

# A Multidisciplinary Team-Based Classroom Exercise for Small Molecule Drug Discovery

Charlotte A. Dodson,\* Stephen E. Flower,\* and Mark Thomas



Cite This: *J. Chem. Educ.* 2023, 100, 3320–3332



Read Online

ACCESS |



Metrics & More



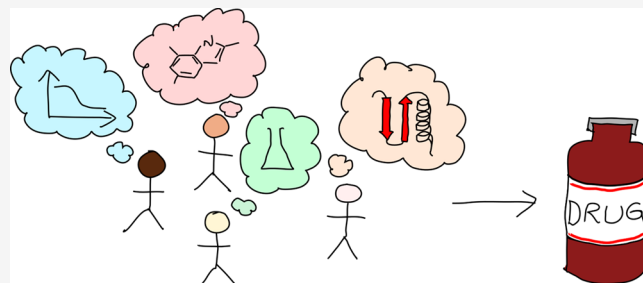
Article Recommendations



Supporting Information

**ABSTRACT:** Industrial drug discovery teams encompass scientists from multiple specialties and require participants to communicate effectively across disciplinary boundaries. In this paper, we present an undergraduate or graduate classroom simulation of this environment. Over a series of five workshops, student teams of mixed scientific backgrounds perform five iterations of the chemistry cycle of small molecule drug discovery. Students analyze physicochemical, structural, and (fictional) assay data and use these to design new compounds for testing. Simulated assay results are returned to students who use the information in the design of subsequent compounds. After workshop 5, each team selects a single lead compound, supported by a potential synthetic route, a portfolio of assay data, and logical scientific decision-making. Our exercise provides students with opportunities for hands-on student-responsive data handling, team-building, and technical knowledge acquisition—all within an industrially relevant scientific scenario.

**KEYWORDS:** Drug Discovery, Upper-Division Undergraduate, Graduate Education/Research, Interdisciplinary, Biochemistry, Organic Chemistry, Hands-On Learning, Problem Solving, Decision Making, Drugs, Pharmaceuticals



## INTRODUCTION

Many chemistry graduates progress to careers within the pharmaceutical industry. Chemists may be appointed to roles across the discovery and development process—from drug discovery, toxicology and clinical development to manufacturing, registration, sales and marketing. Many chemists are employed in research and development (R&D) in the design and synthesis of novel small molecule therapeutics. One scenario encountered within R&D is that of medicinal chemists working alongside assay scientists and structural biologists in the chemistry cycle of compound optimization (Figure 1). This scenario requires chemists to work in a multidisciplinary environment, taking account of biological data to develop high affinity molecules with good physicochemical properties, low predicted toxicity, and a tractable synthetic route—all while maintaining specificity for the target binding site. Traditionally, the first time a newly qualified chemist meets this multidisciplinary multiparameter optimization process is on their first commercial assignment.<sup>1</sup>

To meet the demand for specialist skills in the pharmaceutical industry, many institutions have launched Masters level courses in Drug Discovery, or incorporated this into existing degree offerings.<sup>1,2</sup> Within the MSc in Drug Discovery at our institution, we realized that we had the opportunity to develop a piece of coursework that drew on the diverse scientific background of the class while teaching all students about the chemistry cycle in detail. Material on drug discovery developed

elsewhere has focused on computer-aided drug design,<sup>3,4</sup> on analyzing and designing new compounds,<sup>4,5</sup> on wet laboratory projects<sup>2,6</sup> or on gamification of decision making<sup>7</sup> and has varied in the level of creativity required from the students. We wanted to ensure a high level of student creativity while focusing on the iterative nature of compound optimization and the integration of chemical and biological information. We chose an experiential learning framework because we felt that this approach would encourage the students to use their pre-existing knowledge to create new scientific content, while also learning to work—and perhaps most importantly communicate—in multidisciplinary teams.

This paper presents a virtual drug discovery exercise suitable for Master's or final year undergraduate Bachelor's students. Students from a mix of chemistry- and biology- focused backgrounds work in small teams to analyze the results of an initial screen of 30 fragments, iterate the chemistry cycle five times and develop a scientific case in support of their final lead compound (Figure 2). This is complemented with explicit team-forming activities and a mixture of traditional worksheets and

Received: January 20, 2023

Revised: June 23, 2023

Published: August 1, 2023



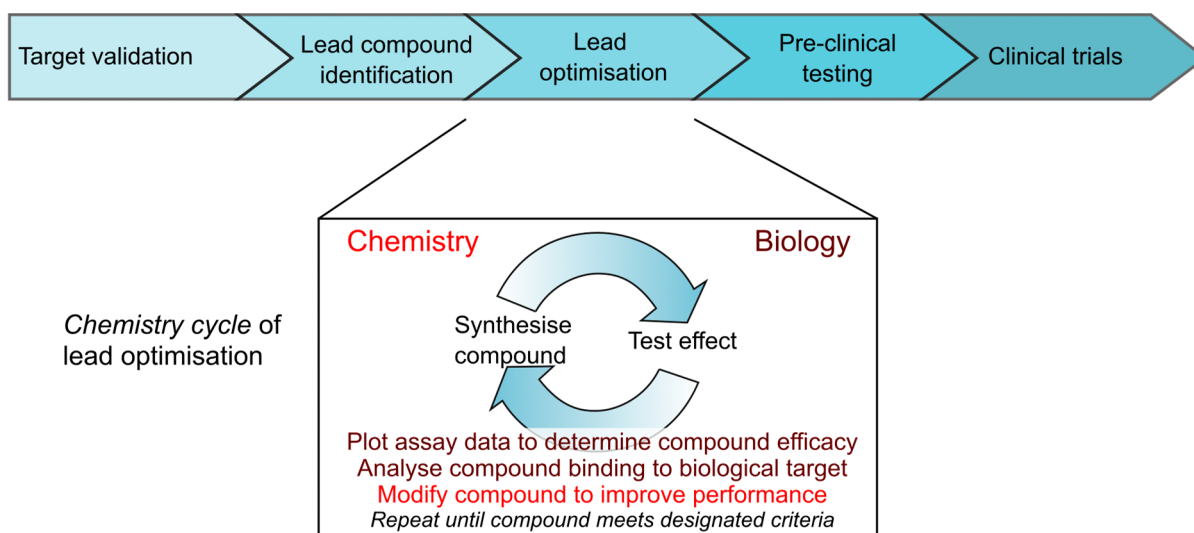


Figure 1. Overview of the Drug Discovery Process and the chemistry cycle of lead optimization.

workshop prereading on related topics using strategies borrowed from team-based learning ([www.teambasedlearning.org](http://www.teambasedlearning.org); Box 2 Figure 2).

## EXERCISE PREREQUISITES

### Software and Database Requirements

Students should have access to a number of software packages and databases. ChemDraw, Excel, SciFinder-n, and Reaxys are pay-for software packages or databases that are included in many University library and software subscriptions. Free alternatives to these software packages are listed below. PyMOL and SciDAVis are available free of charge.

- Excel (or other appropriate software if a different format is chosen for simulated data). Licensed copies of this software are widely available to students on University courses. Free alternatives include Google Sheets (<https://www.google.com/sheets/about/>) and Calc from LibreOffice ([www.libreoffice.org](http://www.libreoffice.org)).
- PyMOL ([www.pymol.org](http://www.pymol.org)) or equivalent protein visualization and manipulation software. Licenses for the commercial build educational use are free and can be obtained by instructor registration at <https://pymol.org/edu/>. PyMOL is available for Microsoft Windows, macOS, and Linux and students can download the software onto their own machines to work on in their own time. OpenSource PyMOL is available free of charge. (<https://github.com/schrodinger/pymol-open-source>) and installation instructions for Windows, MacOS and Linux were available on a number of independent Web sites in January 2023 (including [www.pymolwiki.org](http://www.pymolwiki.org)).
- SciDAVis ([www.sourceforge.net/projects/scidavis/](http://www.sourceforge.net/projects/scidavis/)) or other program suitable for nonlinear least-squares fitting of data. Alternative licensed software includes SigmaPlot or GraphPad Prism. SciDAVis is available for Microsoft Windows, macOS and Linux operating systems and is licensed under GNU General Public License version 2.0.
- ChemDraw (<https://perkinelmerinformatics.com/products/research/chemdraw/>) or other software suitable for drawing chemical structures. Alternatives include Marvin Sketch (<https://chemaxon.com/marvin>; free for academic and personal use) and Chem Sketch (<https://www.acdlabs.com/resources/free-chemistry-software-apps/chemsketch-freeware/>).

[www.acdlabs.com/resources/free-chemistry-software-apps/chemsketch-freeware/](http://www.acdlabs.com/resources/free-chemistry-software-apps/chemsketch-freeware/)). ChemDraw is available for Microsoft Windows and macOS.

