An Intelligent Automation Platform for Rapid Bioprocess Design

Tianyi Wu[1] and Yuhong Zhou[1]

Abstract

Bioprocess development is very labor intensive, requiring many experiments to characterize each unit operation in the process sequence to achieve product safety and process efficiency. Recent advances in microscale biochemical engineering have led to automated experimentation. A process design workflow is implemented sequentially in which (1) a liquid-handling system performs high-throughput wet lab experiments, (2) standalone analysis devices detect the data, and (3) specific software is used for data analysis and experiment design given the user's inputs. We report an intelligent automation platform that integrates these three activities to enhance the efficiency of such a workflow. A multiagent intelligent architecture has been developed incorporating agent communication to perform the tasks automatically. The key contribution of this work is the automation of data analysis and experiment design and also the ability to generate scripts to run the experiments automatically, allowing the elimination of human involvement. A first-generation prototype has been established and demonstrated through lysozyme precipitation process design. All procedures in the case study have been fully automated through an intelligent automation platform. The realization of automated data analysis and experiment design, and automated script programming for experimental procedures has the potential to increase lab productivity.

Keywords

multi-agent system, automation, rapid bioprocess design, integration

Introduction

Recent advances in microscale biochemical engineering have led to the automation of microscale bioprocess experimentation. As a result, a series of microscale experimental methods and a range of novel devices for mimicking large-scale unit operations have been developed.[1] Liquid-handling systems have been used to implement high-throughput microscale experimentation. A liquid-handling system comprises a robotic arm that dispenses a selected quantity of processing material, reagent, samples to a designated container, a set of devices that perform biochemical experiments (such as mixing, centrifugation), and a computer with control software that allows the user to customize the liquid-handling procedures and transfer volumes. The liquid-handling system is an automated system capable of doing most of the microscale experimentation automatically for both upstream bioprocesses such as *Escherichia coli* growth kinetics and recombinant enzyme expression and downstream bioprocesses such as precipitation and chromatography with a prepacked micro column attached to pipetting tips.[1–4] Several examples have been published to indicate that high-throughput technology has shortened the bioprocess development time.[5,6]

Currently, the liquid-handling system carries out the experimentation, and then the data are analyzed to understand the process. Often, 96 experiments are run in parallel by using a 96-well plate and the statistical design of experiment (DoE) method to screen a wide design space. Experiments use material at the microliter volume so it is challenging to achieve the precision of operation at the microscale scale compared with the lab scale. Extra replicates of the experiments may be required to ensure the accuracy of the experiments at the microscale. Therefore, many hundreds of experiments are designed and executed in parallel on such advanced platforms. In this content, experiment design refers to select experimental conditions or experimental choice, and an assay refers to an analytic procedure for quantitatively measuring the amount of a biological entity.

[1]Department of Biochemical Engineering, University College London, London, UK

Received March 20, 2013.

Corresponding Author:
Yuhong Zhou, Department of Biochemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK.
Email: y.zhou@ucl.ac.uk

The existing workflow has two major bottlenecks. First, the development of high-throughput assays does not match the rapid pace of experiment execution. For example, from our previous experience in monoclonal antibody (MAb) precipitation experiments, it took about 30 min to run 96 experiments in a microwell plate but required more than 30 h to detect MAb concentrations and total protein concentrations using high-performance liquid chromatography (HPLC) for those 96 experiments. A large number of parameters for the product and the impurities such as nucleic acid concentration and host cell protein concentration need to be measured to characterize the complex biological molecules. However, most biological assays need sample preparations that are time-consuming. Although advanced HPLC systems of ultra pressure type or low volume type have reduced the assay time, the sequential fashion of biological assays has led to a mismatch between time of experiment execution and time of assay in bioprocess studies.

High-throughput assay development is still in its early stage, and only a few specific high-throughput assays have been reported. The assay of quantitative protein analysis can be sped up by using high-throughput methods such as ultraviolet-visible (UV) spectrophotometry.[7,8] However, these methods are largely based on either mathematical or statistical models, so data treatment of the large data set requires the handling of large amounts of data files manually, which takes a long time to complete.

The second bottleneck is data analysis and experiment design. The liquid handler is capable of carrying out 96 or 384 experiments in parallel by benefiting from microscale experimentation technology. After experimentation and assay, data analysis is the key step in bioprocess design. There are two types of data analysis tasks:

1. To convert raw data from the measurement instruments into biological parameters. For example, the assay of quantitative protein analysis using UV spectrophotometry needs to translate spectrophotometry data into protein concentrations using calibration models.[7,8] The software applications in most instruments include data analysis modules for data treatment. For instance, the algorithm in ChemStation (Agilent Technologies, California, USA), which supports the Agile HPLC instrument, can intelligently recognize the peaks and calculate the areas.
2. To understand the process performance based on the data and, more important, to design the new experiments for process development. Here we use experiment design to refer to determining the experimental conditions, not the experimental procedure. More sophisticated data analysis software applications capable of the design and analysis of experiments such as Design-Expert from Stat-Ease Software (Minneapolis, MN) are, however, stand-alone software and not integrated with automated instrumentations. When the data set is large, researchers often take substantial time for this data analysis. The new experiment design from this data analysis requires the user's input (e.g., which area should be explored, which variables should be studied, and what the limits of variables are). Therefore, this data analysis step becomes a potential bottleneck in bioprocess development.

In a typical precipitation process design, researchers first need to identify the feasible design space according to their knowledge or previous experimental results. Second, they input the information so an optimized experiment design can be produced by computer software. Third, they need to program a command script to control the automated devices performing tasks such as adding buffer, shaking, incubation, and filtration. After an experiment, the researchers have to use software to control assay equipment (e.g., plate reader, HPLC) to do the assays. Finally, they transfer the raw data into analysis software to carry out data treatment and analysis so as to determine the next batch of experiments by various computational methods, if needed, or identify the process solutions.

Hence, an automation platform that integrates the manipulation of experiments and assay instrumentation, as well as automates the data analysis, experiment design, script programming, and raw data treatment, could cut out the time of human involvement and enhance the efficiency of the current practice. The key function of the automated platform should be the capability of making decisions on whether further experiments are needed and, if needed, how to generate the next round of automated experiment designs.

There are many intelligent machine learning methods available that learn from data. In contrast to many data mining problems, ours is how to achieve the process design solution with a minimum number of experiments. Two common sequential experiment design methods are search heuristic methods and model-based methods. They fit well for our needs where only limited experimental data are available for learning. We described these methods briefly below.

Search Heuristic Methods. Search Heuristic Methods are a direct process optimization. Instead of using a mathematical model to describe the process, these types of methods explore the response surface by using data from previous experiments to determine the next best experimental conditions for further study.[9–11] This is in contrast to traditional statistical experiment design methods that define a large set of experiments at the outset to generate the response surface. The search heuristic methods are attractive at the early process development stage because only a small number of experiments need to be carried out at every step, and usually the solution can be found after a reasonable number of iterations.[12,13]

Model-Based Methods. Model-based methods use mathematical models to aid the process understanding and process design.[14] Experiments generate data, which are used to estimate the values of the unknown parameters in the process model. The model with these estimated parameters is then used for simulation and optimization.[15] Starting with an initial experiment design, the strategy is to sequentially revise the parameter estimates according to data generated from the previous experiments until the estimates are within a specified accuracy.[16–18]

Although these sequential methods have been established, the steps in these methods are reported to be taken manually, leading to delay.[15,17,18] However, both the above methods have the potential to be implemented in an automated fashion.

