The Universality of the Solubility Parameter. *Product R&D* **1969**, *8* (1), 2–11. https://doi.org/10.1021/i360029a002.

(3) Hansen, C. M. 50 Years with Solubility Parameters—Past and Future. *Progress in organic coatings* **2004**, *51* (1), 77–84. https://doi.org/10.1016/j.porgcoat.2004.05.004.

(4) Walker, B.; Tamayo, A.; Duong, D. T.; Dang, X.-D.; Kim, C.; Granstrom, J.; Nguyen, T.-Q.; Center for Energy Efficient Materials (CEEM); Energy Frontier Research Centers (EFRC). A Systematic Approach to Solvent Selection Based on Cohesive Energy Densities in a Molecular Bulk Heterojunction System. *Advanced energy materials* **2011**, *1* (2), 221–229. https://doi.org/10.1002/aenm.201000054.

(5) Machui, F.; Langner, S.; Zhu, X.; Abbott, S.; Brabec, C. J. Determination of the P3HT:PCBM Solubility Parameters via a Binary Solvent Gradient Method: Impact of Solubility on the Photovoltaic Performance. *Solar energy materials and solar cells* **2012**, *100* (Journal Article), 138–146. https://doi.org/10.1016/j.solmat.2012.01.005.

(6) Black, S.; Dang, L.; Liu, C.; Wei, H. On the Measurement of Solubility. *Organic process research & development* **2013**, *17* (3), 486–492. https://doi.org/10.1021/op300336n.

(7) Selekman, J. A.; Qiu, J.; Tran, K.; Stevens, J.; Rosso, V.; Simmons, E.; Xiao, Y.; Janey, J. High-Throughput Automation in Chemical Process Development. *Annual review of chemical and biomolecular engineering* **2017**, *8* (1), 525–547. https://doi.org/10.1146/annurev-chembioeng-060816-101411.

(8) Hoelke, B.; Gieringer, S.; Arlt, M.; Saal, C. Comparison of Nephelometric, UV-Spectroscopic, and HPLC Methods for High-Throughput Determination of Aqueous Drug

Solubility in Microtiter Plates. *Analytical chemistry (Washington)* **2009**, *81* (8), 3165–3172. https://doi.org/10.1021/ac9000089.
(9) Reus, M. A.; van der Heijden, A. E. D. M.; ter Horst, J. H. Solubility Determination from Clear Points upon Solvent Addition. *Organic process research & development* **2015**, *19* (8), 1004–1011. https://doi.org/10.1021/acs.oprd.5b00156.
(10) Li, J.; Lu, Y.; Xu, Y.; Liu, C.; Tu, Y.; Ye, S.; Liu, H.; Xie, Y.; Qian, H.; Zhu, X. AIR-Chem: Authentic Intelligent Robotics for Chemistry. *The journal of physical chemistry. A, Molecules, spectroscopy, kinetics, environment, & general theory* **2018**, *122* (46), 9142–9148. https://doi.org/10.1021/acs.jpca.8b10680.
(11) Burger, B.; Maffettone, P. M.; Gusev, V. V.; Aitchison, C. M.; Bai, Y.; Wang, X.; Li, X.; Alston, B. M.; Li, B.; Clowes, R.; Rankin, N.; Harris, B.; Sprick, R. S.; Cooper, A. I. A Mobile Robotic Chemist. *Nature (London)* **2020**, *583* (7815), 237–241. https://doi.org/10.1038/s41586-020-2442-2.
(12) Sou, T.; Bergström, C. A. S. Automated Assays for Thermodynamic (Equilibrium) Solubility Determination. *Drug discovery today. Technologies* **2018**, *27* (Journal Article), 11–19. https://doi.org/10.1016/j.ddtec.2018.04.004.
(13) Goodwin, J. J. Rationale and Benefit of Using High Throughput Solubility Screens in Drug Discovery. *Drug discovery today. Technologies* **2006**, *3* (1), 67–71. https://doi.org/10.1016/j.ddtec.2005.03.001.
(14) Qiu, J.; Patel, A.; Stevens, J. M. High-Throughput Salt Screening of Synthetic Intermediates: Effects of Solvents, Counterions, and Counterion Solubility. *Organic process research & development* **2020**, *24* (7), 1262–1270. https://doi.org/10.1021/acs.oprd.0c00132.
(15) Wenlock, M. C.; Austin, R. P.; Potter, T.; Barton, P. A Highly Automated Assay for Determining the Aqueous Equilibrium Solubility of Drug Discovery Compounds. *Journal of Laboratory Automation* **2011**, *16* (4), 276–284. https://doi.org/10.1016/j.jala.2010.10.002.
(16) Alsenz, J.; Meister, E. v. a.; Haenel, E. Development of a Partially Automated Solubility Screening (PASS) Assay for Early Drug Development. *Journal of pharmaceutical sciences* **2007**, *96* (7), 1748–1762. https://doi.org/10.1002/jps.20814.
(17) Chen, T.; Shen, H.; Zhu, C. Evaluation of a Method for High Throughput Solubility Determination Using a Multi-Wavelength UV Plate Reader. *Combinatorial chemistry & high throughput screening* **2002**, *5* (7), 575–581.
(18) 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. *JALA: Journal of the Association for Laboratory Automation* **2005**, *10* (6), 364–373. https://doi.org/10.1016/j.jala.2005.06.003.
(19) Dinter, C.; Schuetz, A.; Blume, T.; Weinmann, H.; Harre, M.; Neh, H. Automated Solubility Determination Using a Customized Robotic System and a Turbidity Probe. *Journal of Laboratory Automation* **2005**, *10* (6), 408–411. https://doi.org/10.1016/J.JALA.2005.08.006.
(20) Dehring, K. A.; Workman, H. L.; Miller, K. D.; Mandagere, A.; Poole, S. K. Automated Robotic Liquid Handling/Laser-Based Nephelometry System for High Throughput Measurement of Kinetic Aqueous Solubility. *Journal of pharmaceutical and biomedical analysis* **2004**, *36* (3), 447–456. https://doi.org/10.1016/j.jpba.2004.07.022.
(21) Štukelj, J.; Agopov, M.; Yliruusi, J.; Strachan, C. J.; Svanbäck, S. Machine-Vision-Enabled Salt Dissolution Analysis. *Analytical chemistry (Washington)* **2020**, *92* (14), 9730–9738. https://doi.org/10.1021/acs.analchem.0c01068.