- SciFinder-n (<https://scifinder-n.cas.org>). This curated database from the Chemical Abstracts Service (CAS) indexes the chemical literature and enables students to search for appropriate reactions to synthesize their target compound.
- Reaxys (<https://www.reaxys.com>). This index of the chemical literature is licensed by Elsevier and also enables students to search for appropriate reactions to synthesize their target compound.

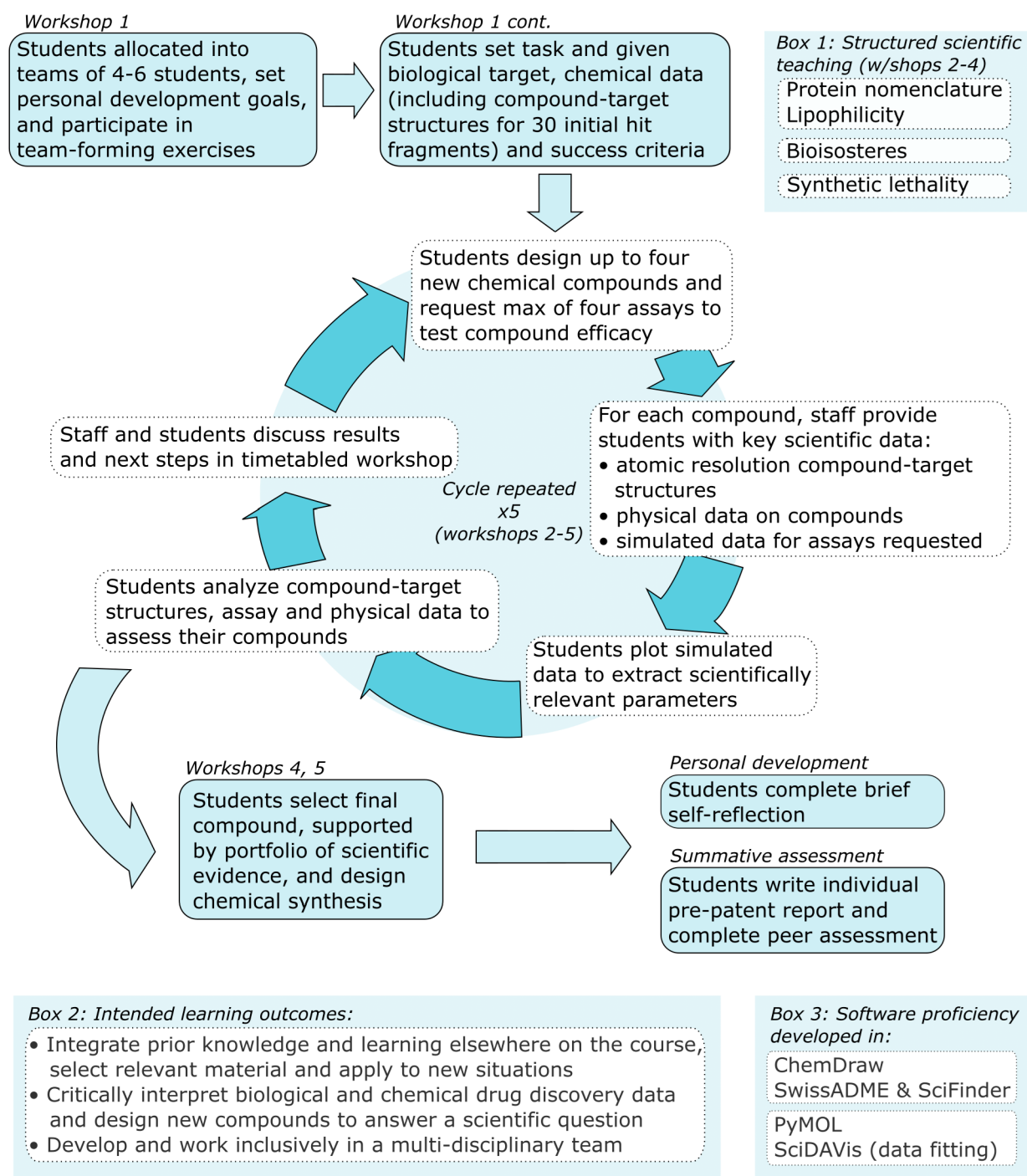
At least one staff member requires access to GraphPad Prism and GOLD, which are commercial pay-for-service software (free alternatives listed below).

- GraphPad Prism ([www.graphpad.com](http://www.graphpad.com)). Alternatives to Prism include programming environments such as Matlab, IDL or Python, and the expertise to write and/or run scripts generating data points from a mathematical equation with Gaussian noise. An example Jupyter notebook (Python) with this functionality is provided in [Supporting Information S21](#). Prism is available for Microsoft Windows and macOS.
- GOLD protein–ligand docking software (<https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/>)<sup>8,9</sup> and the expertise to run simple docking. Alternative free software includes Autodock Vina (<https://vina.scripps.edu>) which is released under an Apache License, Version 2.0 or other ligand/protein docking software. With the example kinase provided here—in which the protein is already prepared for docking—instructors should be able to self-teach using the instructions in [Supporting Information S1](#). As detailed in *Introduction of Student Docking*, an alternative to this would be to use the exercise in a format where students carry out their own dockings. In this case, students will require access to docking software.

There are no wet-lab requirements.

### Student Background

Students should have either chemistry- or biology- focused study at Levels 6 or 7 in the English Qualifications Framework<sup>10</sup> (e.g. a



**Figure 2.** Overview of the virtual drug discovery exercise.

Bachelor's/undergraduate degree in chemistry, biochemistry, pharmacology, pharmacy, or be a final year student in these subjects). The exercise draws on the broad disciplinary background of the students, and so a mix of disciplinary backgrounds within the student cohort is important.

Students should also have a working knowledge of the software packages that they will be using. We achieve this by including three 2 h hands-on tutorial workshops earlier in the MSc course covering ChemDraw, PyMOL, and SciDAVis. In these workshops students carry out exercises to draw chemical reaction schemes, manipulate protein structures or fit a dose-response ( $IC_{50}$ ) curve. Students also create a publication-quality

figure for each data type. The workshops are explicitly linked to the drug discovery exercise, and students are reminded to keep their notes for future reference. Handouts and/or instructor notes for the preparatory workshops are available from the National Teaching Repository.<sup>11</sup>

## ■ VIRTUAL DRUG DISCOVERY EXERCISE

### Overview of Exercise

In this virtual drug discovery exercise, teams of 3–5 students of mixed scientific backgrounds simulate a commercial drug discovery group progressing through the chemistry cycle of lead optimization (Figure 2). Each team is tasked with

developing a small molecule inhibitor against Chk1 kinase using a fragment-based drug discovery approach (Supporting Information S2). The final student-developed lead compound must meet prespecified criteria (Table 1) and its selection must be supported by a logical scientific process, a portfolio of (simulated) assay data, and a plausible chemical synthesis. Criteria for the lead compound are chosen to be intermediate between the Rule of Five<sup>12</sup> and the Rule of Three<sup>13</sup> and an example of student compound progression can be found in Supporting Information S3.

The exercise takes place in five rounds over five 2 h workshops and four 1 h (timetabled but unstaffed) student-led meetings. Except for Round 1 (which includes team forming and task setting instead of points i–iv), each round consists of (i) workshop prereading and a short associated multiple choice test in the workshop (workshops 2–4 only); (ii) class teaching points; (iii) fitting and appraising assay data from the previous round; (iv) workshop discussion of data and team-specific teaching points between students and teaching staff; (v) design of new compounds, selection of associated assay(s) and submission of both to teaching staff; and (vi) data simulation, compound docking and calculating physicochemical parameters by teaching staff for each compound submitted.

Data fitting and appraisal, compound design, assay selection, and scientific discussion (steps iii–v) take place during the workshop, student-led meeting, and in students' private study time. Appraisal of compounds, data simulation, docking, and preparation of physicochemical parameters (step vi) occur prior to each workshop. As a result of this responsive work by teaching staff, three files (Excel spreadsheet of assay data, pdb file of docked compounds, pdf of physical data) are uploaded to the VLE for the students.

At the end of the exercise, students submit a peer assessment of their team members and an individual summative report on the development of their compound written up in a format similar to a paper in the *Journal of Medicinal Chemistry* (example student report provided in Supporting Information S4). An example project timetable is given in Table 2.

### Detail of Exercise

Prior to the start of this exercise, students are preallocated into teams of approximately equal size and balance with respect to chemistry/biological sciences knowledge. We also try to balance the teams in terms of academic strength, gender, and national origin.

Workshop 1 falls into two halves: the first half of the session is dedicated to creating teams, setting goals, and designing ground rules for the exercise; the second half is spent on the scientific task and the specifics of compound modification and assay selection.

