JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

FOCUS — Development of a Global Communication and Modeling Platform for Applied and Computational Medicinal Chemists

Nikolaus Stiefl,*[,†] Peter Gedeck,[‡] Donovan Chin,[§] Peter Hunt,[∥] Mika Lindvall,[⊥] Katrin Spiegel,[∥] Clayton Springer,[§] Scott Biller,[§] Christoph Buenemann,[#] Takanori Kanazawa,[▽] Mitsunori Kato,[§,▽] Richard Lewis,[†] Eric Martin,[⊥] Valery Polyakov,[⊥] Ruben Tommasi,[§] John van Drie,[§] Brian Vash,[§] Lewis Whitehead,[§] Yongjin Xu,[⊥] Ruben Abagyan,[○] Eugene Raush,[○] and Max Totrov[○]

[†]Novartis Institute for Biomedical Research, Postfach, CH-4002 Basel, Switzerland

[‡]Novartis Institute for Tropical Diseases Pte. Ltd., 10 Biopolis Road, #05-01 Chromos, Singapore 138670

[§]Novartis Institute for Biomedical Research, 250 Massachusetts Ave., Cambridge, Massachusetts 02139, United States

[∥]Novartis Institute for Biomedical Research, Wimblehurst Road, Horsham, West Sussex RH12 5AB, United Kingdom

[⊥]Novartis Institute for Biomedical Research, 4560 Horton Street, Bldg. 4, Emeryville, California 94608, United States

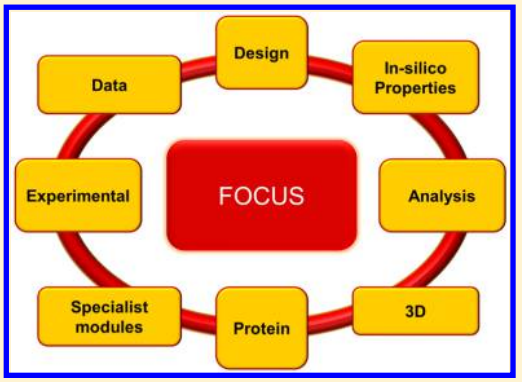
[#]Novartis Institute for Biomedical Research, 11/2 Muthgasse, 1190 Wien, Austria

[▽]Novartis Institutes for BioMedical Research, Ohkubo 8, Tsukuba, Ibaraki 300-2611, Japan

[○]Molsoft L.L.C., 11199 Sorrento Valley Road, S209, San Diego, California 92121, United States

Supporting Information

ABSTRACT: Communication of data and ideas within a medicinal chemistry project on a global as well as local level is a crucial aspect in the drug design cycle. Over a time frame of eight years, we built and optimized FOCUS, a platform to produce, visualize, and share information on various aspects of a drug discovery project such as cheminformatics, data analysis, structural information, and design. FOCUS is tightly integrated with internal services that involve—among others—data retrieval systems and *in-silico* models and provides easy access to automated modeling procedures such as pharmacophore searches, R-group analysis, and similarity searches. In addition, an interactive 3D editor was developed to assist users in the generation and docking of close analogues of a known lead. In this paper, we will specifically concentrate on issues we faced during development, deployment, and maintenance of the software and how we continually adapted the software in order to improve usability. We will provide usage examples to highlight the functionality as well as limitations of FOCUS at the various stages of the development process. We aim to make the discussion as independent of the software platform as possible, so that our experiences can be of more general value to the drug discovery community.



INTRODUCTION

Communication of design ideas and concepts between theoretical (modelers) and applied (bench) medicinal chemists can be a challenge within a drug discovery project. Outside project meetings, the major method used to exchange plans and ideas are slides, chemical drawings on paper, or in electronic form and emails. The specific method used depends on the team member in question. Unfortunately, the tools to communicate within this framework are often only partially overlapping. Depending on the user community, the major workhorse will be more computationally (e.g., PYMOL,[1] Spotfire,[2] MOE,[3] Maestro,[4] Stardrop[5]) or more chemistry (Chemdraw,[6] Instant JChem[7]) driven. In global companies with a large variety of tools and mixed heritage (through mergers and reorganizations), this results in many hours spent in translation between the different media or repeated training in tools that are not used on a regular basis; this is ultimately a very frustrating experience for the end user.

To circumvent these problems, many companies have implemented systems to support the scientists in their projects. Two main strategies can be identified: full in-house development (e.g., OSIRIS,[8] ABCD[9]) and vendor solutions (e.g., D360,[10] Biovia,[11] VIDA[12]). For specific tasks, these tools are often supplemented with specialized user interfaces (e.g., LUCID[13]).