(22) Štukelj, J.; Svanbäck, S.; Agopov, M.; Löbmann, K.; Strachan, C. J.; Rades, T.; Yliruusi, J. Direct Measurement of Amorphous Solubility. *Analytical chemistry (Washington)* **2019**, *91* (11), 7411–7417. https://doi.org/10.1021/acs.analchem.9b01378.
(23) Štukelj, J.; Svanbäck, S.; Kristl, J.; Strachan, C. J.; Yliruusi, J. Image-Based Investigation: Biorelevant Solubility of α and γ Indomethacin. *Anal. Chem.* **2019**, *91* (6), 3997–4003. https://doi.org/10.1021/acs.analchem.8b05290.
(24) Zepel, T.; Lai, V.; Yunker, L. P. E.; Hein, J. E. Automated Liquid-Level Monitoring and Control Using Computer Vision. **2020**. https://doi.org/10.26434/chemrxiv.12798143.v1.
(25) Liu, Y.; Chen, Y.; Fang, X. A Review of Turbidity Detection Based on Computer Vision. *IEEE access* **2018**, *6* (Journal Article), 60586–60604. https://doi.org/10.1109/access.2018.2875071.
(26) Lawler, D. M. Turbidity, Turbidimetry, and Nephelometry. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering* **2013**, No. Generic. https://doi.org/10.1016/B978-0-12-409547-2.11006-6.

Acknowledgments

The authors would like to thank Mettler-Toledo for their contributions to Quantos application and API development. We are also grateful for the UBC Chemistry machine shop assistance with instrument fabrication. For robot control and data processing, we would like to acknowledge the contributors to the Python programming language (Python Software Foundation; https://python.org). The authors acknowledge Mike Lovette at Amgen Inc. for providing reference solubility data obtained by the slurry method. **Funding**: The authors are grateful to the Defense Advanced Research Projects Agency (DARPA) for funding this project under the Accelerated Molecular Discovery Program under Cooperative Agreement No. HR00111920027. The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. Additional financial support for this work was provided by the University of British Columbia, the Canadian Foundation for Innovation (CFI-35883), the Natural Sciences and Engineering Resource Council of Canada (RCPIN-2016-04613) and Natural Resources Canada (EIP2-MAT-001). **Data and materials availability:** Further details on the custom built parts as well as the collected data can be found in supplementary information document. Further data is accessible upon request. **Authors Contribution:** J. E. H. and D. G. conceived the project. J. E. H. supervised the project. P. S. and S. G. developed the robotic workflow. V. L. and T. Z. developed the machine vision. P. S. and V. L. developed the master script and P.S., V. L. and S. G. designed and performed the experiments. J. R. collected the non-robotic solubility data. L. P. E. Y. , H. S., S. S., S. K. and V. L. developed the software packages. S. S. and H. S. developed the robotic hardware. P. S. and F. Y. developed the robotic workflow on Kinova. All authors participated in manuscript writing.

Automated solubility screening platform using computer vision (Supplementary Information)

P. Shiri[1], V. Lai[1], T. Zepel[1], D. Griffin[2], J. Reifman[2], S. Clark, S. Grunert[1], L. P. E. Yunker[1], S. Steiner[1], H. Situ[1], F. Yang, P. L. Prieto[1], J. E. Hein[1,*]
[1]Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada
[2]Amgen Inc. Cambridge, MA, 02141 USA
[*]Corresponding author Email: jhein@chem.ubc.ca

Robotic Hardware

Fig. S1 shows the rendered view of the deck. The robotic deck initially consisted of the robotic arm (N9, North Robotics, www.northrobotics.com), the arm controller box (C9, North Robotics, www.northrobotics.com, not shown in the picture) and the deck. The C9 includes pneumatic outlets to control both the gripper on the arm as well as the gripping station on the deck. The rest of the parts were added as they were needed to build the workflow.

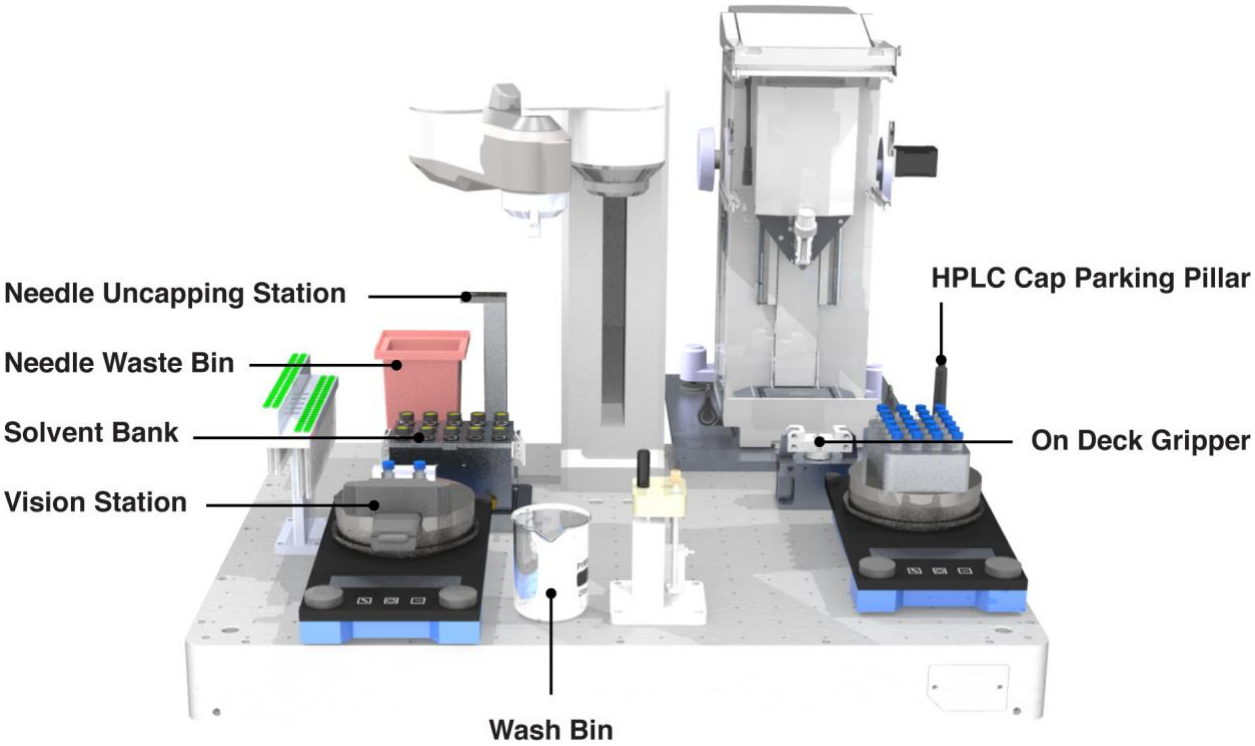


Fig. S1. Rendered illustration of the robotic deck.