We have borrowed some of the strategies of the team-based learning approach ([www.teambasedlearning.org](http://www.teambasedlearning.org)) and, in Workshop 1, teams are asked to create a team name and team charter (up to six rules and up to three consequences for non-

compliance; Supporting Information S5 and S6). Students also spend time considering different roles within a team (e.g., chair, note-taker, meeting organizer) and the range of skills that they may need. To facilitate team introductions and to promote and practice some of the skills we would like the students to develop over the course of the workshops, students carry out an active listening<sup>14</sup> exercise. Students also each set a personal goal for self-development (which may remain private). Example slides covering these activities are provided in Supporting Information S7.

In the second part of Workshop 1, students are introduced to the task and presented with physicochemical data for 30 chemical fragments, (fictional) data from a thermal shift assay of the fragments against Chk1 and a PyMOL (.pse) file showing each fragment docked into the Chk1 ATP-binding site (Supporting Information S2 and S9).

In all workshops, students are set the task of analyzing the data that they have been given in terms of docked pose, hydrogen bonding interactions, thermal shift (or other assay) results, and physicochemical properties of drug-like molecules, e.g., lipophilicity (cLogP, LogD<sub>7.4</sub>, tPSA), solubility, bioavailability, and potentially toxic functional groups (Table 2). Based on their analysis, students are asked to design four new compounds which they would like to test in the next round of the chemistry cycle. We have found that most students find the task challenging and initially daunting. We are explicit about this and reassure the students that the exercise is about learning through doing, and that the point is to go through the process and learn from it rather than to design a chemical compound which will become the next blockbuster drug. In their reflection at the end of the exercise, many students look back on the process and comment on the extent to which their own understanding has developed.

Having designed a round of new compounds, students complete a team submission document in which they detail the chemical structure of each compound and specify one biochemical/biological assay from which they would like to receive simulated experimental data. (In later rounds, we frequently modify this rule to permit two assays per compound; see Table 2, note e.) Students also write a brief justification of the rationale behind their choices (Supporting Information S10). The justification acts both as a reminder for the students when writing up the exercise and to enable staff to be confident that a suitable logic is being employed by the students. The team submission document and an electronic file containing the designed chemical structures (e.g., .sdf or .cdx files) are uploaded to the course VLE by a specified deadline in advance of the next workshop. While uploading an electronic file containing the chemical structures is not essential, our experience is that doing this reduces staff preparation time for the next workshop and reduces the chances of introducing errors by redrawing compounds. We also find it useful to remind students that they should specify the stereochemistry around chiral centers. Student submissions are summarized in Table 3.

Although we provide some pointers, the choice of assay is entirely up to the students as we would like to encourage the students to draw on their existing knowledge and on material they have learned elsewhere in the course rather than to select options from a prescriptive list (likewise the selection and design of compound chemistry is also the choice of the students). In principle, this can provide teaching staff with some surprises, but in practice, we have found that student requests tend to match course material closely—or if not can easily be found online.

**Table 1. Required Properties of Team Lead Compound**

Property	Value
<i>in vitro</i> IC <sub>50</sub> for Chk1	<0.5 μM
cLogP	>3
LogD <sub>7.4</sub>	>~3
LogS	<~3
tPSA	>90 Å <sup>2</sup>

Table 2. Workshop Outline and Example Timetable

	Week Number	Event	Preclass Reading <sup>a</sup>	Classroom Task(s)	Compound/Assay Submission Date
Round 1	1	Workshop 1	None	(a) Initial team-forming (b) Task setting (c) Initial facilitated discussions	Week 2 (1 week prior to Workshop 2)
	2	Student-Led Meeting 1		Assay data analysis, compound design, and assay selection	
Round 2	3	Workshop 2	Lipophilicity & Nomenclature	(a) Short test on prereading and discussion arising (b) Short classroom exercise on protein nomenclature <sup>b</sup> (c) Assay data analysis, fragment development and assay selection	Week 4 (1 week prior to Workshop 3)
	4	Student-Led Meeting 2		Assay data analysis, compound design and assay selection	
Round 3	5	Workshop 3	Bioisosteres	(a) Short test on prereading and discussion arising (b) Overview of summative assessment <sup>c</sup> (c) Assay data analysis, fragment development and assay selection	Week 6 (1 week prior to Workshop 4)
	6	Student-Led Meeting 3		Assay data analysis, compound design and assay selection. Staff should also check uploaded compounds for likely ease of synthesis and for chemical flags re drug metabolism.	
Round 4	7	Workshop 4	Synthetic Lethality	(a) Short test on prereading and discussion arising (b) Introduction to Scifinder-n and Reaxys (c) Synthesis (d) Assay data analysis, fragment development, and assay selection	Week 7 Or 8 (3 days prior to Workshop 5) <sup>d,e</sup>
	7 or 8 <sup>d</sup>	Student-Led Meeting 4		Assay data analysis, compound design and assay selection	
Round 5	8	Workshop 5	None	(a) Synthesis (b) Assay data analysis, fragment development, and assay selection	Week 9 <sup>e</sup>
	9 or 10 <sup>f</sup>			Staff return Round 5 data to students	
	9 or 10 <sup>f</sup>			Students submit structure of final lead compound	
	10			Students submit summative report	

<sup>a</sup>Preclass reading is released on the VLE the week preceding the workshop. <sup>b</sup>Available in [Supporting Information S8](#) (Short classroom exercise on protein nomenclature). <sup>c</sup>This information is also available to students as a short video and FAQ document on the VLE. <sup>d</sup>We teach this exercise in Semester 2 (11 weeks long), and the report submission deadline is in week 10 (to avoid conflict with other deadlines on the course). This makes it impossible to timetable five fortnightly workshops and leave enough time for student submissions, data return, and report-writing. Workshops 4 and 5 are therefore timetabled in consecutive weeks for practical reasons. This impacts on the timetabling of student-led meeting 4, which should be approximately three working days before workshop 5 (to enable staff time to prepare the requested data). Student-led meeting 4 may therefore be in week 7 or 8 (depending on whether the workshops are at the start of end of the week). <sup>e</sup>Depending on the progress made with compound chemistry, we have permitted students to request up to two assays per compound in one or both of Rounds 4 and 5 (compared with one assay per compound in Rounds 1–3). This has permitted students time to develop the chemistry of their compounds while still collecting a rounded portfolio of assay data. <sup>f</sup>Easter usually falls between teaching weeks 9 and 10. This means that there are usually 2 weeks of vacation and 1 week of study between the submission of Round 5 compounds and the final report submission date. This enables staff to return Round 5 data to the students and for the students to upload the structure of their final lead compound at least a week before the deadline for the summative report (most groups have decided what their lead compound will be—or is likely to be—before the end of Workshop 5).

Table 3. Summary of Round-by-Round Student Submissions and Staff Uploads

Files Submitted by Students <sup>a</sup>	Files Uploaded by Staff <sup>a</sup>
Chemical Structure of Four New Compounds (.docx or .pdf and .cdx or .sdf file)	Physicochemical Data (.docx or .pdf file)
Request for One Biological/Biochemical Assay Per Compound (.docx or .pdf file)	Biological/Biochemical Assay Data (.xlsx file)
Brief Justification of Compounds and Assay Request (.docx or .pdf file)	Structural Biology Data (.pse file)

<sup>a</sup>Example formats and files are given in [Supporting Information S10–S12](#)

Early on, students tend to request biochemical IC<sub>50</sub> curves against isolated enzyme or K<sub>d</sub> values. Later on, requests include cell viability assays (healthy or disease model cell lines) or basic ADME/PK assays (e.g., Caco-2 or microsomal stability assays). The important feature of assay selection is for the students to think about selecting assays that enable them to answer a scientific question (e.g., which of these compounds binds most tightly to the target? Are the data from different compounds

comparable?) and, as the project progresses, to widen the range of assays selected to provide a portfolio of evidence in support of further development of their lead compound.

In terms of the chemistry, we encourage students to maximize favorable interactions with the protein binding pocket while keeping the physicochemical properties of the compounds they design within general parameters of lead-like compounds. We use prereading before workshops 2 and 3 to introduce isosteres and bioisosteres that the students can incorporate to optimize these parameters while maintaining the appropriate interactions (Table 2). Students often find it challenging to optimize multiple—sometimes contradictory—properties simultaneously. We address this directly by reminding students that they are tackling the kind of problem that is encountered in a real-world scenario, that there is no “right” answer, and that this is difficult. After round 3, student submissions are briefly assessed to check that a feasible synthetic route can be devised. We also consider whether a prodrug strategy to improve ADME properties or counter unwanted chemical modification during first pass metabolism would be appropriate (Table 2). This gives the

students time to incorporate the feedback (given in workshop 4) in their subsequent rounds of compound design, to discuss a novel synthetic route they have found that staff had not anticipated, or for staff to discourage further development of a particular compound through assigning it a high  $IC_{50}$ .