A robotic scientist called "Adam" has been developed to identify genes encoding orphan enzymes in *Saccharomyces cerevisiae* that has realized the integration of experiment execution and data analysis.[19] "Adam" uses an intelligent system to automatically generate and validate hypotheses in a closed-loop learning manner. The prototype starts with a functional genomics hypothesis based on a system model and a knowledge base. Then a cycle of automated hypotheses generation and experimentation is realized by a combination of statistical machine-learning methods and an integrated robotic system of a liquid handler, an incubator, and a filter. The hypotheses-driven closed-loop learning framework of "Adam" is very similar to the concept of sequential experiment design methods. Furthermore, "Adam" combines a robotic system into sequential methods to make the system overcome the mismatch of data analysis and execution. "Adam" has demonstrated that automation of laboratory equipment can revolutionize laboratory practice and give rise to significantly increased productivity.

The aims of bioprocess development are often defined based on product quality and economic requirements so the objective functions are quantified, and thus the closed-loop integration concept can be realized for bioprocess development. To assist achieving rapidly the process design targets, we propose an intelligent automation platform to enhance the speed of process development. It integrates experimentation, data analysis, and experiment design to perform automated closed-loop learning to achieve the design objectives without human intervention and avoid unnecessary delay.[19] A general framework of the proposed intelligent automation platform is shown in **Figure 1**. It aims to bring all the elements that are performed manually into an automated platform. This includes the physical devices that carry out experiment preparation, experiment execution, and assays and the five types of software needed in this platform. These are (1) software to drive each device timely; (2) software to consolidate the data for each device due to their different formats; (3) software of the sequential design algorithms that realize the process design objectives and are
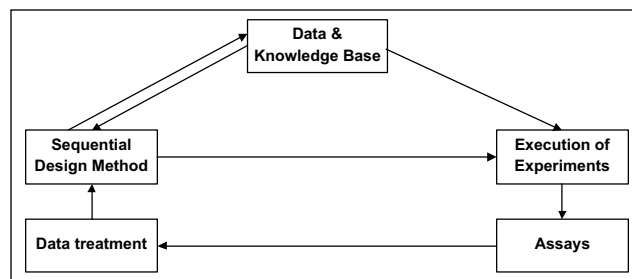
Figure 1. Intelligent automation platform to realize the closed-loop learning.

capable of process evaluation, prediction, and optimization; (4) a database used for storing the experimental data and information on instrument configurations; and (5) software to control and manage the above software. The bottleneck on high-throughput assays is addressed by others.[7,8]

Starting with a set of initial experiments from the user or database, the automation of the platform of the closed-loop learning approach will drive the automation device to perform defined initial experiments and drive the analytical devices to measure the data after the experiments are completed. Then the system will pick up the data from the analytical devices and perform data analysis that turns the raw data into measurement results, such as protein concentrations, and carry out process evaluation based on the optimization objective to achieve the design solution. After evaluation of the design solution, the platform will decide to stop or continue to design the next round of experiments for optimization.

Apart from the experiment execution, there are three stages that need human involvement in the current experiment procedure: (1) program the experimental procedure and initiate the liquid-handling system, (2) data acquisition and treatment from the analysis devices, and (3) the sequential experiment design. Due to the complexity of the dynamic automation platform and the demand of closed-loop learning capability, a sophisticated architecture is required to accommodate the integration of the independent components.[20] There are two types of architectures: distributed and nondistributed. Hierarchy architecture is a typical example of the nondistributed case. In hierarchy architecture, components are linked in a tree format. Commands, tasks, and goals to be achieved flow down the tree from superior components to subordinate components, whereas results flow up the tree from subordinate to superior components. Thus, when adding a new component, the architecture needs to be modified, not only the parts related to the new component but also at each level. Therefore, the entire system needs to be changed every time when new components are introduced. In addition, because the automation devices are controlled by different software from different manufacturers, it is extremely hard to gain direct control of

the devices if an Application Programming Interface (API) is not provided. It will require an agent to monitor the software associated with automation devices and collect data when the task is completed. On the other hand, a distributed architecture has more flexibility when a new component is added. The agent-based architecture has attracted attention for its flexibility. Within the agent-based architecture, an agent is defined as an autonomous unit capable of executing actions to meet specific objectives.[21] The agents are able to cooperate to accomplish a complex task in a distributed way by the capability of interaction and communication. The agent-based architecture technique has shown its advantages in several diverse engineering areas, including problem solving, simulation, and collective robotic control.[22] As for bioprocess development, an agent-based approach has been applied to perform optimization based on models of whole-production process simulation.[23]

Compared with other integration approaches, an agent-based architecture has a number of advantages satisfying the requirements of the intelligent automation platform. The modular approach can be adopted for biopharmaceutical processing using agent-based techniques because it is easier and less costly for system software modification when adding or deleting components.[23,24] The agent-based architecture can divide complex tasks into small and manageable subproblems.[20] The agents can collaborate to solve the presented problem by communication and interaction with each other, allowing the simultaneous facilitation of information exchange. Based on these features, the agent-based architecture provides an ideal approach to integrate all the components involved in the closed-loop learning. The experiment design algorithm, experiment execution, and assay can each be regarded as an autonomous agent to perform defined tasks. The flexibility of the architecture makes it easy to add new algorithms and devices into the platform.

An efficient interaction protocol should be able to support several activities such as the communication of global goals, sharing of knowledge, and avoidance of conflicts.[25] The most widely applied protocols are contract net,[26] market mechanism,[27] and blackboard.[28] The first two are designed to simulate real market activities on the assumption that a global optimum exists for the equilibrium of the market such that the "seller" wants to maximize profit while the "buyer" wants to maximize utility. The agents in these protocols therefore have similar functionalities such as "buy" and "sell." When each agent is capable of a different function, the architecture based on the "blackboard" mechanism is attractive. The blackboard is a database that all relevant agents can access. These agents perform tasks based on the information on the blackboard. The agents publish the results on the blackboard continuously until the goal has been achieved.

In this article, we present the prototype that brings together the currently independent parts: the automated experiment design, the experiment execution by a liquid-handling system, and the assay analysis for protein precipitation studies. In next section, the lysozyme protein precipitation studies, the physical setup of the intelligent automation platform, and software for each type are described. The results on multiagent system software development that controls the whole bioprocess development activities and a protein precipitation case study are then given, followed by a discussion and conclusions.

Materials and Methods

Lysozyme Protein Precipitation Studies

For demonstration purposes, we investigated the solubility of lysozyme to identify the optimal ion strength and pH conditions in precipitation process design. The goal of the experiments was to find an optimal design that maximizes the yield of lysozyme and maximizes the ammonium sulfate concentration. Here maximizing the salt concentration reflects maximizing purity for a crude industrial feed. The objective function, J, is shown in equation (1).
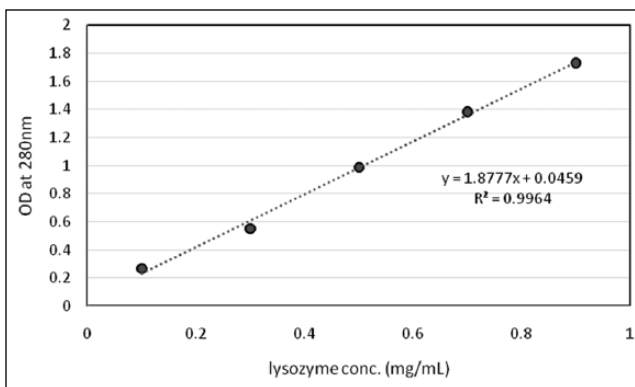
$$J = \alpha(c/c_0) + \beta(s/s_0), \qquad (1)$$

where α and β are weighting factors, c_0 the initial lysozyme concentration, c the final lysozyme concentration in the supernatant, s_0 the maximum ammonium sulfate concentration, and s the ammonium sulfate concentration used in the specific experiment.