We chose a hybrid approach[14] using Molsoft's ICM[15] software, on top of which we built a customized tool,

Received: October 3, 2014
Published: March 27, 2015

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

"FOCUS". We used standard technologies like HTML in combination with ICM's internal scripting language. FOCUS was originally designed for our medicinal chemistry community but is now also used in many other parts of the organization. The main drivers of FOCUS were to distribute a simple tool for decision making to the scientist and to facilitate data and information sharing. We also felt that many standard tasks, such as virtual library enumeration or *in-silico* property calculations, had become robust enough that they could be semiautomated for use by nonexperts. These tasks were later extended by the addition of a 3D editing capability in FOCUS to facilitate structure-based idea generation and designs by medicinal chemists. The development was done over roughly eight years in four stages starting from a functionality-rich client with limited connectivity to external systems to a system that is interfacing with a wide range of computational and data services within the Novartis Institute of Biomedical Research (NIBR) organization. The functionality of the latest version can be best described as a combination of a central hub to gather web service information, complemented by tools developed internally in the ICM scripting language.

Today, FOCUS has evolved to a successfully deployed, widely used, and accepted tool in the scientists' armory at NIBR. However, given the fact that FOCUS is highly optimized for the environment at NIBR and that it is built on a commercial product, describing highly specific technical details is of little value to the global research community.[16] Given these in-house dependencies, depositing the source code (our preferred disclosure strategy)[17] is also not useful; instead, we will concentrate on the challenges we faced during the development of FOCUS, illustrated by different usage examples at each stage. We hope that our experience can help others who are in the process of designing such a strategy to avoid some of the issues we faced and more quickly reach a solution of similar benefit to the researchers.

MATERIALS AND METHODS

Integration. The purpose of FOCUS is to provide an easy to use tool to handle two-dimensional as well as three-dimensional data in a single interface. It accesses project-relevant web services such as a data hub, crystal structure information, property and model calculations, or similarity searches (Figure 1).

The services are accessed using RESTful web services with JSON queries or direct web requests. Data intense web services (like the data hub) are hosted on multiple sites, whereas others are only available from a single location. In addition to common services, like cLogP predictions, some convenience functions such as the extraction and conversion of compound identifiers into chemical structures directly from a text (via cut and paste, for example, directly from an email) are included (for a full list of web services used, see the Supporting Information). Depending on the output, the returned data can be post-processed in a way that maximizes communication and impact (Figure 2). For example, using single click interfaces, numerical values can be reformatted (qualifier handling, log scaling) or color coded by threshold (hERG models, solubility models, etc.); chemical structures can also be annotated (pK_a predictions, predicted sites of metabolism).[18] For specific models, customized output such as images can also be included in the data table (ligand interaction plots, clogD curves, etc.).

Modules. FOCUS is organized around modules with specific tasks. The modules themselves are interconnected so the user can follow them in a workflow of typical research tasks (e.g, import project data → find common scaffold of subseries → do R-group decomposition → analyze). The major modules are built around six areas: Data, Design, InSilico, Analysis, Small Molecule 3D, and Protein (Figure 3). Usage examples will be discussed later.

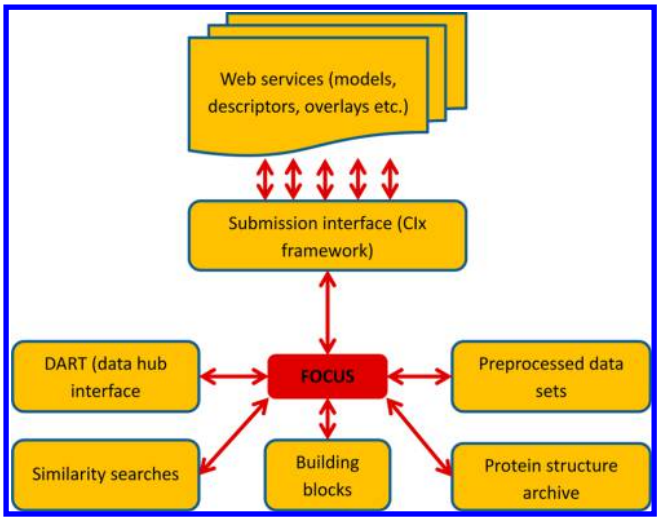


Figure 1. Overview of FOCUS integration within the NIBR environment. All general project-relevant tools are accessible from a single interface. Many of the web services are accessed through a generic submission interface used also by other tools in NIBR, the cheminformatics (CIx) framework.

Additional modules for specific research groups (peptide group, protein structure group, etc.) or the Experimental module are not visible by default but can be enabled by the user. The Experimental module is used for development and deployment of new or infrequently used tasks. Generally, new feature requests can be deployed within hours or days. A list of all available modules and the respective functionality contained in them is provided in the Supporting Information.

All modules have the look and feel of web pages with descriptive text and action links to provide a familiar GUI to inexperienced users. Links can be connected to simple scripts that work automatically on active tables or selected 3D objects. Additionally, more complex graphical interfaces are available for tasks that need specific selections (i.e., subsets of descriptors; Figure 4).

The creation of the customized interfaces is reasonably straightforward using the ICM scripting language to enable active HTML pages, programmable menus, dialogues and panels, and table actions and to translate the front end to the back end commands or external service calls. This specific part of the underlying ICM software was modified and adapted to our needs for an improved user interface during the course of the development of FOCUS.