Prior to workshops 2–5 (and prior to decision of a team's final lead compound), teaching staff prepare physicochemical, assay, and structural biology results for each student-designed compound (see *Preparing Physicochemical and Simulated Data for Students* below). Physicochemical results are a simple table.

Structural biology results take the form of a .pse PyMOL file and represent the highest scoring docked pose of student compounds into the ATP binding site of Chk1. However, consistent with the hands-on nature of the exercise, biological/biochemical assay results are not simple “answers”. Instead, students are asked to determine  $IC_{50}$  values (or equivalent requested numerical value), together with associated errors from experimental replicates, by fitting sigmoid curves to simulated assay data using SciDAVis. For  $IC_{50}$  values,  $K_d$  values, cell viability assays, and similar quantities, assay results are therefore a table of numbers in .xlsx format. For ADME assays (and other data which lie further from the scientific expertise of us as teaching staff), experiments are often “contracted out” to a fictional contract research organization (*ADME Express Ltd.*) and data are provided in a variety of formats. Staff uploads are summarized in Table 3. Example data from a variety of student-requested assays in round 5 are provided in [Supporting Information S12](#).

Since this is a teaching exercise rather than a research project, we consider the simulated assay data we provide as formative feedback on the student compounds and strategy, rather than a reflection of what is a likely result in a laboratory setting (e.g., docking scores indicate that few student compounds are likely to have achieved tight enough binding to satisfy an experimental criterion of  $IC_{50} < 0.5 \mu M$ ). Therefore, the  $IC_{50}$  (or  $K_d$ ) value assigned to each compound is experimentally meaningless. Instead, these values are intended as formative feedback to the students: round 1 compounds tend to have  $IC_{50}$  values of 100s of micromolar and—assuming reasonable logical progression by the students—round 4 or 5 compounds progress to nanomolar  $IC_{50}$ s, even if this degree of improvement is unlikely given the compound chemistry. Nevertheless, within each round (and between rounds where appropriate), docking scores are used to rank the compounds for each team and to assign an initial  $IC_{50}$  value to each compound.

The initial  $IC_{50}$  value may be modified in a number of situations. For example, we tend to reward hydrogen bonding interactions (and likely specificity) over hydrophobic bulk alone by assigning a lower  $IC_{50}$  to compounds with H-bonds. In contrast, should compounds be progressing toward a chemistry for which there is a poor synthetic route, students may be steered away from or toward certain compounds by increasing or decreasing the  $IC_{50}$  values, respectively. Occasionally compound development does not follow a linear pattern (e.g., students return to earlier compounds to develop them further). In these cases, new  $IC_{50}$  values are allocated relative to the other  $IC_{50}$  values of the team using the docking scores as guidance (in order to meet the success criteria of the project, the  $IC_{50}$  values of these compounds often improve quickly!).

At the start of workshops 2–5, all three data sets (physicochemical, assay, and structural) are released to the students using the VLE. Students then examine and fit the data

in the workshop, and this forms the basis of staff-facilitated discussion.

In workshop 4, students begin to consider how their compounds can be synthesized (Table 2). Staff provide a short demonstration of SciFinder-n and Reaxys (chemistry-focused students are usually familiar with these databases already), and during the workshop students begin to create an outline plan for compound synthesis. Introducing synthesis considerations in workshop 4 enables the students to make any necessary modifications to the chemistry in rounds 4 and 5 (and to collect supporting assay data) to enable a successful summative report. Over rounds 4 and 5, chemistry-focused students work on designing a synthesis from plausible starting materials, while biology-focused students ensure that the final compound will be supported by a broad portfolio of assay data.

## ASSESSMENT

Compound strategy, data interpretation, and assay selection are assessed formatively throughout the exercise using the  $IC_{50}$  values allocated to compounds. Formative feedback is also provided on these aspects, on the synthetic route and on the breadth of scientific data in discussions between staff and student teams during the workshops. We assess the students' comprehension of the workshop prereading during the brief workshop multiple choice tests. As part of this, there is opportunity for peer-to-peer discussion and clarification of misunderstandings through use of a team-based MCQ.

Summative assessment is by an individual written report in support of the team's lead compound in the style of a prepatent report or *Journal of Medicinal Chemistry* article ([Supporting Information S4 and S13](#)). The contribution of each student to round by round team decision-making is reflected in their overall score by moderating 20% of their mark with a student-allocated peer assessment score ([Supporting Information S14](#)) as detailed in [eq 1](#).

$$\text{final mark} = 0.8 \times \text{staff mark} + 0.2 \times \text{staff mark} \times \text{peer score} \quad (1)$$

where final mark is the overall assessment mark, staff mark is the initial staff-assigned mark, and peer score a weighting factor based on the proportion of peer assessment score allocated to each student by the other students in their group. The calculation of peer score is detailed in [eq 2](#) and an Excel spreadsheet is provided in [Supporting Information S15](#).

$$\text{peer score}_i = \sum_j \left( \frac{\text{score}_{i,j}}{\sum_i \text{score}_{i,j}} \right) \quad (2)$$

where  $\text{peer score}_i$  is the peer score parameter for student  $i$ , and  $\text{score}_{i,j}$  is the peer assessment score allocated to student  $i$  from student  $j$ .

To reduce the burden of assessment to the students, students are encouraged to keep good records of their team discussions, and all submission documents must record the rationale behind each compound/assay choice. We are also explicit that students may share graphs, synthesis, and other figures between reports within a team and that doing so will not be penalized or constitute collusion. Nevertheless, we are clear that each student takes responsibility for the standard of the work which they submit, and that inclusion of poor quality work (eg poorly fit data, poorly drawn graphs) will result in a mark which reflects this. Prior to the exercise, all students have been trained in

software packages to prepare chemical structures, graphs, and illustrations of protein structure to publication standards and multiple students should have contributed to development of the final synthesis.

The mark scheme that we use is balanced between the biological and chemical aspects of the overall exercise and reflects many of the learning outcomes which we aim to achieve (Figure 2 box 2; Supporting Information S16). Our scheme enables us to differentiate students in all areas. Most—but not all—students are able to use the document template provided and correctly describe the task which they have undertaken, the submissions which they have made and the assay data which they have requested. The depth of discussion around the rationale for group decisions, the results obtained and the integration of different types of information in a narrative style varies widely. There is also considerable variation in the depth and relevance of material in the introduction and in student proficiency in writing an abstract. Strength in one of the introduction or results sections does not always indicate strength in the other. Students are differentiated by the quality of figures which they provide (including legibility of graph axes, presence of units, precision in reported fitted parameters, clarity and consistency in chemical structures, illustrative nature of protein structural information). Although a small number of layout marks are allocated for use of good scientific English, the focus when marking is on the science (and the extent to which this is communicated), not on proficiency in the English language. We do not explicitly assess the learning objectives around interdisciplinary communication (although students' achievement here should be reflected in the depth achieved in their summative report and peer assessment scores) and do not assess students' achievement of their aims for self-development.

## MODIFICATIONS OF THE EXERCISE

### Online Delivery

We have successfully taught this both in person and using online workshops (MSTeams). In both contexts, student compounds were submitted to staff and data returned to students using a VLE (in our institution, this ensures a centrally backed up record of the teaching). Irrespective of the mode of teaching delivery, we have found it useful to integrate MSTeams within the exercise, because this facilitates document-sharing and team chat and enables the submission of MCQ answers during workshops using MSForms. Although we formally return data to students using the institutional VLE for archiving purposes, we additionally upload these to a Team folder within MSTeams.

To prepare MSTeams in the context of this teaching, we first created an MSTeam for the entire class, and then created channels for each student team within this (for large cohorts, self-enrolment into Teams using email distributed links may prove helpful). Staff members were added to all channels. Data, chat, and week-to-week MCQ tests were then held within each channel.

For online delivery, each workshop was held as a timetabled MSTeams meeting, and students used the "Meet now" function within each channel to move into discussion groups. Staff members could then drop into and leave ongoing group discussions as required, and students could tag staff (using @staff\_member in the chat) as necessary in a manner analogous to circulating small group discussions.

## Introduction of Student Docking

Our most recent modification has been to increase student skills and reduce staff preparation time by asking students to dock their own compounds against Chk1 (ie to generate their own structural biology data). Students were taught how to use the docking software (in our case GOLD) during a hands-on 2 h workshop earlier in the course and were provided with a pdb of Chk1 prepared for docking, a configuration file of the procedure, and a file defining the target binding pocket in workshop 1 (Supporting Information S17–S19). To ensure that compounds were docked appropriately and that scores and poses were available to staff to use when ranking compounds and simulating assay data, students were required to upload two files to the VLE with their compound submission: the 3D coordinates of the docked compound (.pdb or .sdf file) and the bestranking.lst output file from GOLD containing the docking scores of the best poses (Table 4). Since all docking was carried out against the same .pdb, the relative orientations of the student compounds and the target protein were preserved. This made it straightforward to superpose student-docked compounds on the target protein for visual inspection.