For our experiments, the ammonium sulfate, monosodium phosphate, and disodium phosphate were obtained from Sigma-Aldrich. (Munich, Germany) The ammonium sulfate solution was prepared at a concentration of 4 mol/L. The phosphate buffer was chosen to maintain the desired pH value. The monosodium phosphate and disodium phosphate solutions were prepared at the concentration of 300 mM and then diluted. The lysozyme was obtained from Sigma-Aldrich and prepared at 20 mg/mL. The 96-well Multiscreen filter plates were sourced from Millipore (Hertfordshire, UK) and the 96-well flat-bottom Costar UV Microwell plates from Corning (Leicestershire, UK).

The experiment involved adding 100 μL of lysozyme solution and 50 μL of buffer solution (a mixture of monosodium phosphate and disodium phosphate) into the microwell. The total volume was 300 μL by adding the 150-μL mixture of water and the ammonium sulfate solution. Thus, the initial lysozyme concentration in each well was 6.67 mg/mL. The ammonium sulfate concentration ranged from 0 to 2 mol/L. The range of pH was from 6 to 8. The interval of ammonium sulfate concentration was 0.1 mol/L, and the interval of pH was 0.1.

All the liquid-handling tasks were executed by a TECAN Freedom 200 liquid-handling system (TECAN, Männedorf,
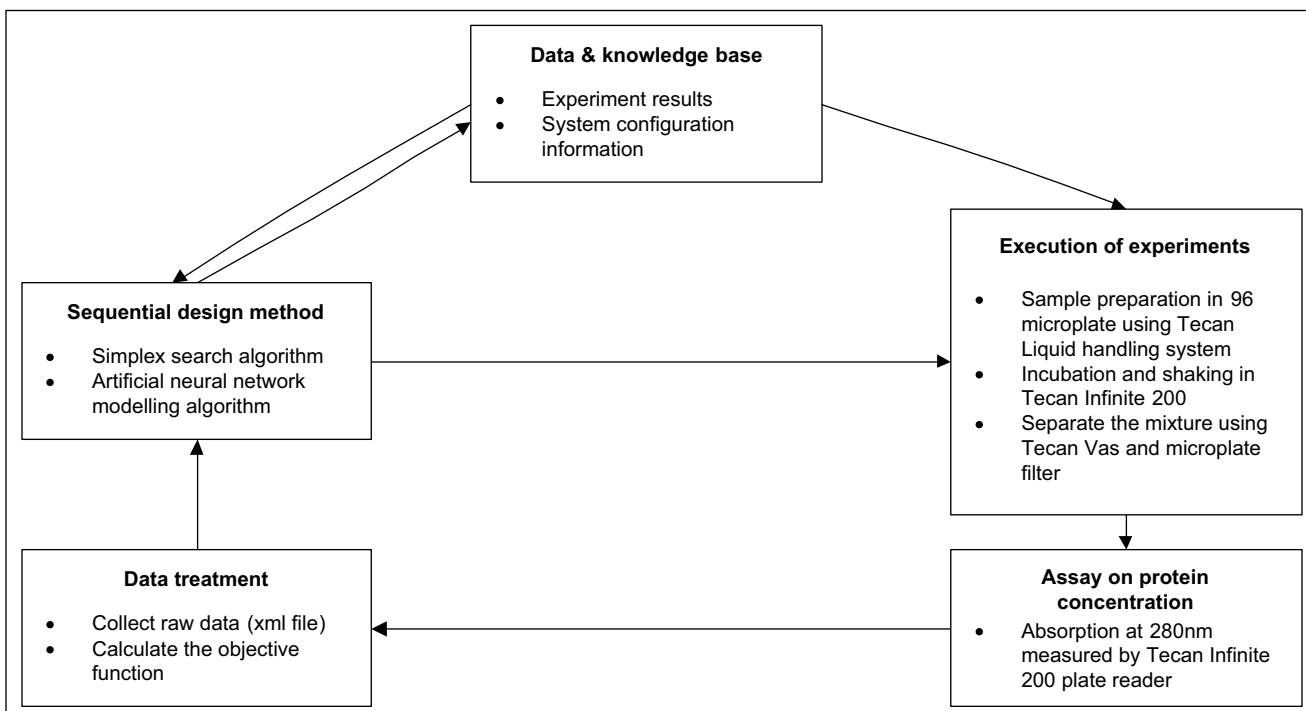
Figure 2. Calibration curve of the absorption at 280 nm and lysozyme concentration.

The simplex search method and the artificial neural network (ANN) model-based method were chosen as the sequential experiment design methods in this study. The simplex algorithm, first proposed in 1965,[11] was designed for minimizing an objective function in a multidimensional space. The technique has been applied to solve the bioprocess optimization problem[12] and is very powerful because only the experimental data are needed to calculate the objective function values. ANN is a computer algorithm inspired by how the human brain processes information and represents a promising modeling technique for data sets having nonlinear relationships. ANN has been used to analyze biopharmaceutical problems,[29] for example, in the solubility of lysozyme.[30] In this case study, a three-layer back-propagation network was used to model the relationship between ammonium sulfate concentration, pH, and lysozyme solubility.

Intelligent Automation Platform for Protein Precipitation Experiments

Based on the framework shown in **Figure 1**, the prototype for lysozyme precipitation process development has been established and is illustrated in **Figure 3**. The data and knowledge base initially stores the system configuration information of the devices (e.g., TECAN configuration). The configuration information in the database consists of both connection parameters to realize communication among devices and the dimension and layout information to guide the movement of the robotic arms during the

Switzerland). Samples were first prepared in a 96 microplate by adding different solutions before shaking. After shaking, samples were transferred to a 96 microfilter plate for filtration with a microplate beneath for filtrate collection. The filtered samples were then diluted 10-fold using two wells in the same microplate before transferring to a UV-transparent plate for measurement. Absorption at 280 nm was measured by a TECAN Infinite 200 plate reader to determine the lysozyme concentration. The calibration curve shown in **Figure 2** confirmed that the absorption at 280 nm was good for lysozyme measurement. The filtration operation was performed by a vacuum separator from the TECAN VacS series.

Figure 3. Prototype of the intelligent automation platform for precipitation experiments.

execution of the experiments. Any future experiments designed later can be translated into executable programs to drive the devices. The experimental results are stored in the database. The database uses an entity-relationship (ER) model in Microsoft Access (Microsoft Corp., Redmond, WA).

The experiments carried out included sample preparation in a 96 micro filter plate using the robotic arm, incubation and shaking in the plate reader, and liquid and solid separation of the samples by a vacuum pump sucking out the liquid from the bottom of the micro filter plate (**Fig. 4**). All of these have been carried out by the TECAN liquid-handling system and driven by designated experimental procedure software.

A plate reader measuring protein concentration was located next to the TECAN and within the reach of the robotic arm. The lysozyme optical density was measured at 280 nm UV absorption and the data stored in an XML file. The assay agent converted the raw data into lysozyme concentration based on calibration curves and stored them in the database.

The sequential design method uses the experimental data to calculate the objective function and decide if it is necessary for another round of experiments; it uses the simplex search algorithm or the artificial neural network modeling algorithm to design any next round of experiments required. The experiment results and the newly designed experimental conditions are stored in the database for further use.

A multiagent architecture has been developed to control all of the software to realize the closed-loop learning. The whole software is programmed in C#, and the control of the TECAN liquid-handling system and other devices is realized through the API provided by EVOware and i-Control from TECAN. The ANN is programmed based on NeuronDotNet Library. Sample source code for these is provided in the supplementary documentation.

Results

Multiagent Architecture for the Intelligent Automated Platform

In our platform, each agent is capable of a different function, and thus the architecture has been established based on the blackboard mechanism, as shown in **Figure 5**. Hence, as an example, the architecture for the lysozyme precipitation process development has been established, as shown in **Figure 6**.