3D Editing. The 3D editor was designed to provide a simple interface for chemists to explore analogs of an initial ligand within a protein−ligand complex. Versions were iteratively optimized during the development of FOCUS. In very early versions, only two-dimensional modifications, followed by constrained local minimizations of ligand molecules in the protein pocket environment without any scoring, were possible. Over time, this was extended by (a) the introduction and evaluation of "on-the-fly" modifications to the ligand structure in 3D, (b) the prediction of novel ligand poses, and (c) the

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

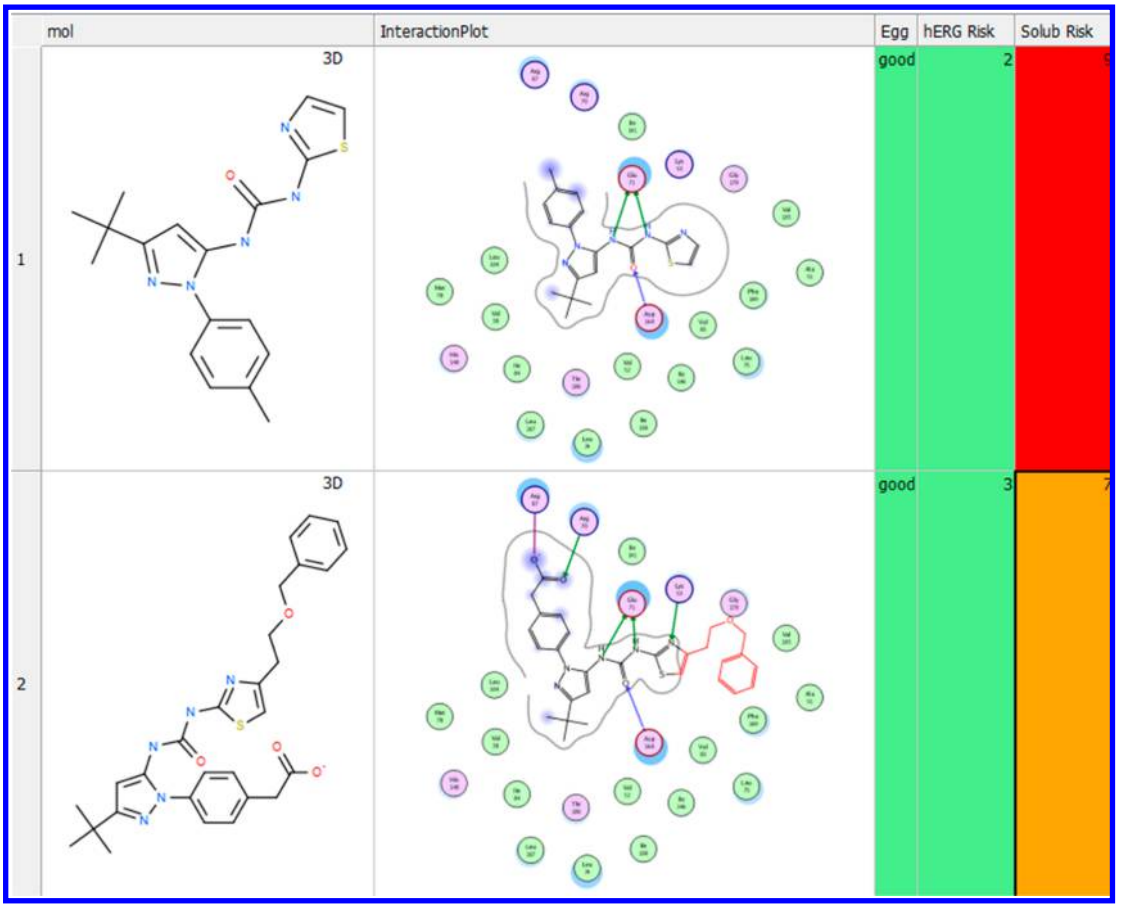


Figure 2. Example of two public domain structures extracted from PDB entries 3LFB and 3LFC.[19] The table contains color-coded calculated property columns as well as information integrated from other external software and methods (i.e., MOE interaction plots).

evaluation of their relative energies. These operations are described below.

Typically, 3D editor sessions are set up by modeling experts and then distributed to medicinal chemistry teams for use. Experience shows that this is especially important because results can be of very low quality if the initial model is not set up properly.

Local minimization of the ligand in the binding pocket is still one of the most commonly used user actions. This feature eliminates minor steric conflicts and/or optimizes polar interactions while retaining the overall pose of the ligand. The main usage of this feature is the exploration of minor modifications to a known ligand to get an impression of size constraints of ligands inside the binding pocket. Such minimization can be performed using the full-atom representation of the receptor pocket or with approximations using grid potentials.

(a). Ligand Editing Functionality. The 3D molecular editor implements functionality for editing the ligand *in-situ*. This includes substitution of groups (one-by-one or in a screening mode), linkers and scaffolds connecting the specified groups (important for scaffold hopping), or editing the ligand structure in 2D. Each modification is synchronously introduced in the 3D model of the ligand, conformationally optimized, and evaluated in terms of a docking score.