Quantos accessories

A Quantos Balance XPE205 (220 g capacity, 0.01 mg readability) with powder dosing unit and anti-static kit were purchased from Mettler Toledo. The Quantos enables dosing on a z-axis through a variety of dosing heads designed for different texture of solid and dosing values. BMNP and LNMP dosing heads used for the implementation of the caffeine solubility workflow. Quantos control was automated using a custom Python package developed by our group. Using this package, we were able automate taring, powder dosing and weighing as well as gain control over door position, z stage position and dosing commands.

Z stage control

A dosing gun was required to automatically approach the opening of the vial. An Arduino controller was used to drive the z stage by a mounted stepper motor. The motor was mounted by a machined supporting plate and the assembly is shown in **Fig. S2**. As the z stage goes up to home position, it hits a limit switch mounted on inside Quantos. The switch mounting plates were waterjet cut from aluminium sheet and the assembly is shown in **Fig. S3**. Then, the number of steps needed to lower the stage to the right height is specified in the code.

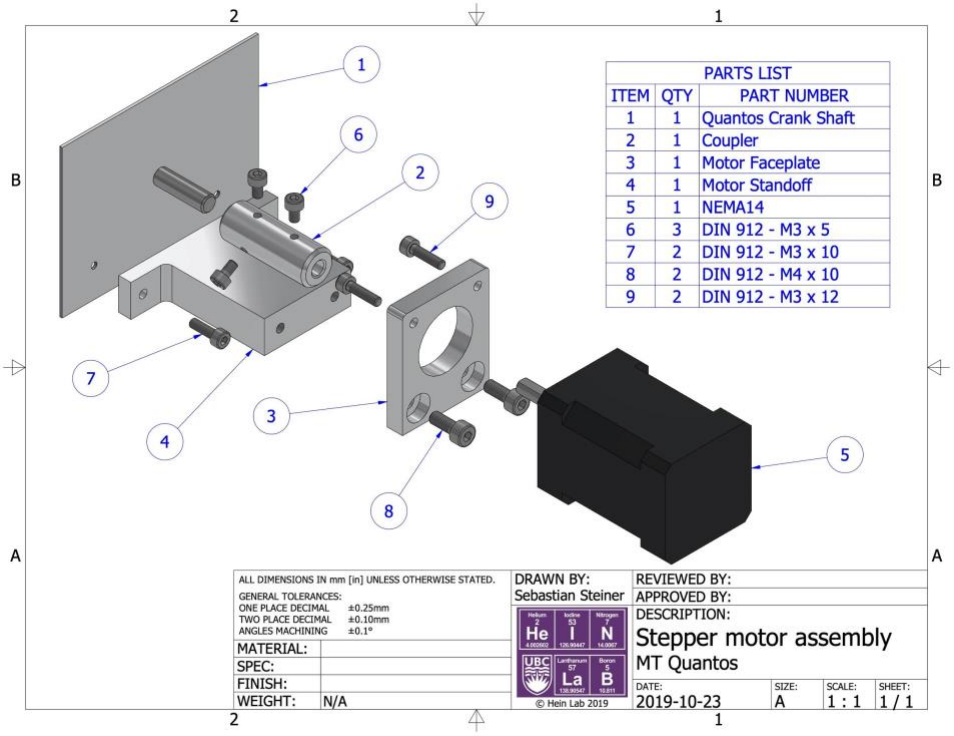


Fig. S2. Stepper motor assembly

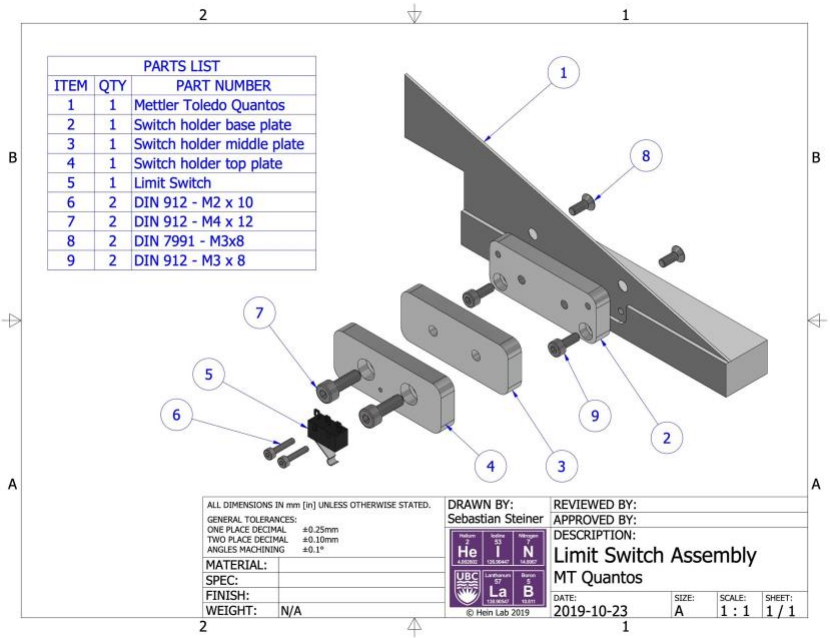


Fig. S3. Limit switch assembly

Vial holder and mounting plates

A 3D printed HPLC vial holder was designed and placed inside the Quantos (**Fig. S4**). In order to insert vials inside Quantos with the arm, the Quantos was mounted on two auxiliary aluminium plates screwed into the deck and supported by a back foot. The aluminium plates were cut by waterjet.