The coordinate agent retrieves the historical experimental data from an ER database and writes the new executed experimental results into the same database. It also translates the information in the database into the message format for the blackboard. The blackboard is the place to publish two types of information: configuration data (e.g., platform layout, database connection parameters) and

experimental data. The data of each experiment are published in the format similar to that stored in the ER database. The message format in any experiment has a tree structure where the parameters are defined. It includes experiment number and attributes of experimental settings (e.g., parameters and assays) and results. An example of a message on the blackboard for a lysozyme precipitation experiment is shown in **Figure 7**. The coordinate agent selects all relevant experimental data from both the database and other agents and displays them on the blackboard. Other agents can access the blackboard to update specific fields (e.g., experimental result) when they are available.

The execution agent controls the robotic arm to add the lysozyme solution, phosphate buffer, and ammonium sulfate solution into the microwell plate, with the different pH values of each experiment achieved by adding calculated volumes of monosodium phosphate and disodium phosphate solution to the microwell plate. Then the robotic arm puts the microwell plate into the incubator, and the execution agent controls the shaking speed and time. After incubation and shaking, the execution agent controls a vacuum pump on the platform to separate the precipitate and liquid. Finally, it controls the robotic arm to take the samples of the permeate to the plate reader for detection of the lysozyme concentrations.

The assay agent controls the plate reader. The coordinate agent communicates with the assay agent when the experiments have completed. Then the assay agent initiates the plate reader for the detection of protein concentrations. Based on the calibration results, the agent reads data from the plate reader and calculates the lysozyme concentration of each well. Finally, it posts the lysozyme concentrations data back to the blackboard.

The experiment design agent comprises two algorithms that are a simplex search algorithm and an ANN. Based on these algorithms, the agent publishes the newly designed experiments on the blackboard for execution.

Starting with three random initial points in the design space of pH and ammonium sulfate concentration, the simplex algorithm decides the next needed experimental point based on the evaluation of the objective function at the vertices of the simplex. The flowchart in **Figure 8A** shows how the experiment design agent uses the simplex search algorithm to design the next round of experiments and the data communication. The experiment design agent picks up the experimental data from the blackboard and then computes the objective function values at the vertices of the simplex. It then selects the next round of experiments by reflection, extension, or contraction of the vertex with the lowest objective function value relative to the other vertices. Finally, the experiment design agent posts back the newly designed experimental conditions unless they correspond to an experiment that has been previously performed. The flowchart in **Figure 8B** illustrates how the experiment
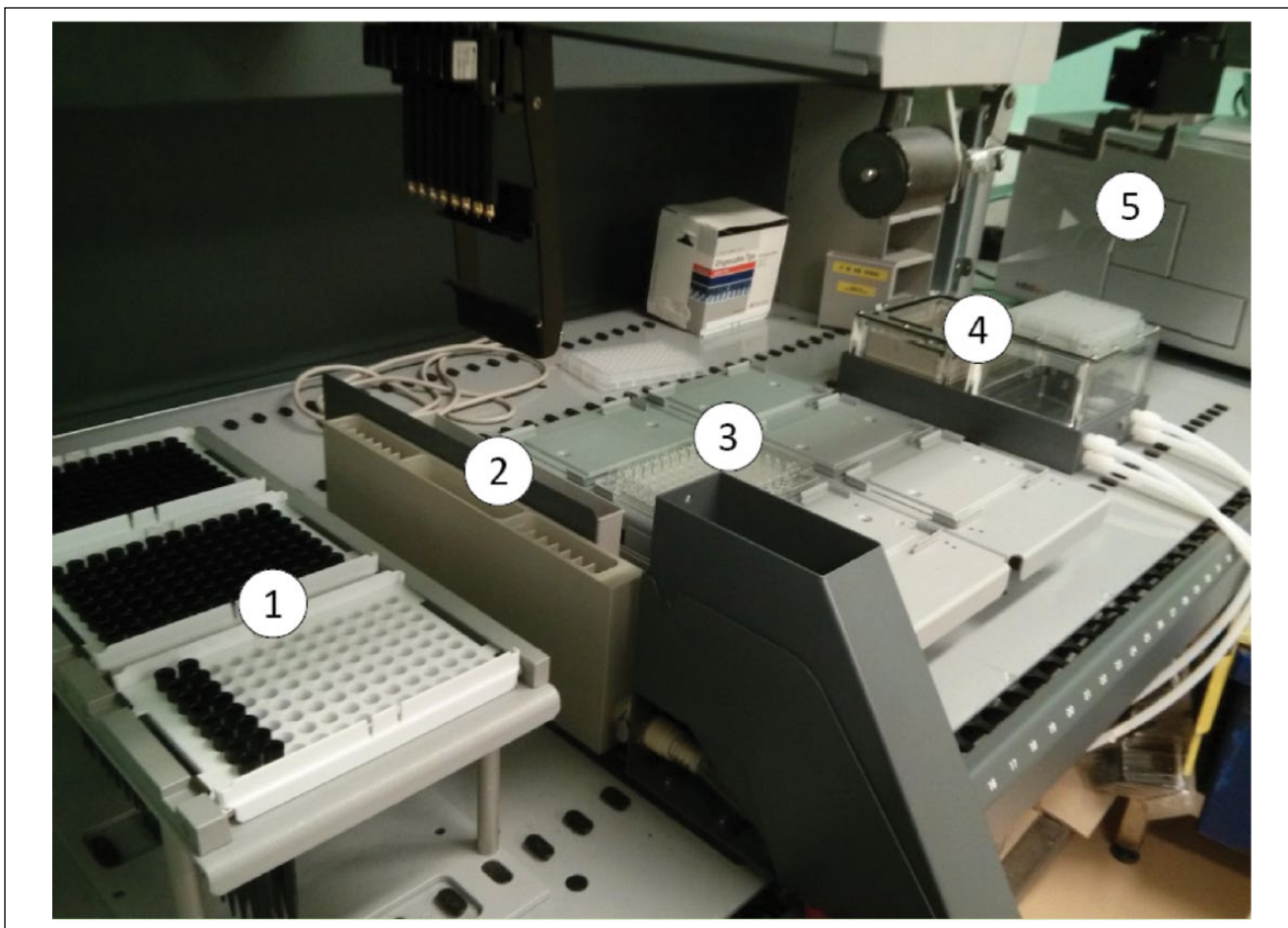
Figure 4. Hardware configuration of the intelligent automation platform for precipitation. (1) Disposable tips rack; (2) tube carrier; (3) 96-well microplates carrier; (4) vacuum filtration; (5) UV plate reader.

design agent carries out the ANN experiment design method. The experiment design agent picks up the previous experimental results from the blackboard and then uses them to train the ANN model. Instead of modeling from a data set with a given size, the experiment design agent trains and tests the network after every five experiments until the desired accuracy is achieved. If the required accuracy is not achieved, it generates a new batch of five experiments and posts them back to the blackboard.

Optimization of Precipitation

For our given process, we have obtained the following results using our described methods. The TECAN liquid-handling system and the plate reader were driven by the execution agent and assay agent automatically. The communication between the coordinator agent and other agents was smooth in facilitating data exchange. Two lysozyme precipitation studies were carried out. The experiment design agent used the simplex search algorithm in the first study and the ANN algorithm in the second study.