(b). Pose Prediction. For interactive use, it is important that this operation is performed in near real-time. Ligand pose prediction is performed by stochastic global optimization of the flexible ligand molecule (represented in internal coordinates) inside a set of precalculated grid potential maps representing the binding pocket of the receptor.[20,21] To further reduce runtimes, only a subset of possible grid potential maps are used to approximate interactions (i) van der Waals (represented as Lennard–Jones potential for hydrogen, carbon, and "large atom" probes mapped onto a 0.5 Å grid), (ii) electrostatic potential (calculated by the Coulomb formula with the distance-dependent dielectric constant of $4r$; charges on the protein atoms are scaled according to the SChEM approach[22] using polarization charge densities calculated by boundary element method[23]), (iii) hydrogen bonding potential (combining the donor and acceptor fields

$$E_{hb} = 2.5 \times e^{-(\vec{r}-\vec{r}_0)^2/1.96}$$

where vector $\vec{r}_0$ represents the ideal position of the hydrogen bonding partner and is placed 1.7 Å away from the donor/acceptor atom), and (iv) apolar surface energy (

$$E_{hp} = 3 \times e^{-d^2/2.8^2}$$

where d is the shortest distance from the grid point to the solvent accessible surface around the hydrophobic atoms of the receptor).

Prior to sampling, multiple starting poses of the ligand are generated by exhaustively sampling the ligand *in-vacuo* and overlaying each of the resulting conformations onto the binding pocket in four principal orientations. The sampling phase is performed using biased probability Monte Carlo optimization[24] of rotational, translational, and torsional variables of the ligand

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896–908

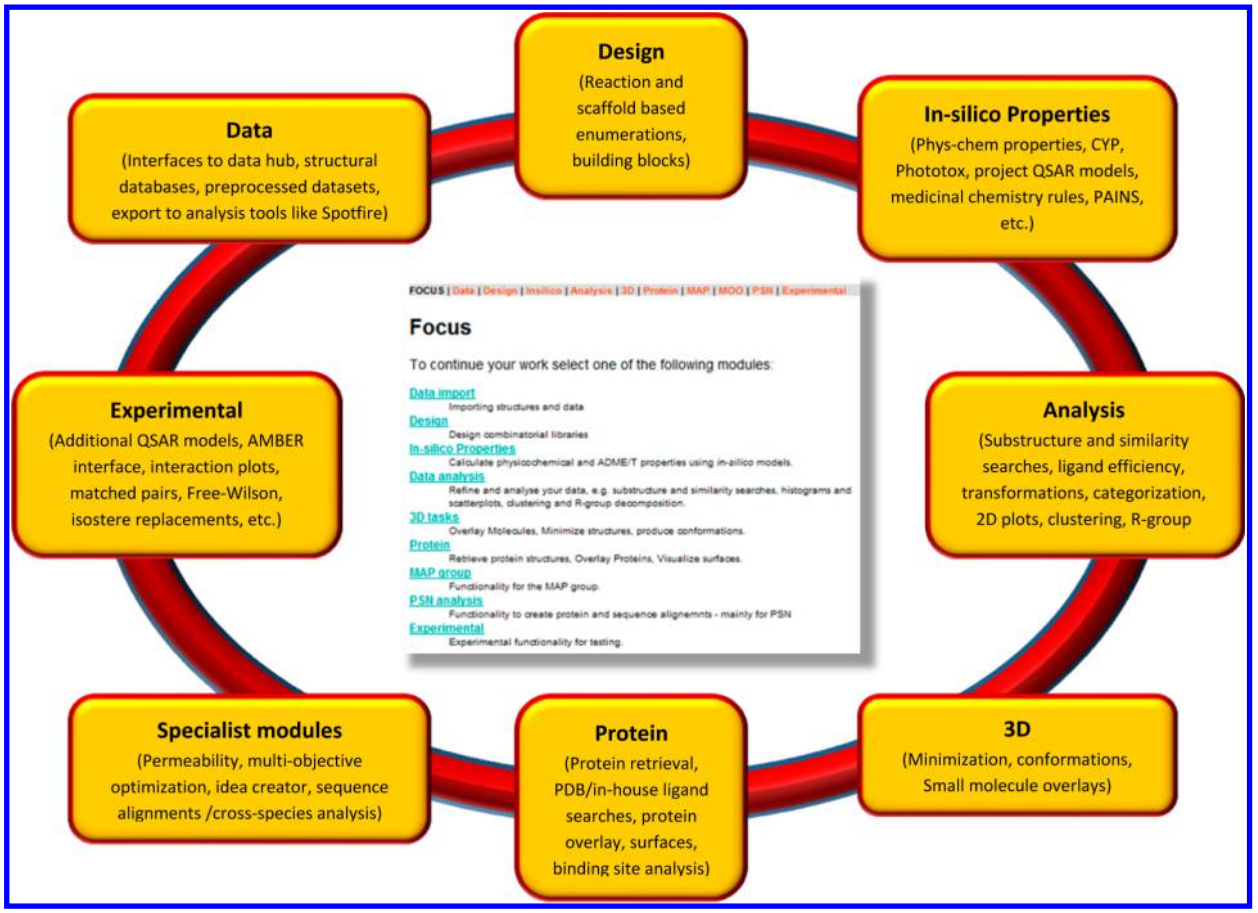


Figure 3. FOCUS modules and their interconnections. Typical project tasks such as patent analysis or property filtering can be easily performed through working with multiple modules.