The introduction of student-performed docking reduced the staff time needed to prepare between rounds, as intended. However, the requirements for student record-keeping increased, and students needed to be reminded repeatedly to ensure that their compounds were named consistently across each of their uploads.

### Other Modifications

There are, of course, many further possible modifications of this exercise. One of us (SEF) has taught a similar exercise to final year Chemistry students which omits biological/ADME data and focuses on structural biology, physicochemical properties, compound design, and synthesis. Alternatively, the focus could be moved away from chemical synthesis by focusing instead on compound design, physicochemical properties, and structural biology (with or without simulated wet-lab data).

## PREPARING PHYSICOCHEMICAL AND SIMULATED DATA FOR STUDENTS

### Physicochemical Data

This is calculated for each student-submitted compound by inputting a SMILES string for the compound into both SwissADME (<http://www.swissadme.ch/>)<sup>15</sup> and the Chemicalize tool developed by ChemAxon (<https://www.chemicalize.com/>; requires free registration). cLogP, LogS, and tPSA values are the consensus cLogP, ESOL, and TPSA values, respectively,

**Table 4. Summary of Round-by-Round Student Submissions and Staff Uploads (student docking)**

Files Submitted by Students	Files Uploaded by Staff
Chemical Structure of Four New Compounds (.docx or .pdf file)	Physicochemical Data (.docx or .pdf file)
Request for One Biological/Biochemical Assay Per Compound (.docx or .pdf file)	Biological/Biochemical Assay Data (.xlsx file)
Brief Justification of Compounds and Assay Request (.docx or .pdf file)	
Docking Output File (bestranking .lst file for GOLD software)	
Pose of Highest Scoring Docking for Each Compound (.pdb or .sdf file)	

Table 5. Graphpad Prism Simulation Parameters

Parameter	Value For IC <sub>50</sub>	Value for Mtr/Cck8 Assay	Value for Fp Binding Assay	Value for Cyp Inhibition Assay	Value for Comet Assay	Notes
Generate a Series of X-values Start at X =	0.005	0.005	0.005	0.001	0.005	This will depend on the IC <sub>50</sub> (so that the data points spread across the whole sigmoidal transition). Units should match those used for the IC <sub>50</sub> (we use M throughout).
Each value equals the previous value	Times 0.2	Times 0.2	Times 0.2	Times 0.2	Times 0.2	ie., A 5X dilution series
Generate values	8	8	8	8	8	Equivalent to #Rows on a 96-well plate
Equation	Generate Y Values from Equation List → Dose-Response-Inhibition → [Inhibitor] vs Response (Three Parameters)	Generate Y Values from Equation List → Dose-Response-Inhibition → [Inhibitor] vs Response (Three Parameters)	Generate Y Values from Equation List → Dose-Response-Inhibition → [Inhibitor] vs Response (Three Parameters)	Generate Y Values from Equation List → Dose-Response-Inhibition → [Inhibitor] vs Response (Three Parameters)	Generate Y Values from Equation List → Dose-Response-Inhibition → [Inhibitor] vs Response (Three Parameters)	This is a built-in equation in Graphpad Prism. The mathematical form of the inhibition equation is $Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + (X/IC_{50}))$ . The mathematical form of the horizontal line is $Y = \text{Mean} + 0 * X$ (all equations in Prism must contain X).
Parameter values Simulate... data sets	4	4	4	4	4	Or number of data sets required—one per compound
For each data set generate... replicate values	3	3	3	3	3	ie., assay carried out in triplicate
Bottom	50	0.01	0	0	N/A	Arbitrary value held constant across all compounds to simulate high inhibition control
Top	200	0.7	200	100	N/A	Arbitrary value held constant across all compounds to simulate low/no inhibition control
Midpoint (IC <sub>50</sub> , EC <sub>50</sub> , K <sub>d</sub> )	As chosen for compound—see main text	As chosen for compound—see main text	As chosen for compound—see main text	As chosen for compound—see Table 6	N/A	Units should match units used for X
Mean	N/A	N/A	N/A	N/A	20	Same as the value for zero compound.
Random error	Gaussian, absolute	Gaussian, absolute	Gaussian, absolute	Gaussian, absolute	Gaussian, absolute	Adding random error to each data point to simulate experimental scatter
SD	15	0.03	5	8	8	Arbitrary value chosen so that the simulated data resembles plausible experimental values.



from SwissADME.  $\text{LogD}_{7.4}$  is obtained from the Lipophilicity box in Chemicalize.

### IC<sub>50</sub> Values

Unless explicitly requested otherwise, we assume that we have run an ADP<sub>glo</sub> assay (Promega)<sup>16</sup> against recombinant Chk1 *in vitro* following the manufacturer's instructions. This commercially available assay measures [ATP], and reports [ADP] present in solution at the end of the 1 h assay run-time. A standard IC<sub>50</sub> graph will therefore plot [ADP] against [compound] (using a log scale on the *x*-axis) and give a sigmoidal plot which decreases in *y*-value with increasing *x*.

Experimental data is simulated using the "simulate data" function of GraphPad Prism. Briefly, we create a blank XY data table with *y* values in triplicate. This table is then "analyzed" using the "Simulate XY data" option in the "Simulate Data" family of analyses with the parameters given in Table 5. The values in the resulting data table are cut/pasted into Excel and uploaded to the VLE as experimental assay results for fitting by the students. A Prism file with example data sets is provided in Supporting Information S20. Alternatively, a python Jupyter notebook with similar functionality is provided in Supporting Information S21.

### Cell Viability Assays

Our students have requested either MTT<sup>17</sup> or CCK8<sup>18</sup> (Abcam) assays.<sup>19</sup> Both are colorimetric assays in which a tetrazolium salt is reduced by NAD(P)H in the cell to form a colored formazan product. The raw data are therefore a measured absorbance value at 460 nm, indicating the amount of formazan product produced. A low absorbance value indicates dead cells (low conversion to formazan), and a high absorbance indicates live cells (metabolizing cells giving high conversion to formazan).

Cell viability plots show the percentage of live cells (i.e., a percentage of an untreated control) against [compound], so the data we provide to students are simulated absorbance values for both a compound dilution series and control cells (zero compound). In practice, we generate data for the dilution series in Prism using the parameters in Table 5 and then rerun the calculation and copy an additional set of low-concentration values to act as the values for control cells. A Prism file with example data sets is provided in Supporting Information S20.

When choosing an EC<sub>50</sub> value for cell viability assays, the cLogP value<sup>20</sup> of each compound is used to determine whether or not the compound is likely to cross the cell membrane (or—occasionally—whether the compound is so hydrophobic it will remain in the membrane). If a compound is likely to enter the cell ( $1 < \text{cLogP} < 5$ ),<sup>21</sup> the EC<sub>50</sub> value is set to be 2–5 times larger than the IC<sub>50</sub> value (since the midpoints for cellular and *in vitro* assays are not directly comparable, likely due to different concentrations of ATP or physiological binding partners).

### Fluorescence Polarization (FP) Binding Assay

When students decide to measure the binding affinity of the compounds to the kinase directly, we assume that a fluorescence polarization competition assay<sup>22</sup> has been carried out. The physical basis of this assay is the different tumbling times of a fluorescently labeled Chk1 ligand free in solution and bound to Chk1: free ligand will tumble rapidly in solution, bound inhibitor will tumble more slowly. When the solution is illuminated with plane polarized light, the emitted fluorescence is also polarized. Polarization of the emitted light will be lower for unbound ligand molecules compared with bound ones, since

the rapid tumbling of the unbound inhibitor will remove the direction dependence of the emission more quickly.

We assume that the FP assay has been performed in a competition format, where an unlabeled test compound is titrated into a mixture of Chk1 and a tight-binding fluorescent tracer compound (at fixed concentration) under nonstoichiometric conditions. Unlabeled test compound competes with the fluorescent tracer and the measured fluorescence polarization signal gives a traditional sigmoidal IC<sub>50</sub> curve of signal against [compound] (using a log scale on the *x*-axis), decreasing in intensity with increasing *x*. Under these conditions, IC<sub>50</sub> is mathematically equivalent to the dissociation constant of the compound for the enzyme ( $K_d$ ),<sup>23</sup> and so the binding affinity of the compound for Chk1 has been determined experimentally.

In practice, experimental data are simulated using the "simulate data" function of GraphPad Prism to provide triplicate fluorescence polarization values for a test compound titration using the parameters in Table 5. A Prism file with example data sets is provided in Supporting Information S20.