Simplex algorithm results. The initial simplex was selected randomly to mimic a novel process design task when no prior knowledge was available. The sequential experiment design agent then designed 16 new experiments based on the simplex algorithm. The optimal point was found after 16 iterations. The maximum value of the objective function was 0.70, where the ammonium sulfate concentration was 0.8 mol/L and the pH was 7.1. The corresponding lysozyme concentration was 6.05 mg/mL. As shown in **Figure 9A**, the objective value jumped from 0.14 to 0.50 after only the first set of iterations, which proves that the simplex algorithm is very effective when searching for the extreme point in the given space. When the ammonium sulfate concentration was lower than 0.8 mol/L, the solubility of lysozyme stayed almost the same so that the objective function value decreased. When the ammonium sulfate concentration was higher than 0.8 mol/L, the rapid decrease of the solubility of lysozyme made the objective function decrease; also, the impact of pH on the solubility was less than that of the ammonium sulfate concentration. **Figure 9B** shows the response surface extrapolated from these 18 data points. It
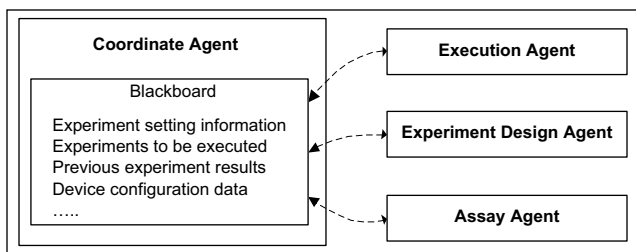
Figure 5. Multiagent architecture for the intelligent automation platform.

confirms that the optimum point found reached a relatively high objective function value.

ANN results. The initial five experiments were selected randomly. After four iterations, the model achieved the desired accuracy. The response surface of predicted lysozyme concentration in the soluble phase that was achieved is shown in **Figure 10A**. When the ammonium sulfate concentration was higher than 0.8 mol/L, the lysozyme concentration decreased rapidly, similar to the results from the simplex algorithm. At same time, pH had little impact on the solubility of lysozyme. According to the model generated by the network, the highest objective function value was found at 0.7 mol/L of ammonium sulfate concentration and pH of 7.0, as shown in **Figure 10B**.

The results from two precipitation studies demonstrated the capability of experiment design agents. Both of the sequential experiment design methods delivered the process design solution successfully. At each iteration, the communication was also very efficient since the time used for communication, data analysis, and experiment design in the above studies was less than 1 min, which significantly saved the process design time.
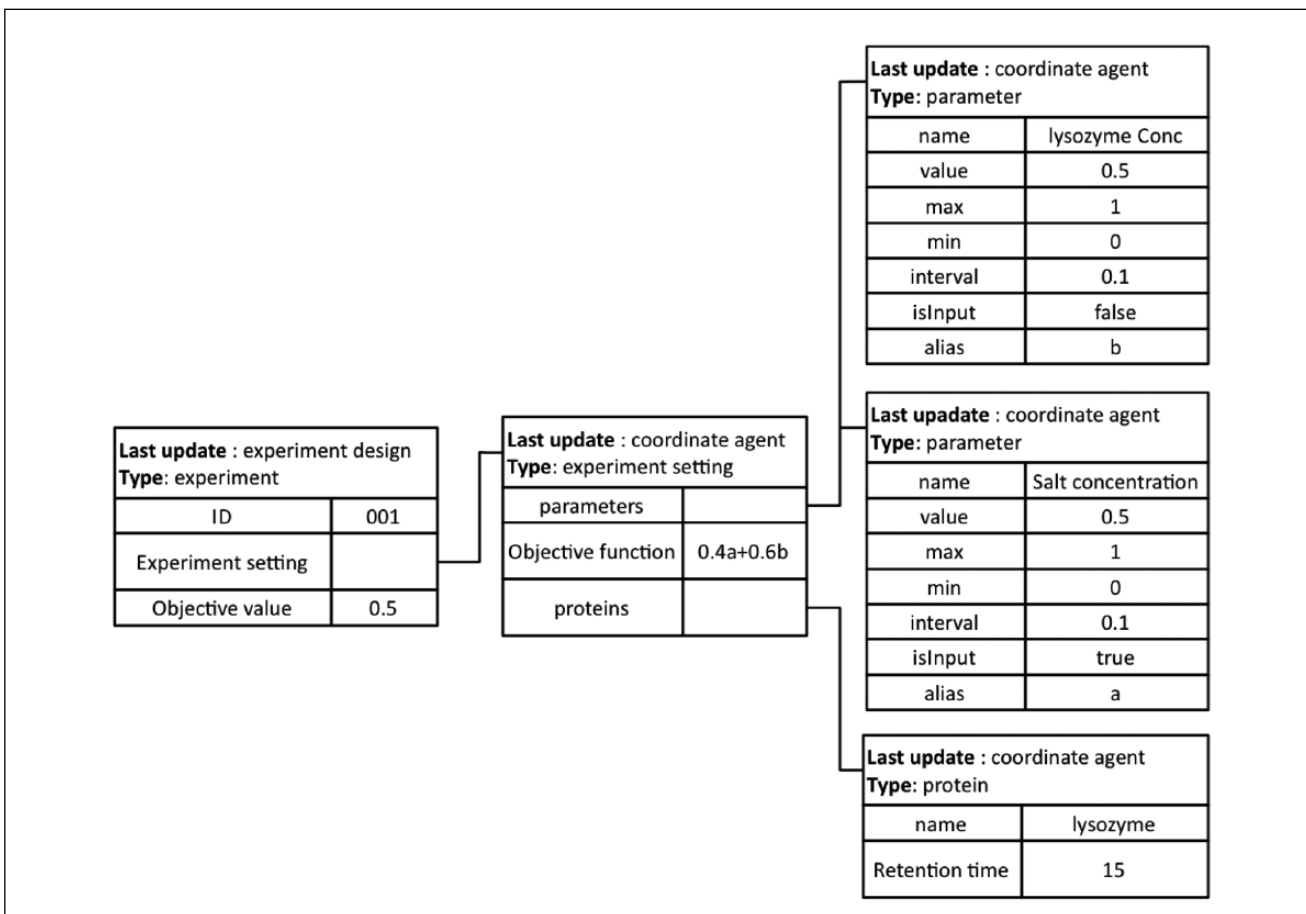
Discussion