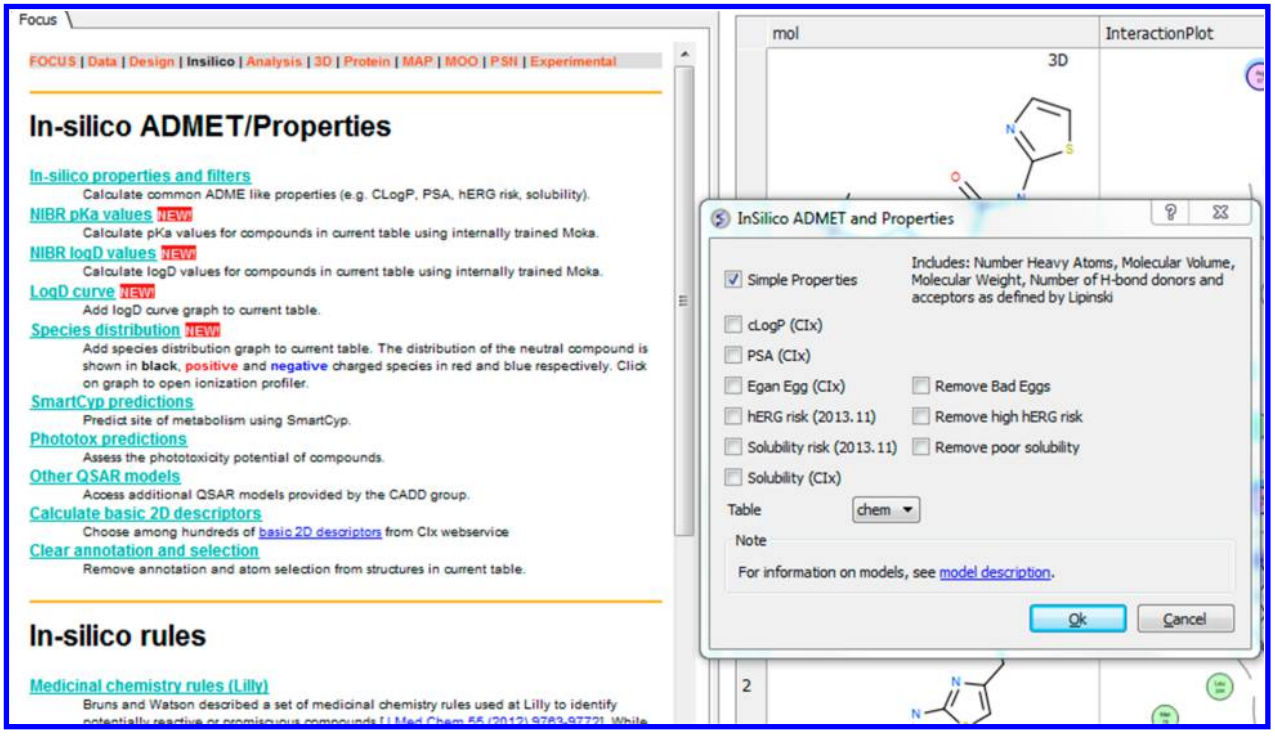


Figure 4. FOCUS example module (here the "InSilico" module). Short help texts are combined with action links. More elaborate help texts are also available as links. Display level of help texts is customizable. The separate checkbox interface obtained when clicking the "*In-silico* properties and filters" is shown. This display hierarchy enables complex interfaces within a familiar GUI.

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

(and explicit receptor parts if applicable) in an attempt to achieve the global minimum of the objective energy function. This energy function is a combination of ligand interactions with the receptor grid potentials and the explicit (full-atom) interactions within the ligand itself ("ligand strain").

Specific modifications and constraints can be applied by the user to the protein or the ligand. For proteins, those options are the exclusion or down-weighting of uncertain and unimportant protein atoms (e.g., those with low occupancy or high B-factor), usage of multiple conformations of the binding pocket,[25] and exclusion of fragments from grid potential maps calculation but inclusion at the simulation stage as full-atom models allowing explicit conformational sampling. For ligands, positional and distance restraints can be imposed onto the ligand atoms. Positional restraints ensure that the Cartesian coordinates of an atom do not significantly change throughout the simulation, while distance restraints maintain the specified distance between two atoms (ligand−receptor or ligand−ligand) regardless of their Cartesian coordinates

(c). Pose Scoring and Graphical Evaluation. Following the sampling phase, top-ranking predicted poses of the ligand are re-evaluated using full-atom representation of the receptor pocket, here the ICM ligand binding score.[26] To the user, the score is reported with and without the ligand strain term because for some complicated ligand chemotypes, the inaccuracies of strain calculation introduce unwanted noise. In addition, specifically highlighting the ligand strain is useful for the medicinal chemists to further analyze the docking poses provided.