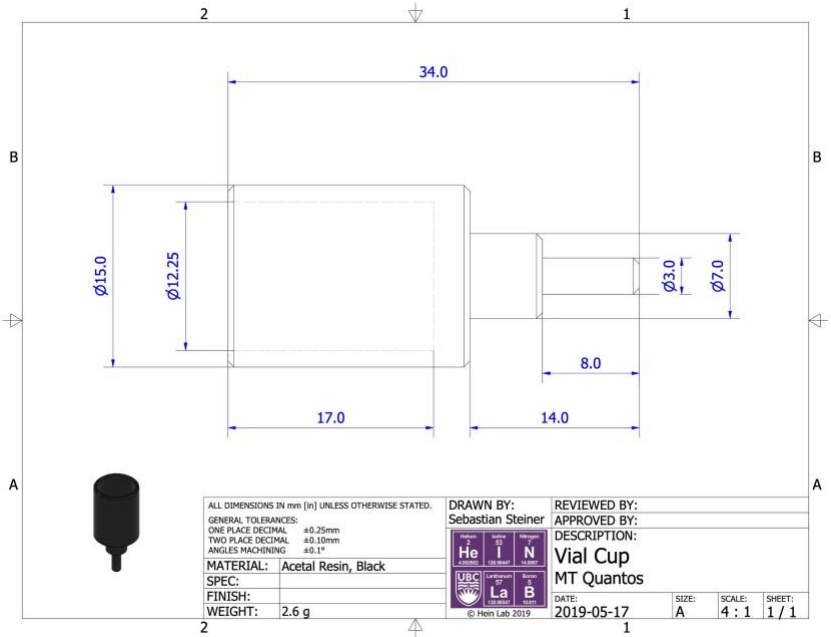


Fig. S4. Quantos vial holder

Uncapping

The ability to uncap and recap the vial is a key capability in an automated platform. For our automated workflow, vials needed to be uncapped prior to solid dosing and recapped prior to turbidity monitoring due to the volatility of the solvents. For any uncapping and recapping, the vial was first moved to a separate static gripper to hold onto it. The N9 arm was capable of uncapping and recapping the vial caps by performing a simultaneous upward and rotational movement at high speed of 100 m/s, which enabled it to perform this movement in a second. In order to be able to park the cap so that the arm's gripper was free to handle vials, a pillar was designed and 3D printed (**Fig. S5**). The cap was parked on this pillar after uncapping and was picked up right after uncapping. In order to ensure the vials were properly capped, we implemented torque feedback reading from the N9 arm's gripper. First, the reading associated with properly recapping a vial was identified. Then that value was used as a threshold after the recapping step: if the torque read was smaller than the threshold, the arm's gripper tightened the cap by a 1 degree rotation. The tightening continued until the torque threshold was reached.

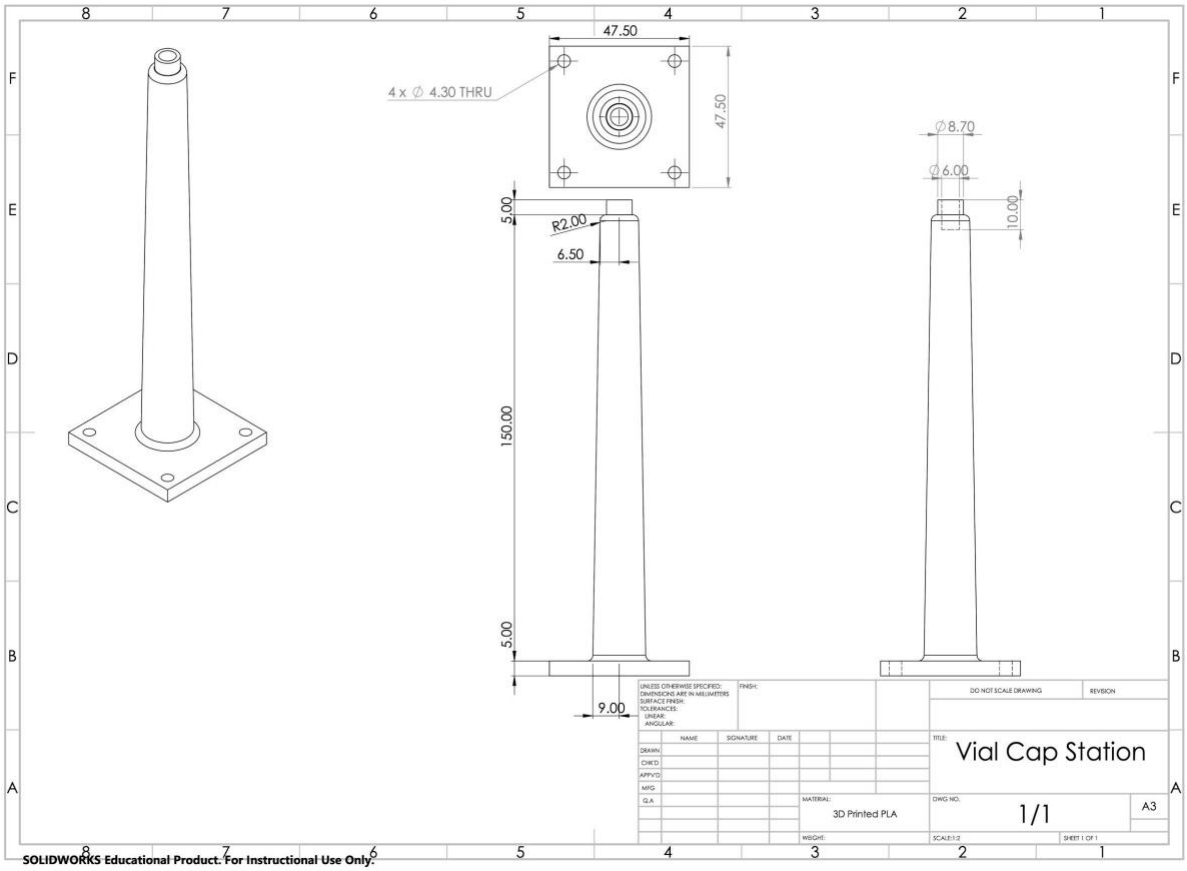


Fig. S5. HPLC cap parking pillar

Vision station

Our vision station is a custom designed and 3D printed vial holder (**Fig. S6**) placed on the stir plate to enable simultaneous stirring, imaging, and solvent addition. A collar acted as a stopper

to keep capped vials in place when piercing through them with a needle to dose liquids. The open front enables seeing through the vial as it stirs. The collar has a 0.5 mm margin of freedom for vial placement and pickup by the N9 arm.