This case study was chosen to illustrate the functions of an intelligent automated platform. Although a lysozyme precipitation experiment is very simple, it has several operations such
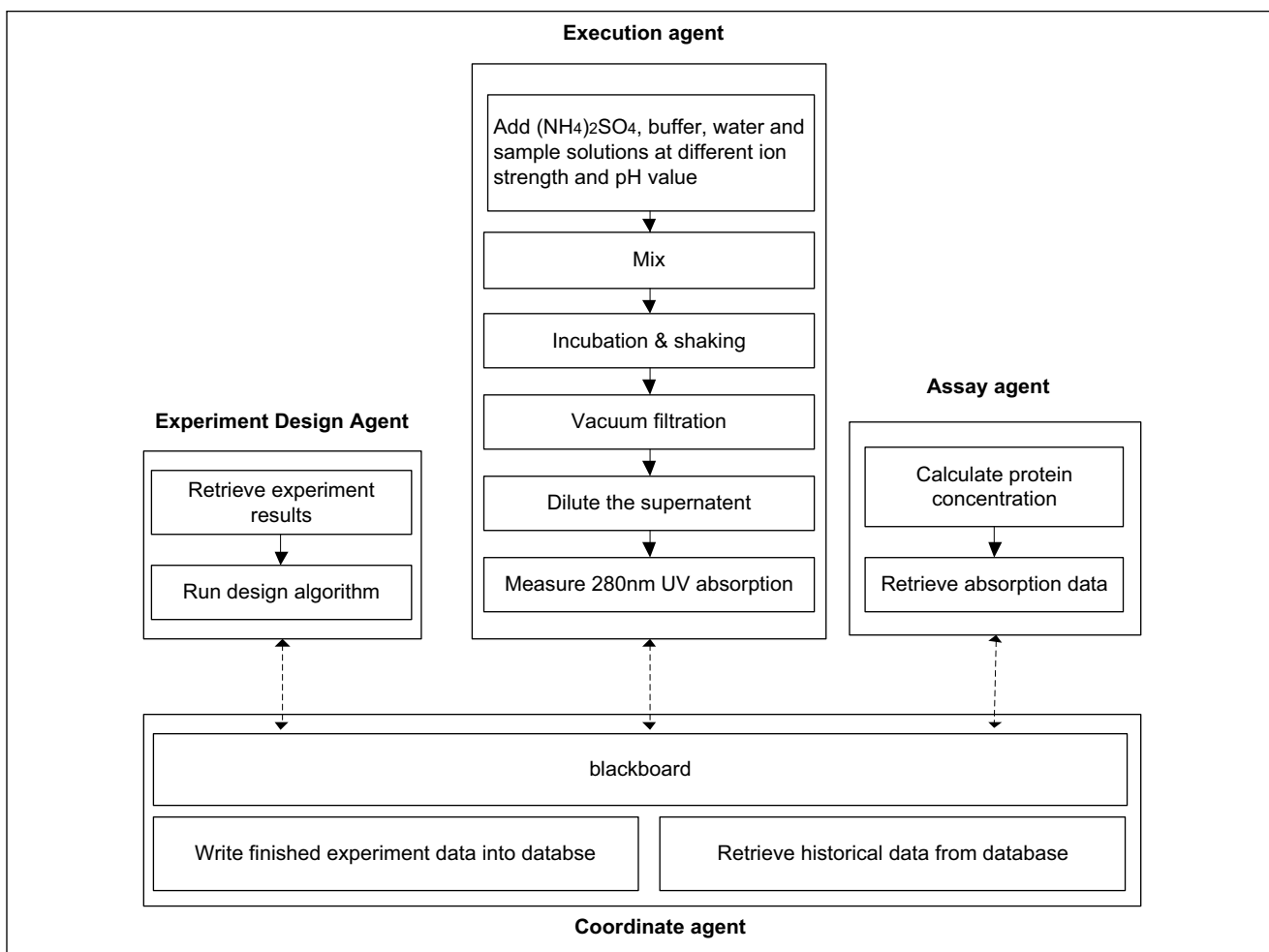
Figure 6. Multiagent architecture for the lysozyme precipitation process development.

Figure 7. An example of messages on the blackboard showing an experiment published by an experiment design agent. The experiment's objective function is $0.4a + 0.6b$, where a is the salt concentration and b is the lysozyme concentration. The protein involved in the experiment is lysozyme, and its peak appeared at 15 min when using high-performance liquid chromatography to measure the concentration. The parameters investigated in the experiment are salt concentration and lysozyme concentration. The parameters consist of design space information (maximum and minimum values and interval) and the actual value.

as mixing, incubation, and filtration in sample preparation in common with most experiments. Also, the measurement of lysozyme concentration can be easily realized by a plate reader. The objective function in this case study illustrates the typical bioprocess design trade-off between yield and purity in a complex multicomponent precipitation process.

The concept of an intelligent automation platform to realize the closed-loop learning is general and can be applied to each bioprocess design provided microscale experimentation can be carried out on a robotic platform (e.g., microscale fermentation, chromatography, centrifugation). The prototype developed for precipitation, when applied to other bioprocess design, stays the same apart from the experiment details, but naturally the experiment details may be different.

While the "Adam" system is a specialized system to generate research hypothesis automatically, our system is to address a much more fully defined bioprocess design problem, as well as aim to increase lab automation and reduce the process design time by using intelligent communication and intelligent experiment design methods. Our system is based on a commercially available liquid handler and other common lab instrumentations; hence, our method is more practical and can be applied easily in the research and development labs.

Platform Architecture

The level of interactions among the agents in this case study was relatively low. However, the required interactions will be much higher in the whole bioprocess sequence design when more than 10 unit operations are involved and each unit operation has a high number of design variables. The adopted multiagent architecture has the flexibility to allow adding or modifying elements in the system (e.g., adding a new data analysis instrument or a new experiment design algorithm). This allows the further development of agent negotiation in resource allocation, activity scheduling, and reaching agreement during process evaluation to accommodate the needs in the design of a whole bioprocess sequence. The advantage of
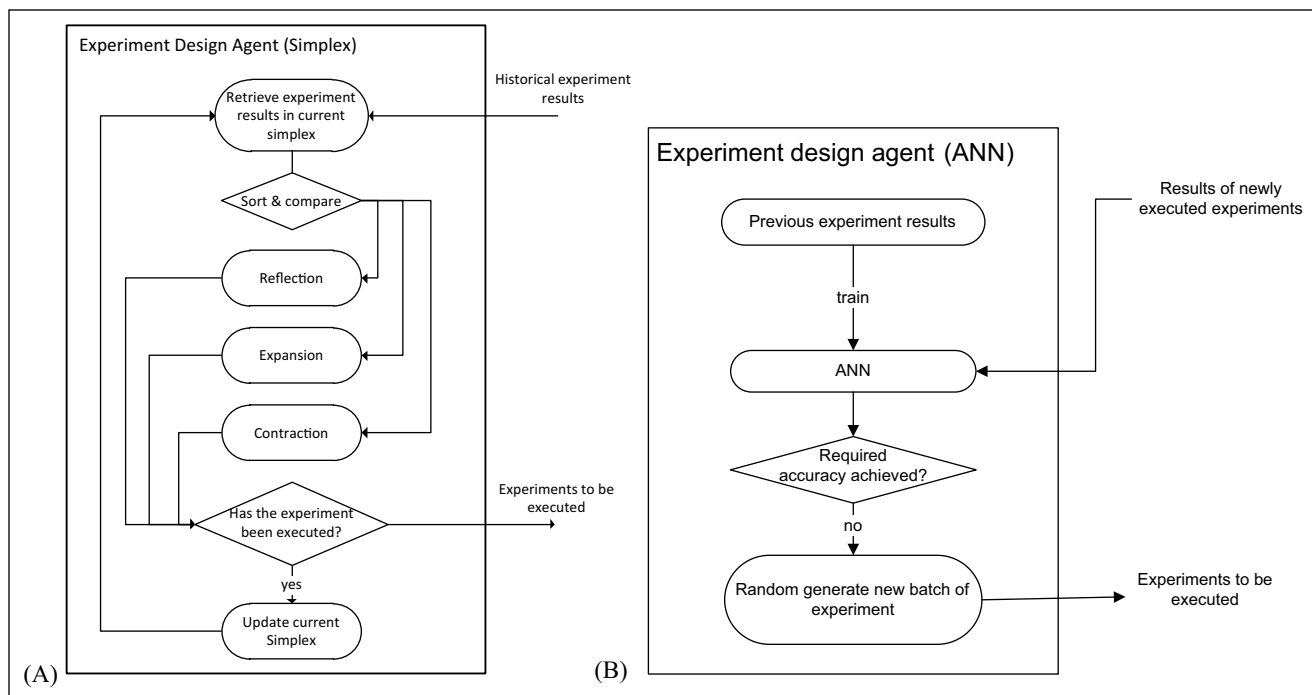
Figure 8. (**A**) The experiment design agent to facilitate the simplex search algorithm. (**B**) The experiment design agent to facilitate the artificial neural network (ANN) algorithm.
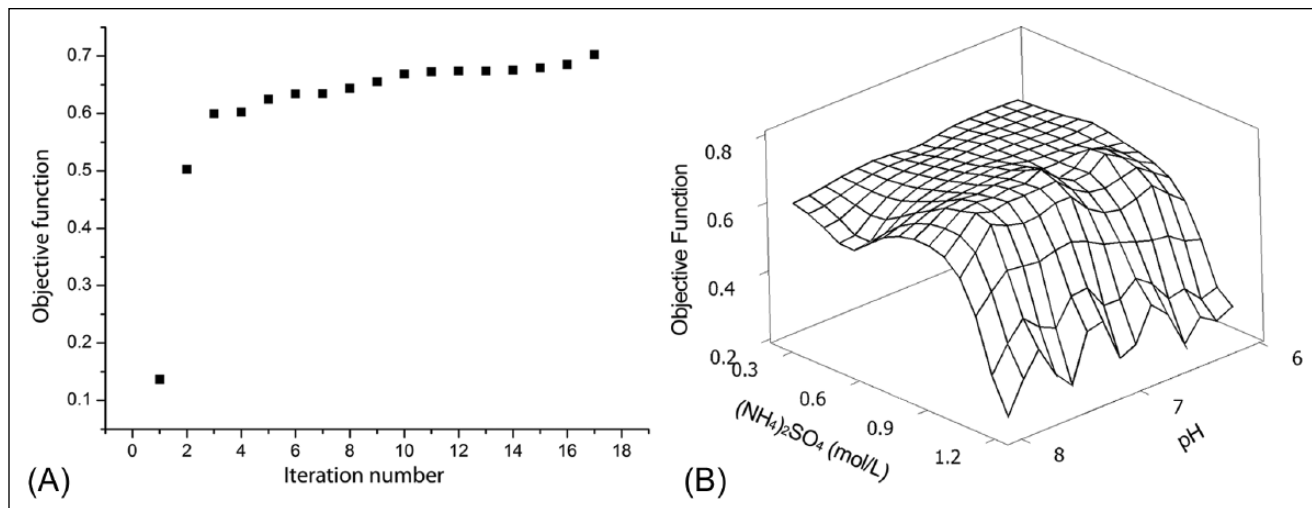
Figure 9. (**A**) The objective function increases with the number of iterations in the simplex search algorithm. (**B**) The estimated surface of the objective function based on all 18 experimental points.