A lot of extra details are given to provide an improved and interactive binding mode analysis for nonexpert modelers (Figure 5).

"Ligand surface" is a graphic mesh display of the Lee−Richards surface[28] representing the imaginary space where ligand atom centers should be placed for optimal interactions with the receptor. The surface is colored according to preferred atom types at each location such as "nonpolar", "aromatic", "hydrogen bond donor", or "hydrogen bond acceptor". "Atomic energy circles" provide estimates of per atom scores and allow quick visualization of local favorable and unfavorable (e.g., steric clash) contacts between the ligand and the pocket. The hydrogen bond display shows, in a quantitative fashion, strong and weak hydrogen bonds and allows visualization of polar buried atoms that are not engaged in hydrogen bonds ("unsatisfied hydrogen bonds") that introduce unfavorable contributions to the binding energy. In addition, it is possible to locally minimize the bound conformation to the next local energy minimum *in-vacuo* (without protein interactions) to estimate the conformational constraints.

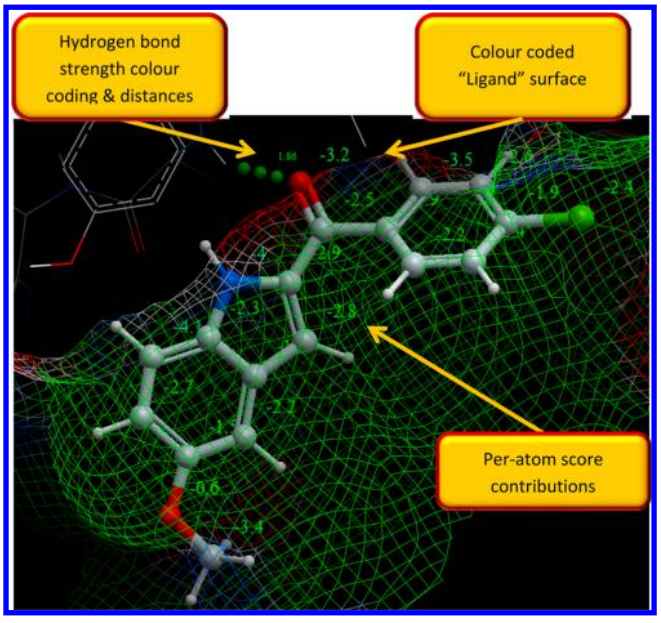


Figure 5. Close-up of a modified compound in a binding pocket (here PDB entry 3AD4;[27] see also Figure 8). Color coded ligand surfaces (green/white for more hydrophobic and red/blue for polar areas), hydrogen bond strengths, and per atom score contributions allow nonexperts to interpret the model and consequently come up with novel ideas more easily.

Deployment. The FOCUS library is a set of ICM function definitions available as separate files. Those are stored in a Subversion[29] repository and for Windows and Linux globally deployed to local file servers using automated Jenkins[30] services without the necessity of any user interaction. At NIBR, only few chemists use Mac OS X. For those users, a specific, well-documented setup process is available for the manual installation and update by the end users. Upon startup of ICM, all relevant FOCUS functionality is loaded from these shared local network drives. During the startup process, the standard ICM GUI is modified, macros and dialogues are read, and FOCUS specific variables are defined.

The network drives are standard, fast access storage systems that are also used for other shared project areas containing, for example, presentations or data spreadsheets. No specific server setup is necessary for this task because the library comprises ASCII files only and less than 10 MB of data are read upon startup. This setup allows the developers to produce very fast modifications to the FOCUS code without the organizational burden of involving multiple groups in the dissemination of these updates.

All FOCUS functions collect information on their usage. For this, each FOCUS function is either tagged as a helper function (e.g., table column name processing) or a method function (e.g., cLogP calculation). Information on usage of method functions is sent to a central server for further auditing. After careful revision, this information is used to identify and remove unused functionality. This process helps to declutter the interface, identify training or documentation needs, and adjust budget priorities for the underlying software licenses.

The same usage information is also partly used in the deployment of new functionality. After discussion of requirements and implementation, new user requests are initially deployed through the Experimental module. After an optimization phase, a review on the general applicability of the method is performed. Here, it is decided in cooperation with users whether the functionality will be removed, transferred to a production module, or kept in an experimental state. In addition, a more general review of any functionality is done on a regular basis.