### DMPK Data

The following DMPK assays are "outsourced" to a fictional contract research organization *ADME Express Ltd.* for the practical reason that these assays are well beyond our personal expertise. As such, *ADME Express* provides the results of the assays rather than the raw data for data plotting. Students tend to request these assays in the final rounds of the exercise when ensuring that they have a balance of assay data to support further development of their compound. Therefore, the intended learning outcomes here relate to assay selection rather than data handling.

### Caco-2 Permeability Assays

*In vitro* Caco-2 permeability assays monitor the extent to which compounds cross a polarized monolayer of Caco-2 cells and act as a model of intestinal permeability (for the uptake of orally administered compounds).<sup>24</sup> When students request that a Caco-2 assay is carried out on their compound, we provide values for two commonly reported parameters: the apparent permeability of the compound,  $P_{\text{app}}$  (apical-basolateral), and the efflux ratio,  $P_{\text{app}}$  (basolateral-apical)/ $P_{\text{app}}$  (apical-basolateral).

To select values for these parameters, we input the student compound into the online server SwissADME (<http://www.swissadme.ch/>)<sup>15</sup> using a SMILES string (easily output from ChemDraw). The server reports a parameter "GI permeability" (based on a prediction by reference 25) and the output from this (e.g., "high") is used to choose a value for  $P_{\text{app}}$  according to the guidance in Table 6.

In order to determine the value that is selected for the efflux ratio, we use the SwissADME prediction for whether the compound is a substrate for the P-glycoprotein 1 efflux pump  $P_{\text{gp}}$ . Compounds that are  $P_{\text{gp}}$  substrates will be pumped out of intestinal epithelial cells back into the intestinal lumen in an ATP-dependent manner. An efflux ratio greater than one indicates that efflux occurs, an efflux ratio less than one indicates that efflux does not occur and we have selected values of around 0.92 for compounds with good efflux ratios. In practice, efflux ratios less than two are acceptable, so for those compounds that are predicted to be  $P_{\text{gp}}$  substrates, we have chosen efflux ratios close to this upper limit. This is for the practical reason that teams tend to request ADME assays in rounds 4 or 5 (i.e., late in the development of their compounds), and we would like to enable students to write up a reasonably successful drug discovery story, rather than one that is fully realistic.

Table 6. Guide Values for ADME Parameters

Property	$P_{\text{app}}/\text{nm s}^{-1}$	$\text{CL}_{\text{int}}/\mu\text{L min}^{-1} \text{mg protein}^{-1}$	$\text{IC}_{50}/\mu\text{M}$
	Caco-2 and PAMPA Assays	Microsomal Stability	CYP Inhibition
High Permeability <sup>26</sup>	>100		
Medium Permeability <sup>26</sup>	34–100		
Low Permeability <sup>26</sup>	<34		
High Metabolism <sup>a</sup>		>47.0	
Low Metabolism <sup>a</sup>		<8.6	
Potent Inhibition <sup>b</sup>			<1
Moderate Inhibition <sup>b</sup>			1–10
No or Weak Inhibition <sup>b</sup>			>10

<sup>a</sup>Values calculated as described in the main text. <sup>b</sup>Inhibition bands from <https://www.cypotex.com/admepk/in-vitro-metabolism/cytochrome-p450-inhibition>

### PAMPA Assays

The *in vitro* parallel artificial membrane permeability assay (PAMPA) measures the extent to which a compound crosses an artificial hexadecane or lipid membrane<sup>24</sup> and acts as a model of passive compound absorption across the gut wall. Once more, PAMPA reports an apparent permeability of the compound ( $P_{\text{app}}$ ) and so the SwissADME prediction of GI permeability<sup>15</sup> is used to inform the value that we select for student compounds (guided by the values in Table 6).

### Liver Microsomal Stability Assays

The majority of drug metabolism in the human body occurs in the liver, and estimates of the metabolic stability of a compound are often obtained by *in vitro* experiments on different subcellular fractions of liver cells. Assays using liver microsomes are popular since microsomes contain many of the relevant oxidative enzymes (including CYP enzymes), and these assays report an intrinsic clearance value,  $\text{CL}_{\text{int}}$ .<sup>24</sup>

When students request a liver microsomal stability assay, the online prediction tool PredMS (available at <https://predms.netlify.app/>)<sup>27</sup> is used to obtain a probability of compound stability,  $P_{\text{stable}}$ . Compounds for which  $P_{\text{stable}} < 0.5$  are likely to be unstable, compounds for which  $P_{\text{stable}} > 0.5$  are likely to be stable. Based on the PredMS output, we then choose a value of  $\text{CL}_{\text{int}}$  to report to the students using the guide values in Table 6 (calculated as detailed below).

The relationship between  $\text{CL}_{\text{int}}$ , extraction ratio ( $E$ ), liver blood flow ( $Q_{\text{H}}$ ), microsomal protein per gram of liver (MPPGL), liver mass ( $m_{\text{liver}}$ ), and fraction of drug unbound in blood plasma ( $f_{\text{u}}$ ) is given in equation 3.<sup>28</sup> Using the approach taken by the Cypotex (<https://www.cypotex.com/admepk/in-vitro-metabolism/microsomal-stability>) and given in ref 29 and assuming liver blood flow of  $1450 \text{ mL min}^{-1}$ ,<sup>30</sup> liver mass of  $1800 \text{ g}$ ,<sup>30</sup> and completely free drug (i.e.,  $f_{\text{u}} = 1$ ), boundary values of  $E = 0.3$  and  $E = 0.7$  were used to calculate clearance values for low and high clearance respectively.

$$\text{CL}_{\text{int}} = \frac{E \cdot Q_{\text{H}} \cdot \text{MPPGL}}{f_{\text{u}} \cdot (1 - E) \cdot m_{\text{liver}}} \quad (3)$$

### CYP Inhibition

Cytochrome P450 (CYP) superfamily enzymes are monooxygenases, a handful of which are responsible for the majority of oxidative drug metabolism and elimination.<sup>31</sup> When students request that assays be carried out to determine CYP inhibition, we assume that an  $\text{IC}_{50}$  value has been determined for the inhibition of the CYP enzymes in liver microsomes by the compound of interest. In this assay, the amount of a defined CYP substrate remaining after the assay duration (expressed as a percentage) is determined using LC-MS/MS.

We use SwissADME to predict whether the student compound is likely to be a substrate for five different CYP enzymes (CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4) and use this to choose an overall  $\text{IC}_{50}$  value for the *in vitro* assay based on the guide values in Table 6. We then simulate  $\text{IC}_{50}$  data in GraphPad Prism using the parameters in Table 5.

### Plasma Protein Binding

Small molecule drugs may bind to proteins in human blood (eg human serum albumin or  $\alpha_1$ -acid glycoprotein). When complexed with blood proteins, compounds will not be taken up into tissues and so will be pharmacologically inactive.<sup>24</sup>

We assumed that equilibrium dialysis against blood plasma proteins has been carried out. In this assay, the concentration of free drug either side of an equilibrated dialysis membrane is measured and the fraction of compound unbound is reported.<sup>24</sup> This assay is requested infrequently, but when it is we submit student compounds to the PreADMET server (accessed via <https://preadmet.qsarhub.com/>)<sup>32</sup> and report the fraction unbound ( $f_{\text{u}}$ ) (calculated as  $(100 - \text{Plasma\_Protein\_Binding parameter})/100$ ).

### Comet Genotoxicity Assay

The Comet assay is a conceptually simple genotoxicity assay in which single cells are immobilized in agarose, lysed, and undergo electrophoresis.<sup>33</sup> Expansion of cellular DNA away from a compact nucleoid attached to the nuclear matrix toward the anode occurs most rapidly when the DNA has multiple sites of breakage (and thus relaxation of supercoiling). When visualized using fluorescence microscopy, this expansion resembles the tail of a comet (hence, the assay name).

The degree of tail expansion is quantified as a relative intensity value (% of tail DNA) and plotted against compound concentration to visualize toxicity: a flat line (no change with compound concentration) indicates no toxicity while a linear increasing relationship between % tail DNA and compound concentration indicates toxicity. At zero compound, the value of % tail DNA is around 20%.

Students tend to request a Comet assay toward the close of the project, so if there are no obvious flags to the contrary (which should have been corrected as a result of earlier discussions), results for a negative Comet assay are simulated in Prism by creating a set of compound dilution data for a flat line using the parameters in Table 5.