the multiagent system in the integration solutions is that the agents are designed to achieve a common goal, although they have a certain level of autonomy. This may require the agents to have more intelligence—for example, the experiment execution agent currently only implements the experiments following predefined procedures, and more intelligence may be required to plan multiple procedures dynamically and/or enable parallel execution of experiments to shorten the total experimentation time.

The measurement of lysozyme concentration can be realized by a plate reader. The assay in this case is simple—only the plate reading. We designed the assay agent to be responsible for the sample assay. The assay agent not only carries out measurements but also has a capability of designing experiments for calibration (e.g., calibration curve regression for 280-nm absorbance). In addition, it is responsible for data treatment (e.g., translate UV absorbance into protein concentration). When multiple measurements are required in
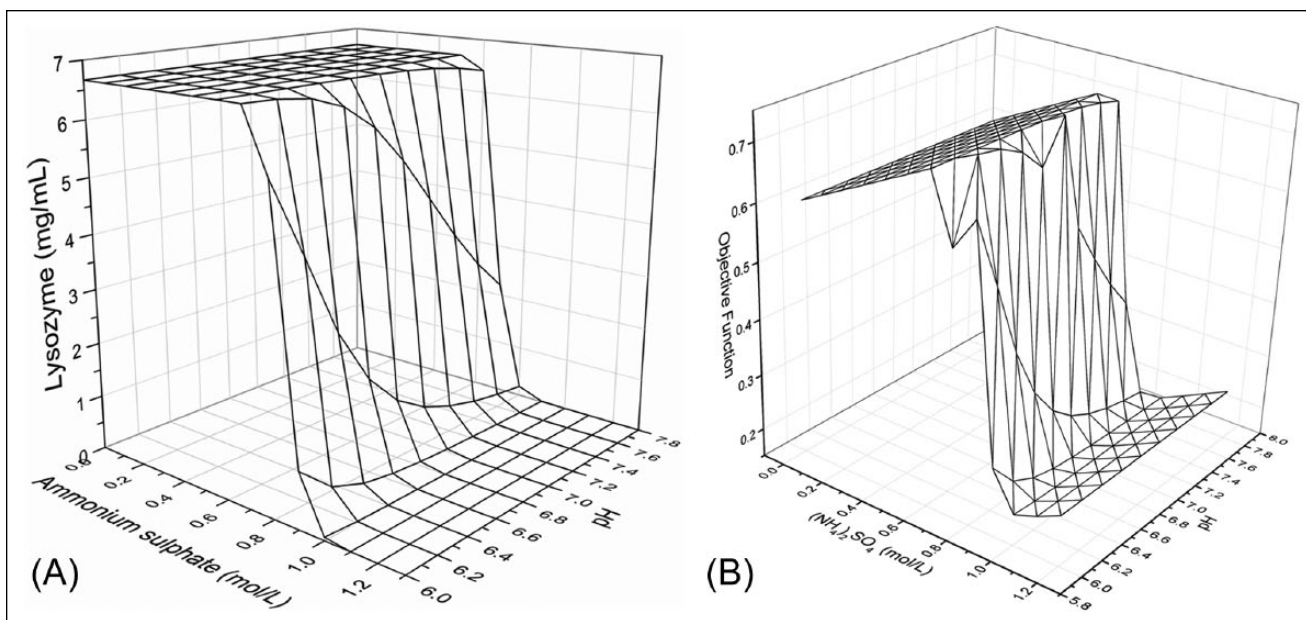
Figure 10. (**A**) The predicted lysozyme concentration response surface at various ammonium sulfate concentrations and pH values by the artificial neural network (ANN) method. (**B**) The predicted objective function surface from the ANN method.

bioprocess design, multiple assay agents will be required to work in parallel.

The communication between agents can be crucial, especially when the devices are not in the same location. Although devices for assays are available in most laboratories, they may not have been connected to the liquid-handling system where manual material handling is still needed. The communication among the agents in these cases may be realized through the computer network, although the throughput might be limited.

High-Throughput Equipment and Assays

The success of the intelligent platform largely depends on the development of high-throughput automation technology. Although a large number of devices have been developed and introduced into bioprocess development, most are for the experiment execution. Therefore, the bottleneck of the total platform efficiency is the assays, and more novel high-throughput assays are desirable to significantly improve the efficiency of the platform. Sometimes, advanced instruments with sophisticated data analysis software are required in research and development activities. The ability to automate such sophisticated data analysis will enhance the efficiency.

Experiment Design Algorithms

The method of the experiment design depends on the experimental objectives. In some cases, more than one algorithm can be combined to achieve better outcomes. For example,

a simplex search algorithm can be used to narrow down the design space before implementing DoE methods to investigate the properties of the response surface. New experiment design methods that address process dynamic behaviors are important to bioprocess development. It is also desirable for the experiment design agent to consider the objectives of the experiments and intelligently choose a suitable experiment design method.

The replications in the experiments needed for each process design depend on the accuracy of experimentation and assay. Error estimation-based statistical analysis may be used to aid the experiment design agent to determine the replications.

In conclusion, this article presented a method to establish an intelligent platform that automates every task to achieve rapid bioprocess design. The method is general and can be applied to each bioprocess design, provided microscale experimentation can be carried out on a robotic platform. A first-generation prototype of an intelligent automation platform to perform precipitation experiments automatically has been established. The platform is able to combine a liquid-handling robotic, a vacuum pump, a shaker, and a plate reader together to perform all procedures such as assembly adding, shaking, incubation, and UV absorption measurement automatically without human intervention. The prototype developed for precipitation is used as a vehicle to demonstrate the method. When applied to other bioprocess designs, the prototype stays the same while the experimental details may be different.