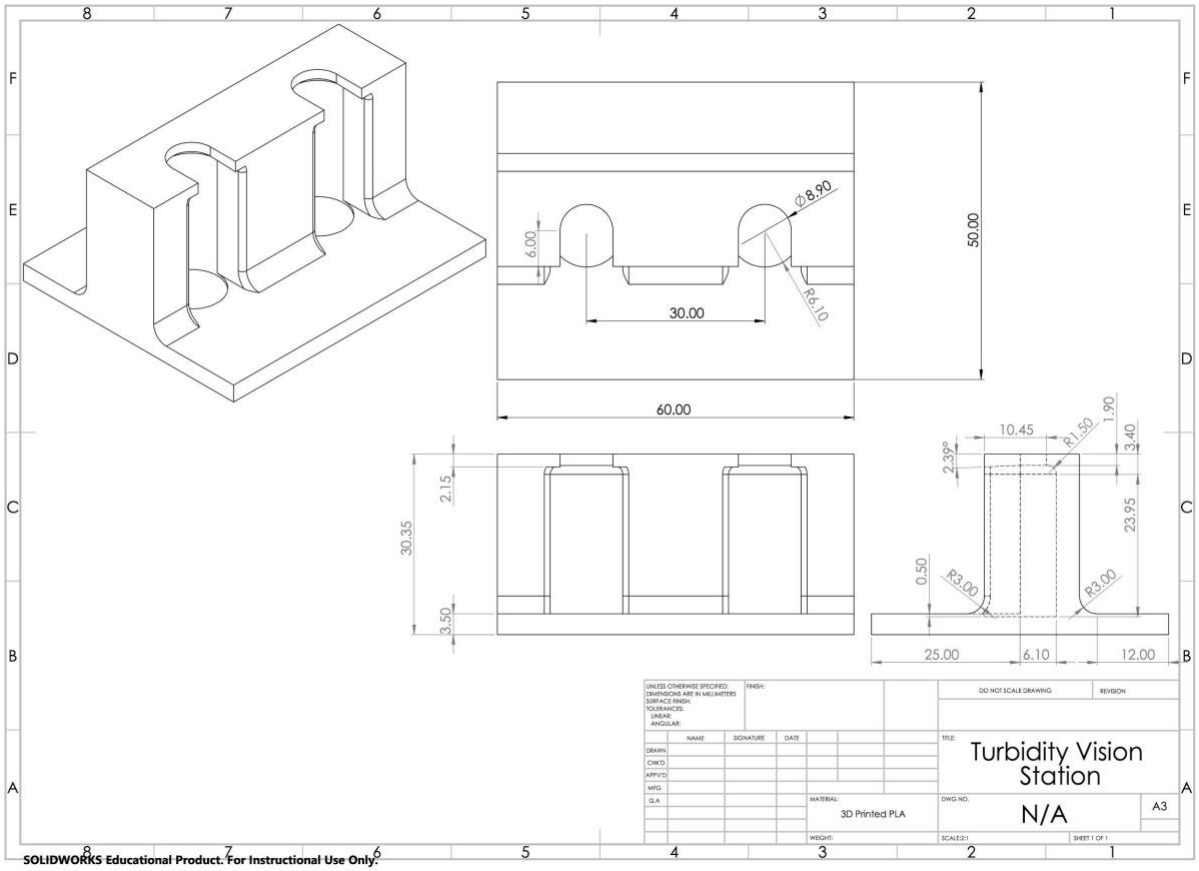


Fig. S6. Vision station

Remote Control and Development

The system computer was accessed via Chrome remote desktop while a surveillance camera provided a visual of the platform's actions during remote operation. Once the instruments were placed and their robot reach locations recorded, development and testing occurred remotely. Our system also made use of Slack, an online messaging platform, which enabled on-the-fly control of the system and provided real time monitoring capabilities along with updates and final reports (**Fig. S7**) for each experiment. **Table S1** shows the programmed commands that can be sent via Slack to access and control the system during an active run.

Table S1. Slack commands for live control and manipulation of the system

vial image	Send the current image from camera aimed at vials in the vision station
deck image	Send the current image from camera aimed at the deck
get needle	Get and uncap a needle
uncap needle	Uncap a needle already on the probe
dump needle	Remove a needle
main resume	If the run paused at a specific point, use this to resume the run
graph	Send a graph of turbidity vs. time
add solvent	Add a step addition volume of solvent into a vial in the vision station
clean up vial	Clean up the current vial
start next vial	Start the next experiment, clean up the current vial if required
don't start next vial	Don’t start the next experiment
roi	Send the current image from camera with the region of interest drawn on the image aimed at the vision station
change stir rate [number]	Change the stir rate to the specified number
set dissolved reference [number]	Change the fully dissolved turbidity reference to the specified number
turbidity video	Send a video from all the images used to make turbidity measurements from the camera aimed at the vision station for the current run
use next solvent	For the next run, use the next solvent in the sequence to test
pause solvent addition	Pause solvent addition
resume solvent addition	Resume solvent addition if paused
pause monitoring	Pause turbidity monitoring
resume monitoring	Resume turbidity monitoring if paused

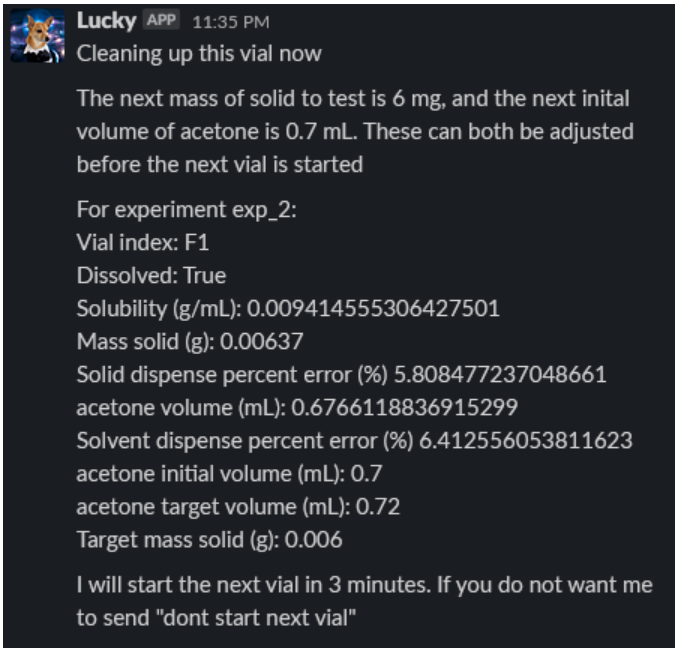


Fig. S7. Sample report generated and sent through slack

Vision Checks

After picking up and uncapping or removing a needle, the N9 arm moves in front of a webcam for a visual needle check to ensure the action was completed correctly. This is done using the colour matcher and ROI modules in our HeinSight Python package. If the needle was not properly picked up and uncapped or removed, the automated workflow is paused and the researcher is notified to correct the situation via Slack. See the Visual Needle Detection section of SI for more detail.