Collaboration between Development Teams. Teams at NIBR and Molsoft collaborated closely to extend the needed capabilities of ICM in scriptable GUI programming, startup hooks, interfaces to external services such as JSon or SOAP, and gaps in functionality and speed. Some important or CPU-intensive functionality such as the 3D editor (see above) was implemented directly into the underlying ICM code. For the 3D editor, design cycles around topics involving, for example,

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

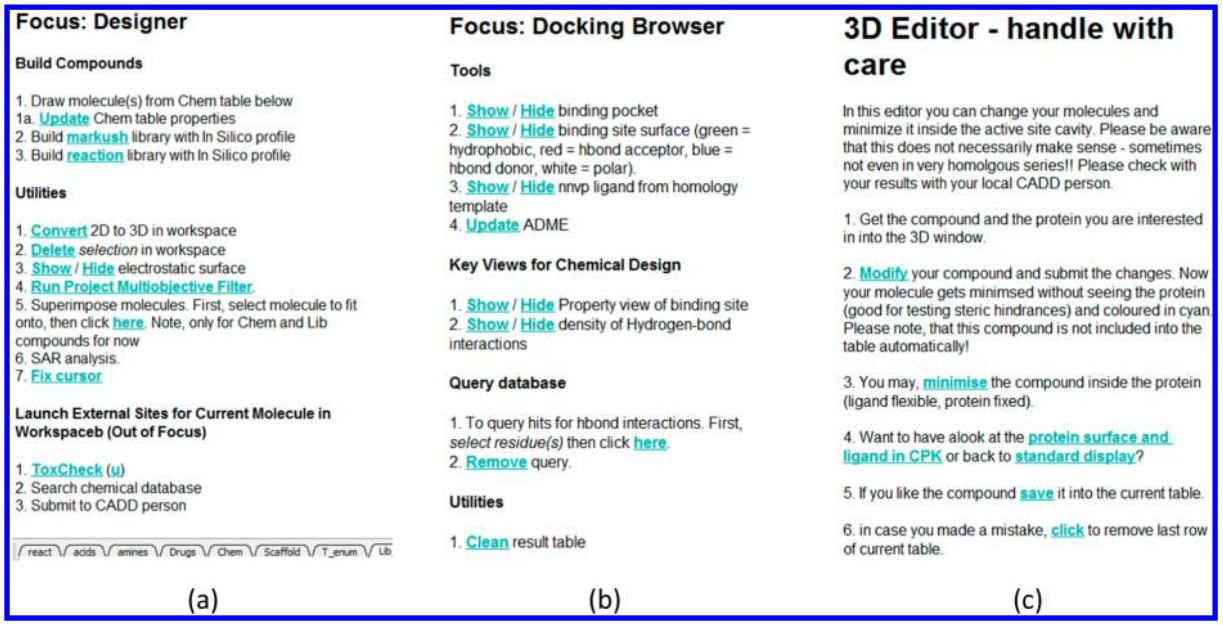


Figure 6. Typical FOCUS sessions from the early development stages. Action links had to be manually modified and could not be transferred between projects without changing the underlying code. (a) Very early small molecule design session for library enumeration, property calculation, and alignment of results. (b) Docking browser to analyze results provided by modeler. (c) Editor for compounds inside the binding pocket.

fast scoring approximations, interface design, and library extensions were necessary.

RESULTS AND DISCUSSION

The development process of FOCUS can be subdivided into four major stages. Depending on the respective time frame, different use patterns were obtained and hence different conclusions were drawn at each stage. Here, we want to summarize these findings, provide usage examples, and give guidelines for future developments.

Evaluation, Prototyping, and Initial Rollout. The original ideas behind FOCUS (tighter communication in medicinal chemistry projects especially between modelers and bench chemists; desktop virtual medicinal chemistry and information sharing) were, for the most part, favorably received at NIBR by leadership reviews. Given the strong medicinal and computational chemistry components of the project, the choice was made that FOCUS would be developed by the NIBR global chemistry community, after consultation with our research IT department. At NIBR, we decided that it was not viable to build a system with the scope of FOCUS from scratch but rather buy high-level components and assemble them into the capabilities we wanted. These high-level components for molecules and modeling have become so advanced and specialized that recreating them would have been prohibitively expensive. We therefore chose a commercial modeling application that also provided a scripting environment as a quick and cost-effective way to customize the tool to the needs we had deemed most appropriate.

Initially, different software packages (among others CCG's MOE, Schroedinger's Maestro, OpenEye's VIDA, Biovia's DiscoveryStudio, PYMOL, and Molsoft's ICM) were compared using several criteria: flexibility/robustness of the product, 2D/3D capabilities, interfaces, multiplatform nature, scriptability for both front end and back end, internal expertise, cost, and vendor profile (including support and ability to develop new science). The different vendors had strengths and weaknesses; ICM was chosen as the best balance of these criteria. Prior experience with the development of similar systems to FOCUS by other biotech companies also played a major role in the decision process. We stress that our conclusion should not be seen as a rigorous evaluation but as a decision at a point in time.