The experiment design agent uses two sequential experiment design algorithms, the simplex algorithm and an ANN,

demonstrating the automation of data analysis and experiment design in this novel integration approach to plan experiments within closed-loop learning. The simplex algorithm shows its power when searching for extreme points in a defined space, and the neural network modeling approach gives a better understanding of the effect of lysozyme solubility at various ion strength and pH values. Both of these methods can be applied to other bioprocess design.

The integration of experimental devices, a database, and algorithms for experiment design demonstrates a powerful tool for high-throughput process development. The "intelligent" experiment design algorithms keep the number of experiments low to reduce the amount of time and materials. The elimination of time delay by automating experimental procedure, data analysis and experiment design can reduce time further and make the platform a very effective process optimization tool.

The intelligent platform will need to be tested with more bioprocess design case studies. The requirements in whole bioprocess design will need more agents, and the interactions among these agents need to be much more dynamic to work together to achieve global goals in the future.

Acknowledgments

Declaration of Conflicting Interests

Funding

References

 1. Micheletti, M.; Lye, G. J. Microscale Bioprocess Optimisation. *Curr. Opin. Biotechnol.* **2006**, *17*, 611–618.
 2. Ferreira-Torres, C.; Micheletti, M.; Lye, G. Microscale Process Evaluation of Recombinant Biocatalyst Libraries: Application to Baeyer–Villiger Monooxygenase Catalysed Lactone Synthesis. *Bioprocess Biosyst. Eng.* **2005**, *28*, 83–93.
 3. Knevelman, C.; Davies, J.; Allen, L.; et al. High-Throughput Screening Techniques for Rapid PEG-Based Precipitation of IgG4 mAb from Clarified Cell Culture Supernatant. *Biotechnol. Prog.* **2010**, *26*, 697–705.
 4. Titchener-Hooker, N.; Dunnill, P.; Hoare, M. Micro Biochemical Engineering to Accelerate the Design of Industrial-Scale Downstream Processes for Biopharmaceutical Proteins. *Biotechnol. Bioeng.* **2008**, *100*, 473–487.
 5. Chhatre, S.; Titchener-Hooker, N. J. Review: Microscale Methods for High-Throughput Chromatography Development in the Pharmaceutical Industry. *J. Chem. Technol. Biotechnol.* **2009**, *84*, 927–940.
 6. Wiendahl, M.; Schulze Wierling, P.; Nielsen, J.; et al. High Throughput Screening for the Design and Optimization of Chromatographic Processes—Miniaturization, Automation and Parallelization of Breakthrough and Elution Studies. *Chem. Eng. Technol.* **2008**, *31*, 893–903.
 7. Arteaga, G. E.; Horimoto, Y.; Li-Chan, E.; et al. Partial Least-Squares Regression of Fourth-Derivative Ultraviolet Absorbance Spectra Predicts Composition of Protein Mixtures: Application to Bovine Caseins. *J. Agric. Food Chem.* **1994**, *42*, 1938–1942.
 8. Hansen, S. K.; Jamali, B.; Hubbuch, J. Selective High Throughput Protein Quantification Based on UV Absorption Spectra. *Biotechnol. Bioeng.* **2013**, *110*, 448–460.
 9. Mandenius, C.-F.; Brundin, A. Bioprocess Optimization Using Design-of-Experiments Methodology. *Biotechnol. Prog.* **2008**, *24*, 1191–1203.
10. Gheshlaghi, R.; Scharer, J. M.; Moo-Young, M.; et al. Medium Optimization for Hen Egg White Lysozyme Production by Recombinant *Aspergillus niger* Using Statistical Methods. *Biotechnol. Bioeng.* **2005**, *90*, 754–760.
11. Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308–313.
12. Chhatre, S.; Konstantinidis, S.; Ji, Y.; et al. The Simplex Algorithm for the Rapid Identification of Operating Conditions during Early Bioprocess Development: Case Studies in FAb′ Precipitation and Multimodal Chromatography. *Biotechnol. Bioeng.* **2011**, *108*, 2162–2170.
13. Susanto, A.; Treier, K.; Knieps-Grünhagen, E.; et al. High Throughput Screening for the Design and Optimization of Chromatographic Processes: Automated Optimization of Chromatographic Phase Systems. *Chem. Eng. Technol.* **2009**, *32*, 140–154.
14. Berkholz, R.; Guthke, R. Model Based Sequential Experimental Design for Bioprocess Optimisation—An Overview. In *Focus in Biotechnology, Engineering and Manufacturing for Biotechnology*; Hofmann, M., Thonart, P., Eds.; Kluwer Academic: Norwell, MA, 2002, vol. 4, pp. 129–141.
15. Franceschini, G.; Macchietto, S. Model-Based Design of Experiments for Parameter Precision: State of the Art. *Chem. Eng. Sci.* **2008**, *63*, 4846–4872.
16. McKenzie, P.; Kiang, S.; Tom, J.; et al. Can Pharmaceutical Process Development Become High Tech? *AIChE J.* **2006**, *52*, 3990–3994.
17. Abu-Absi, S. F.; Yang, L. Y.; Thompson, P.; et al. Defining Process Design Space for Monoclonal Antibody Cell Culture. *Biotechnol. Bioeng.* **2010**, *106*, 894–905.
18. Konstantinidis, S.; Chhatre, S.; Velayudhan, A.; et al. The Hybrid Experimental Simplex Algorithm—An Alternative Method for 'Sweet Spot' Identification in Early Bioprocess Development: Case Studies in Ion Exchange Chromatography. *Anal. Chim. Acta.* **2012**, *743*, 19–32.
19. King, R. D.; Rowland, J.; Oliver, S. G.; et al. The Automation of Science. *Science.* **2009**, *324*, 85–89.
20. Sycara, K. P. Multiagent Systems. *AI Magazine.* **1998**, *19*, 79–92.
21. Wooldridge, M.; Jennings, N. R. Intelligent Agents: Theory and Practice. *Knowledge Eng. Rev.* **1995**, *10*, 115–152.

22. Ferber, J.; Drogoul, A. Using Reactive Multi-Agent Systems in Simulation and Problem Solving. *Distributed Artif. Intell*. **1992**, *5*, 53–80.

23. Gao, Y.; Kipling, K.; Glassey, J.; et al. Application of Agent-Based System for Bioprocess Description and Process Improvement. *Biotechnol. Prog*. **2009**, *26*, 706–716.

24. Genesereth, M. R.; Ketchpel, S. P. Software Agents. *Commun*. ACM 1994, 37, 48–53, 147.

25. Davidsson, P.; Wernstedt, F. A Multi-Agent System Architecture for Coordination of Just-in-Time Production and Distribution. *Knowledge Eng. Rev*. **2002**, *17*, 317–329.

26. Smith, R. G. The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. *IEEE Trans. Comput*. **1980**, *100*, 1104–1113.

27. Wellman, M. P. A Computational Market Model for Distributed Configuration Design. *Artif. Intell. Eng. Design Anal. Manufact*. **1995**, *9*, 125–134.

28. Nii, P. The Blackboard Model of Problem Solving. *AI Magazine*. **1986**, *7*, 38–53.

29. Agatonovic-Kustrin, S.; Beresford, R. Basic Concepts of Artificial Neural Network (ANN) Modeling and Its Application in Pharmaceutical Research. *J. Pharm. Biomed. Anal*. **2000**, *22*, 717–727.

30. Zhang, X.; Zhang, S.; He, X. Prediction of Solubility of Lysozyme in Lysozyme–NaCl–H_2O System with Artificial Neural Network. *J. Crystal Growth* **2004**, *264*, 409–416.