Visual Needle Detection

Visual needle detection uses the colour matcher and ROI modules in our HeinSight Python package and allows for real time error handling/correction of needle pick up with Slack integration. The N9 will pick up a needle then move to a specific position on the deck where the webcam takes a photo. The colours in this photo are compared to a reference photo where the needle was picked up successfully (Fig. S10). If the average colour within the region where the needle is expected to be is similar to the colour in the reference photo, then it is concluded that the needle is on the N9 probe. In testing, we have seen that this method for checking for the needle is able to correctly detect when the needle was not picked, and when it was not uncapped successfully.

Real time visual needle detection combined with Slack integration will prevent the script from continuing until a researcher has sent a Slack command that it can. In the case where a needle was not uncapped, picked up, or removed properly, this would be detrimental to the rest of the experiment, but through Slack the researcher can command the N9 arm to perform all these actions and then allow the script to resume.

The visual check only needs to be set up once before it can be used multiple times. For set up, the N9 needs to pick up a needle on the probe and bring it to the check location for three different needle states: capped, uncapped, and after the needle has been removed. Then the researcher is presented with the image where the needle is on the probe and is prompted to interactively select the region in the image that will be used for the needle check. Afterwards, a file with all the parameters for needle detection are saved, which can be loaded in for future use. The results of the needle detection on the three collected images are displayed to the user, and if the results from detection do not match up with the actual state, the researcher might have to adjust the needle detection parameters, lighting conditions, or the position the N9 goes to for the check.

Sampleomatic and Liquid Handling Setup on Kinova

When duplicating the workflow to use with the Kinova arm, a slightly different liquid handling system was implemented. **Fig. S8** shows a diagram of the set-up. In this set-up, a number of solvent bottles are connected to a selection VICI valve, which is connected to a Tecan Cavro pump. In this system, instead of single use needles, a custom design sampler was used called the “sampleomatic” (**Fig S9**). The sampleomatic consists of a top and a base part. The top part consists of a needle, driven by an Arduino controlled motor, with a handle which enables the robot arm to pick it up. The base has a hole through and a drain outlet to allow flushing the line for rinsing the needle. The handle was initially designed for N9 gripper. As this was used with the Kinova arm setup, a different handle compatible with the Robotiq gripper was 3D printed and used.