### Compound Docking

Molecular docking is a computational technique that approximates the process of a small molecule ligand binding to a protein active site. It uses iterative rounds of ligand pose prediction and scoring to optimize protein–ligand interactions and provide physically informed models of the 3D structure of the final ligand-bound complex. The magnitude of the docking score is used as a qualitative estimate of the strength of the

protein–ligand interaction and—in the absence of information to the contrary—the model with the highest docking score is accepted as a model of the likely ligand binding pose.<sup>8,34</sup>

Instructions to enable generalist staff to dock student-drawn compounds into Chk1 are included in [Supporting Information S1](#). Briefly, compounds are designed by the students and uploaded as ChemDraw files. Students are reminded to specify stereochemistry around chiral centers and staff check for unspecified stereocenters prior to carrying out compound docking. Student-designed molecules are copy/pasted into Chem3D by staff members and a short round of MM2 energy minimization is carried out. The energy-minimized compound is saved as a .sdf file. Ten rounds of docking into the ATP binding pocket of Chk1 are carried out for each compound using GOLD software and the highest scoring pose is used as a model of the final structure. A .pdb file of Chk1 prepared for docking, a file defining the ATP binding cavity, and a configuration file for docking in GOLD are included in the Supporting Information ([Supporting Information S18 and S19](#)).

## ■ PREPARATION OF INITIAL DATA FOR STUDENTS AND FOR MOLECULAR DOCKING

The 6FCK crystal structure of Chk1 kinase was taken from the Protein Data Bank and prepared for docking by removing sulfate ion, ligand ((S)-2-phenyl-4-(piperidin-3-ylamino)-1H-indole-7-carboxamide) and water molecules. Hydrogens atoms were added, assuming a pH of 7.0 ([Supporting Information S17](#)). The binding site was defined as a sphere of 10 Å radius centered on the centroid of the ligand present in the original crystal structure ([Supporting Information S19](#)).

Thermal shift data were generated by first docking a selection of chemical fragments against this prepared molecule. Fragments were ranked according to the GOLD docking score and then allocated a nominal  $\Delta T_m$  in the thermal shift assay. The  $\Delta T_m$  was calculated as (docking score  $-4$ )/12 + 0.2 with a manual adjustment of 0.1 or 0.2 °C for a few fragments. Final  $\Delta T_m$  values ranged from 1.5–4.2 °C.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available at <https://pubs.acs.org/doi/10.1021/acs.jchemed.3c00066>.

README; **File S1**: Brief instructions for docking ligands into Chk1.docx/pdf; **File S2**: Team brief and fragment screen.docx/pdf; **File S3**: Example of compound progression.pdf; **File S4**: Example of excellent student work.pdf; **File S5**: Blank Team Charter.docx/pdf; **File S6**: Example team rules and commitments.docx/pdf; **File S7**: Sample slides for workshop 1.pptx/pdf; **File S8**: Workshop 2 nomenclature.pdf; **File S9**: Fragment screen hits.pse; **File S10**: Team submission document.docx; **File S11**: Blank physicochemical data form.docx; **File S12**: Example data provided to students.xls; **File S13**: Summative assessment template.docx; **File S14**: Peer assessment and feedback.docx/pdf; **File S15**: Peer evaluation.xlsx; **File S16**: Example mark scheme.docx/pdf; **File S17**: Chk1.pdb; **File S18**: gold.conf; **File S19**: cavity.atoms; **File S20**: Template Prism file for data simulation.pzf; **File S21**: Python code for simulating experimental data.zip (ZIP)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Charlotte A. Dodson** – Department of Life Sciences, University of Bath, Bath BA2 7AY, U.K.; [orcid.org/0000-0002-8403-726X](https://orcid.org/0000-0002-8403-726X); Email: [c.a.dodson@bath.ac.uk](mailto:c.a.dodson@bath.ac.uk)

**Stephen E. Flower** – Department of Chemistry, University of Bath, Bath BA2 7AY, U.K.; [orcid.org/0000-0003-2659-3118](https://orcid.org/0000-0003-2659-3118); Email: [s.e.flower@bath.ac.uk](mailto:s.e.flower@bath.ac.uk)

### Author

**Mark Thomas** – Department for Health, University of Bath, Bath BA2 7AY, U.K.

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jchemed.3c00066>

### Notes

The authors declare no competing financial interests.

## ■ ACKNOWLEDGMENTS

We are grateful to James Taylor and Steven Bull (Department of Chemistry), for their prior development of a simpler exercise used with final year chemistry undergraduates at University of Bath. We are also grateful to Begoña Delgado-Charro and Nikoletta Fotaki (School of Pharmacy, Department of Life Sciences) for discussions and for access to their lecture material on ADME and DMPK, and to Patrick Taylor (Department of Chemistry) for the Jupyter notebook. We are particularly grateful to the Efficacious Four (Grace V, Adam H, Siqi W, Maria P) for permission to include their work as examples of student compound progression, to Siqi W for the basis of [Supporting Information file S3](#), and to Adam H for granting us permission to include his project report as [Supporting Information file S4](#). This exercise is part of several related Drug Discovery teaching projects at University of Bath and materials are available at <https://doi.org/10.25416/NTR.c.6388596>. Version controlled updates for this project are available at <https://doi.org/10.25416/NTR.21908502>. The work in this paper was funded by the University of Bath and a CC-BY license is applied to the version of record arising from this submission.

## ■ REFERENCES

- (1) Rafferty, M. F. No Denying It: Medicinal Chemistry Training Is in Big Trouble. *J. Med. Chem.* **2016**, *59* (24), 10859–10864.
- (2) McNally, T.; Macdonald, S. J. F. Unusual Undergraduate Training in Medicinal Chemistry in Collaboration between Academia and Industry. *J. Med. Chem.* **2017**, *60* (19), 7958–7964.
- (3) Clent, B. A.; Wang, Y.; Britton, H. C.; Otto, F.; Swain, C. J.; Todd, M. H.; Wilden, J. D.; Tabor, A. B. Molecular Docking with Open Access Software: Development of an Online Laboratory Handbook and Remote Workflow for Chemistry and Pharmacy Master's Students to Undertake Computer-Aided Drug Design. *J. Chem. Educ.* **2021**, *98* (9), 2899–2905. Rodrigues, R. P.; Andrade, S. F.; Mantoani, S. P.; Eifler-Lima, V. L.; Silva, V. B.; Kawano, D. F. Using Free Computational Resources To Illustrate the Drug Design Process in an Undergraduate Medicinal Chemistry Course. *J. Chem. Educ.* **2015**, *92* (5), 827–835.
- (4) Tantillo, D. J.; Siegel, J. B.; Saunders, C. M.; Palazzo, T. A.; Painter, P. P.; O'Brien, T. E.; Nuñez, N. N.; Nouri, D. H.; Lodewyk, M. W.; Hudson, B. M.; Hare, S. R.; Davis, R. L. Computer-Aided Drug Design for Undergraduates. *J. Chem. Educ.* **2019**, *96* (5), 920–925.
- (5) Veale, C. G. L.; Krause, R. W. M.; Sewry, J. D. Blending problem-based learning and peer-led team learning, in an open ended 'home-