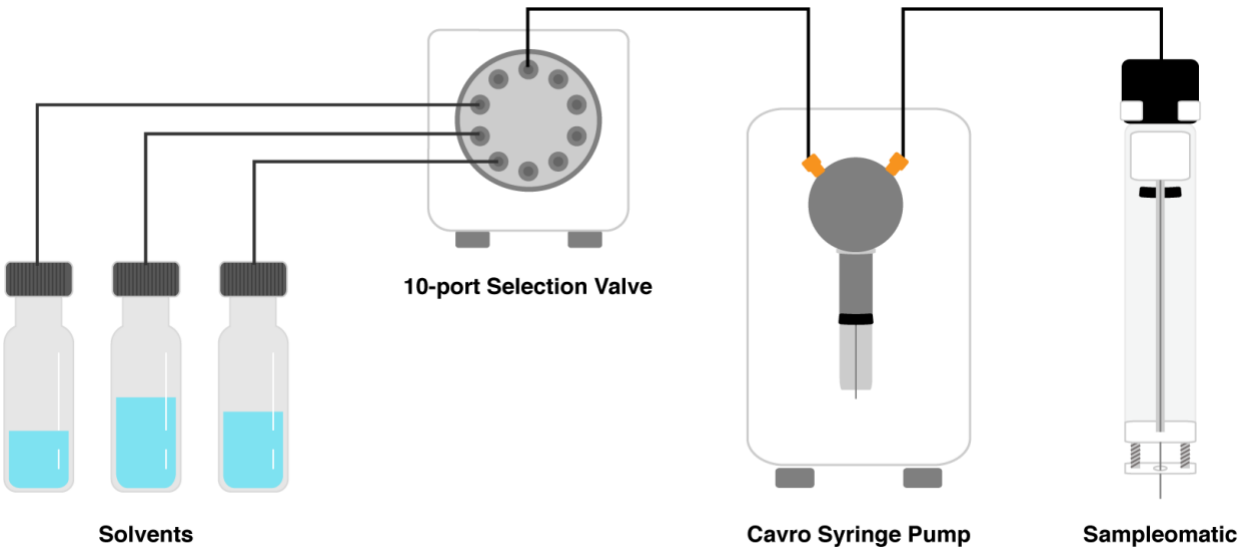


Fig. S8. Diagram of liquid handling system used in Kinova platform

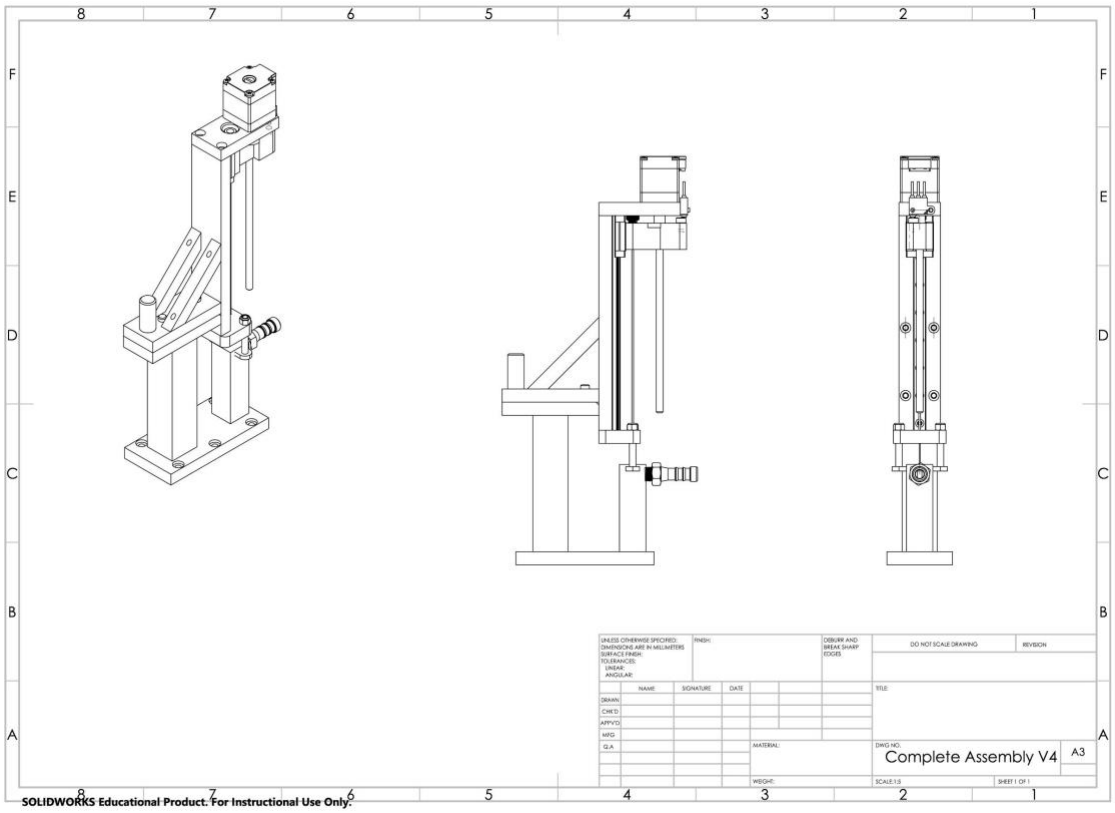
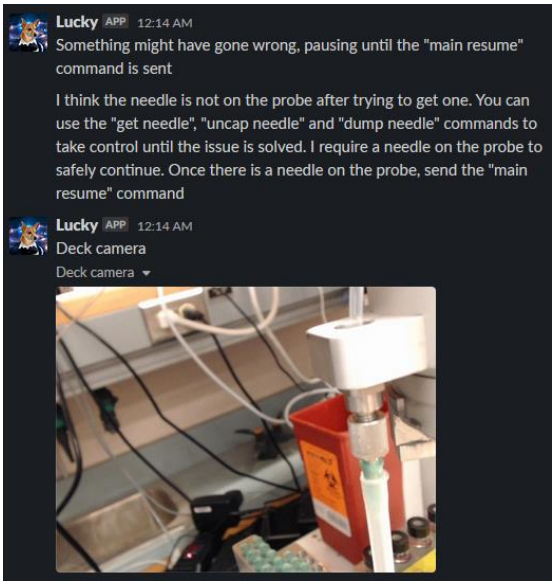


Fig. S9. Sampleomatic assembly, top and base

Uncapping vials with Kinova

As mentioned earlier, uncapping HPLC vials on the N9 platform happened with the assistance of an on-deck vial gripper. In the new setup, O-ring press fit vial locations were used to grip the vial while the gripper on the arm performed the vertical twist motion for uncapping procedure. Due to the arm speed and limitation, the uncapping was broken down into a few movements, rotating about 90 degrees at a time until capping and uncapping was achieved. A screw top was used as a parking station for the caps.

(A)



(B)

(C)

(D)

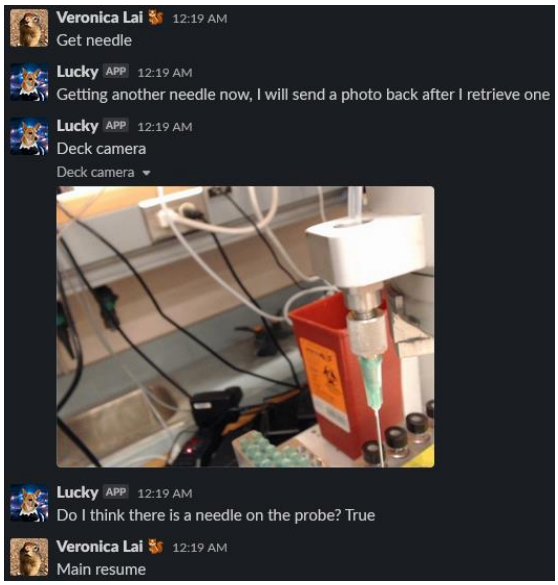


Fig. S10. Needle cap and uncapping control with vision: **(A),** after unsuccessfully getting and uncapping a needle, a message was sent by the Slack app detailing the issue and prompting the researcher to take control to remedy the issue. **(B)**, a command to uncap the needle was sent but it looks like there is an issue with trying to uncap the needle, most likely the cap was on too tight. **(C),** the command to remove the needle is sent and was executed successfully. **(D)**, the command to get and uncap a needle was sent and was executed successfully, and the command to resume the experiment was sent.

Computer vision for turbidity, challenges

There are limitations in our visual proxy measurement of turbidity. As the algorithm uses brightness as a proxy for turbidity, it works best when there is high contrast between the solution being measured and the background behind it. When the solution looks white when turbid, the contrast makes the algorithm give clear and turbid solutions distinct values. The algorithm always measures brightness in a fixed region of the image, so if the camera or the solution container shifts, the resulting measurements might not reflect the solution's turbidity. For example, we observed that if the stir bar moved in and out of the region of interest this caused lots of noise and unreliable turbidity measurements. Light glare off the glass and the general lighting environment will also contribute to the turbidity measurement, so these are factors that need to be kept constant.

Solubility data

Table S2. Overall data from the runs with the N9 platform

Solvent	Mass solid (mg)	Volume solvent (mL)	Solubility (g/mL)
acetone	8.23	0.82174463	0.0100152769
acetone	8.51	0.82667509	0.0102942499
acetone	8.14	0.80240202	0.0101445407
acetone	8.05	0.84791403	0.00949388698
water	7.29	0.4279123	0.01703669
water	8.99	0.4892356	0.01837694
water	8.35	0.444	0.01880631
water	8.60	0.4609	0.01865915
ethanol	4.71	0.85361217	0.00551773
ethanol	4.53	0.84334601	0.00537146
ethanol	4.55	0.87084918	0.00522479
ethanol	4.29	0.79632446	0.00538725
methanol	6.3	0.71163085	0.0088529
methanol	6.9	0.77079646	0.00895178

methanol	6.16	0.79620733	0.00773668
methanol	6.02	0.76207332	0.0078995
acetonitrile	9.21	0.51272265	0.01796293
acetonitrile	21.57	0.52697201	0.0215761
acetonitrile	20.34	0.110005	0.01849
acetonitrile	9.36	0.49987277	0.01872476

For solubility screening, if the measurement of what a fully dissolved solution looks like is inaccurate, all ensuing decisions on when dissolution is achieved will be inaccurate. To remedy this, the turbidity measurement that indicates the fully dissolved state can be adjusted through Slack. As we use computer vision to measure turbidity, the images the algorithm uses are the same input that we would rely on to make decisions about turbidity ourselves.