- grown' pharmaceutical chemistry case study. *Chemistry Education Research and Practice* **2018**, *19* (1), 68–79.
- (6) Faville, S. C.; Harris-Hamdscomb, K.; Harker, O.; Mattison, S.; Tamorite, H.; Bristowe, J.; Daly, D.; Ege, R.; He, H.; Jones, J.; McCorkindale, A.; Mei, K.; Monson, A.; Moree, L.; Perkovic, F.; Rickerby, G.; Robinson, J.; Rudkin, F.; Whibley, L.; Worthington, R.; Ennis, C.; de la Harpe, S.; Brind, T.; Hopkins, A.; Winefield, K.; Hendrickx, S.; Caljon, G.; Perry, B.; Vernall, A. J. Open Synthesis Network Research in an Undergraduate Laboratory: Development of Benzoxazole Amide Derivatives against Leishmania Parasite. *J. Chem. Educ.* **2022**, *99* (4), 1682–1690. Daniels, D.; Berkes, C.; Nekoie, A.; Franco, J. Fighting Tuberculosis in an Undergraduate Laboratory: Synthesizing, Evaluating and Analyzing Inhibitors. *J. Chem. Educ.* **2015**, *92* (5), 928–931. García-Caballero, M.; Moya-García, A.; Torres-Vargas, J. A.; García-Ponce, A. L.; Rodríguez-Quesada, A. A course-based undergraduate research experience to illustrate the early stages of the drug discovery process. *Biochemistry and Molecular Biology Education* **2022**, *50* (5), 437–439.
- (7) Yap, W. H. Role Play in Drug Discovery and Development: Biosciences student performance and perspectives when role play experiential learning activities are integrated into learning. *Prog. Drug Discov Biomed Sci.* **2019**, *2* (1), a0000032. Lounsbury, N.; Chow, N. K.; Jackson, K. C.; Chbib, C. Creation and implementation of a drug discovery and development game. *Currents in Pharmacy Teaching and Learning* **2022**, *14* (2), 229–234. Covey, J. R.; Guarascio, A. J.; O'Donnell, L. A.; Tidgewell, K. J. Development and Implementation of a Global Health Elective with a Drug Discovery Game for Pharmacy Students. *Pharmacy* **2017**, *5*, 49.
- (8) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267* (3), 727–748.
- (9) Liebeschuetz, J. W.; Cole, J. C.; Korb, O. Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. *J. Comput. Aided Mol. Des.* **2012**, *26* (6), 737–748.
- (10) UK Quality Code for Higher Education (Part A: Setting and Maintaining Academic Standards), 2014. <https://www.qaa.ac.uk/the-quality-code>.
- (11) (a) Dodson, C. A. A self-paced tutorial for visualising protein structures using PyMOL. *National Teaching Repository* **2022**, DOI: 10.25416/NTR.21610089. Dodson, C. A. A self-paced tutorial to least-squares data fitting using SciDAVis. *National Teaching Repository* **2022**, DOI: 10.25416/NTR.21623028. (b) Taylor, J. ChemDraw Key Skills Workshop. *National Teaching Repository* **2023**, DOI: 10.25416/NTR.22574944.v1.
- (12) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46* (1–3), 3–26. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44* (1), 235–249.
- (13) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discov Today* **2003**, *8* (19), 876–877.
- (14) Rogers, C. R.; Farson, R. E. Active Listening. In *Communication in Business Today*; Newman, R. G., Danziger, M. A., Cohen, M., Eds.; Heath and Company, 1987.
- (15) Daina, A.; Michielin, O.; Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717.
- (16) Zegzouti, H.; Zdanovskaia, M.; Hsiao, K.; Goueli, S. A. ADP-Glo: A Bioluminescent and homogeneous ADP monitoring assay for kinases. *Assay Drug Dev Technol.* **2009**, *7* (6), 560–572.
- (17) Mosmann, T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* **1983**, *65* (1–2), 55–63.
- (18) Ishiyama, M.; Miyazono, Y.; Sasamoto, K.; Ohkura, Y.; Ueno, K. A highly water-soluble disulfonated tetrazolium salt as a chromogenic indicator for NADH as well as cell viability. *Talanta* **1997**, *44* (7), 1299–1305.
- (19) Riss, T. L.; Moravec, R. A.; Niles, A. L.; Duellman, S.; Benink, H. A.; Worzella, T. J.; Minor, L. Cell Viability Assays. In *Assay Guidance Manual*; Markossian, S., Grossman, A., Brimacombe, K., Arkin, M., Auld, D., Austin, C. P., Baell, J., Chung, T. D. Y., Coussens, N. P., Dahlin, J. L., Devanarayan, V., Foley, T. L., Glicksman, M., Hall, M. D., Haas, J. V., Hoare, S. R. J., Inglese, J., Iversen, P. W., Kales, S. C., Lal-Nag, M., Li, Z., McGee, J., McManus, O., Riss, T., Saradjian, P., Sittampalam, G. S., Tarselli, M., O. Trask, J., Wang, Y., Weidner, J. R., Wildey, M. J., Wilson, K., Xia, M., Xu, X., Eds.; Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2004.
- (20) Leo, A. J.; Hoekman, D. Calculating log P(oct) with no missing fragments; The problem of estimating new interaction parameters. *Perspectives in Drug Discovery and Design* **2000**, *18* (1), 19–38. Mannhold, R.; Poda, G. I.; Ostermann, C.; Tetko, I. V. Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *J. Pharm. Sci.* **2009**, *98* (3), 861–893.
- (21) Alamudi, S. H.; Satapathy, R.; Kim, J.; Su, D.; Ren, H.; Das, R.; Hu, L.; Alvarado-Martinez, E.; Lee, J. Y.; Hoppmann, C.; Pena-Cabrera, E.; Ha, H. H.; Park, H. S.; Wang, L.; Chang, Y. T. Development of background-free tame fluorescent probes for intracellular live cell imaging. *Nat. Commun.* **2016**, *7*, 11964. Bennion, B. J.; Be, N. A.; Mc Nerney, M. W.; Lao, V.; Carlson, E. M.; Valdez, C. A.; Malfatti, M. A.; Enright, H. A.; Nguyen, T. H.; Lightstone, F. C.; Carpenter, T. S. Predicting a Drug's Membrane Permeability: A Computational Model Validated With in Vitro Permeability Assay Data. *J. Phys. Chem. B* **2017**, *121* (20), 5228–5237.
- (22) Lea, W. A.; Simeonov, A. Fluorescence polarization assays in small molecule screening. *Expert Opin Drug Discov* **2011**, *6* (1), 17–32.
- (23) Huang, X. Fluorescence polarization competition assay: the range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand. *J. Biomol. Screen* **2003**, *8* (1), 34–38.
- (24) *Everything You Need to Know about ADME*; Cyprotex.
- (25) Daina, A.; Zoete, V. A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem* **2016**, *11* (11), 1117–1121.
- (26) Egan, W. J.; Merz, K. M.; Baldwin, J. J. Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* **2000**, *43* (21), 3867–3877.
- (27) Ryu, J. Y.; Lee, J. H.; Lee, B. H.; Song, J. S.; Ahn, S.; Oh, K. S. PredMS: A random Forest model for predicting metabolic stability of drug candidates in human liver microsomes. *Bioinformatics* **2022**, *38*, 364.
- (28) Houston, J. B. Utility of in vitro drug metabolism data in predicting in vivo metabolic clearance. *Biochem. Pharmacol.* **1994**, *47* (9), 1469–1479.
- (29) (a) Cyprotex. *Microsomal Stability Assay*, 2022. <https://www.cyprotex.com/admepk/in-vitro-metabolism/microsomal-stability> (accessed 12/05/2022). (b) Barter, Z. E.; Bayliss, M. K.; Beaune, P. H.; Boobis, A. R.; Carlike, D. J.; Edwards, R. J.; Houston, J. B.; Lake, B. G.; Lipscomb, J. C.; Pelkonen, O. R.; Tucke, G.; Rostami-Hodjegan, A. Scaling factors for the extrapolation of in vivo metabolic drug clearance from in vitro data: reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. *Curr. Drug Metab* **2007**, *8* (1), 33–45. (c) Iwatsubo, T.; Suzuki, H.; Shimada, N.; Chiba, K.; Ishizaki, T.; Green, C. E.; Tyson, C. A.; Yokoi, T.; Kamataki, T.; Sugiyama, Y. Prediction of in vivo hepatic metabolic clearance of YM796 from in vitro data by use of human liver microsomes and recombinant P-450 isozymes. *J. Pharmacol. Exp. Ther* **1997**, *282* (2), 909–919.
- (30) Davies, B.; Morris, T. Physiological parameters in laboratory animals and humans. *Pharm. Res.* **1993**, *10* (7), 1093–1095.
- (31) (a) Wilkinson, G. R. Drug metabolism and variability among patients in drug response. *N Engl J. Med.* **2005**, *352* (21), 2211–2221. (b) Zhao, M.; Ma, J.; Li, M.; Zhang, Y.; Jiang, B.; Zhao, X.; Huai, C.; Shen, L.; Zhang, N.; He, L.; Qin, S. Cytochrome P450 Enzymes and Drug Metabolism in Humans. *Int. J. Mol. Sci.* **2021**, *22* (23), 12808.
- (32) Lee, S. K.; Lee, I.H.; Kim, H.J.; Chang, G.S.; Chung, J.E.; No, K.T. The PreADME Approach: Web-based program for rapid

prediction of physico-chemical, drug absorption and drug-like properties. In *EuroQSAR 2002: Designing Drugs and Crop Protectants: Processes, Problems and Solutions*; 2002.

(33) (a) Gharsalli, T. Comet Assay on Toxicogenetics; Several Studies in Recent Years on Several Genotoxicological Agents. *J. Environ. Anal. Toxicol.* **2016**, DOI: 10.4172/2161-0525.1000418. (b) Collins, A. R.; Annangi, B.; Rubio, L.; Marcos, R.; Dorn, M.; Merker, C.; Estrela-Lopis, I.; Cimpan, M. R.; Ibrahim, M.; Cimpan, E.; Ostermann, M.; Sauter, A.; Yamani, N. E.; Shaposhnikov, S.; Chevillard, S.; Paget, V.; Grall, R.; Delic, J.; Goñi-de-Cerio, F.; Suarez-Merino, B.; Fessard, V.; Hogeveen, K. N.; Fjellsbø, L. M.; Pran, E. R.; Brzicova, T.; Topinka, J.; Silva, M. J.; Leite, P. E.; Ribeiro, A. R.; Granjeiro, J. M.; Grafström, R.; Prina-Mello, A.; Dusinska, M. High throughput toxicity screening and intracellular detection of nanomaterials. *WIREs Nanomedicine and Nanobiotechnology* **2016**, 9 (1), No. e1413.

(34) Pagadala, N. S.; Syed, K.; Tuszynski, J. Software for molecular docking: a review. *Biophys Rev.* **2017**, 9 (2), 91–102.