Turbidity monitoring dissolution point

When the algorithm monitors turbidity the check to determine when the solution has fully dissolved happens in multiple stages. Samples are considered dissolved if the turbidity has stabilized at the reference value or if the five most recent minutes of turbidity are all below the reference value. To account for cases where large solid particles remain but do not contribute to the solubility measurement, the stir rate is increased to strongly agitate and break down any remaining chunks. After five minutes of strong stirring the stir rate is reduced back to the normal rate and the turbidity is monitored for another five minutes. After this second round of monitoring, the dissolution check is repeated. If the sample is considered dissolved, the workflow ill proceed to the clean-up stage. If not, then turbidity monitoring is resumed.

Ambient light study

The large number of windows in our laboratory lead to significant fluctuations in ambient light. To investigate the degree to which ambient light affects the turbidity value output by the algorithm, we used the system to monitor the turbidity of water at different time intervals throughout the day (Fig. S11). Fluctuations in sunlight and researchers moving around the benchtop set-up during daylight hours result in a far more inconsistent lighting environment than at night.

Fig. S11 shows data gathered during each run along with an image of the ROI and normalization region selections. As expected, relatively turbidity varied less during the evening (top three graphs) than during the day (bottom graph). The increased variance in inconsistent lighting environments is a result of the turbidity algorithm’s reliance on brightness as a proxy for

turbidity. Measurements are thus most reliable when a consistent lighting environment can be ensured.

The difference in turbidity values between each run can be explained by the different ROI and normalization region selections for each run. Because the determination of when a solution is dissolved is based on monitoring turbidity against a reference for each individual run, the turbidity inconsistency between runs does not introduce error.

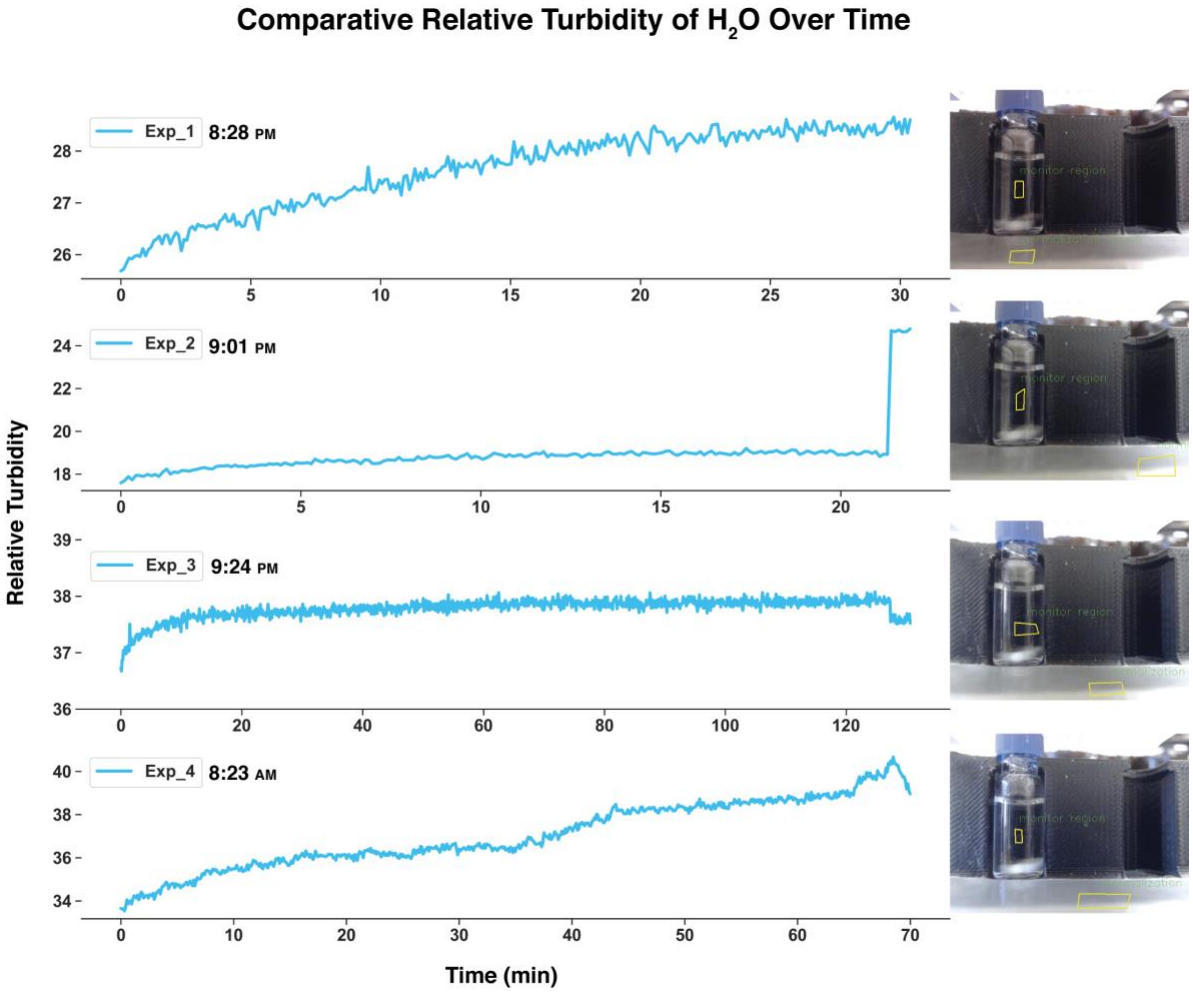


Fig. S11. Relative turbidity of H_2O at different daylight intervals. For each graph a corresponding image of the vision station and with the ROI and normalization region annotated on the image are presented. The top three graphs correspond to evening runs. The bottom graph corresponds to a run conducted during the day when fluctuations in sunlight caused fluctuations in ambient light in the room. In the second top graph, Exp_2, the sharp increase at the end is due to the bright lock screen wallpaper of the computer when it was woken from sleep and used to stop the experiment; the computer directly faced the front of the vision station. For the third graph, Exp_3, there is a slight decrease at the end, which is likely due to the lock screen wallpaper.