The initial team was made up of members with a variety of skills, proficiency in scripting, knowledge of the problem space (drug discovery), and awareness of the local customs and environments of the distributed sites at NIBR. However, in the long run, it became clear to us that not involving an expert software developer as a consultant in the decision process at NIBR resulted in a range of problems during this early phase. The addition of new functionality was sometimes cumbersome because the underlying design model of the software was not known to the internal developers and was not a good fit to the requirements. Another positive aspect we missed by not involving a scientific software development expert was around guidance of the project team members to set up tools like bug tracking systems for the vendor/NIBR collaboration and a thorough documentation review early on in the project.

After an initial learning phase of ICM scripting, prototyping commenced, and we learned what could and could not be achieved without modifications to ICM. It should be highlighted, that especially in this phase, a very close contact with the developers on the vendor side and almost immediate response was of great importance. Given that a substantial part of the internal development happened in Europe and as Molsoft is based on the west coast of the United States the working day did not overlap very much. Specifically, at this early stage, greater colocation or time zone overlap would have helped.

From project initiation up to this stage, expectations in the organization on the prospective functionality of FOCUS could differ. We therefore recruited a broad set of medicinal chemists beyond the early adopters to create and prioritize specific deliverables. This was important for the more agile software development process that we adopted for FOCUS.

In the very early phase of FOCUS, prototypes with embedded functionality were created for testers and stored

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

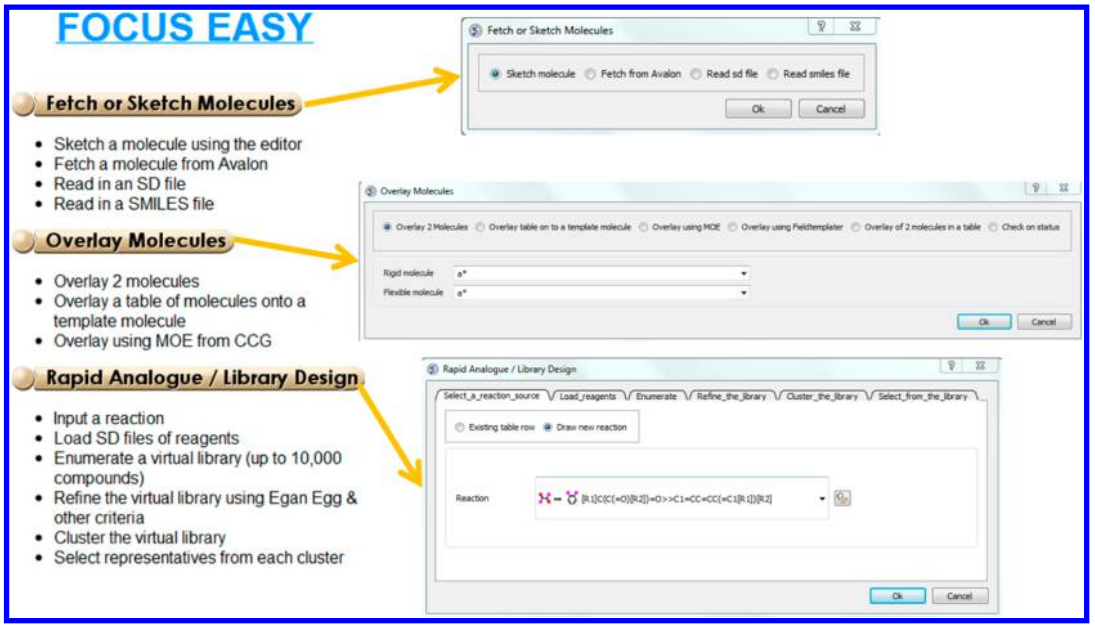


Figure 7. FOCUSEasy interface provided generic functionality bundled into topics like data handling, molecule alignment, and library design. Each "module" provided functionality that could be applied either separately (e.g., "Load reagents") or as whole workflows (e.g., library enumeration and profiling). Later, it became clear that one of the drawbacks of this setup is that infrequent users felt that functions were hidden.

into binary files—so-called FOCUS sessions. Typically, those sessions were highly customized for the respective medicinal chemistry projects and could only be reused to a certain extent. This created a lot of additional work on the side of the modeling community involved at this stage.

Figure 6 shows typical examples of such sessions. It is shown that there is very little user flexibility. New functionality had to be specifically added and was restricted to local tables. For example, in Figure 6a, the superimposition of molecules (Utilities, point 5) can only be applied to tables "Chem" and "Lib". As a consequence, usage patterns of this stage were limited. They mostly comprised calculation of simple property calculations such as MolLogP or PSA as implemented in ICM or small molecule alignments. For web tools that were accessible outside of FOCUS (e.g., ToxCheck in Figure 6a), users were typically sending single compounds. Other workflows used at this stage comprised the analysis of lists from virtual screening with calculated properties and filters specific to the binding pocket in questions (Figure 6b) as well as the rudimentary modification of compounds in binding pockets (Figure 6c).

Sometimes, sessions were sent around within project teams. Because the sessions were static, different versions of the same task could be in the user community. These versions became outdated, and testers sometimes experienced crashes if the underlying software was modified in an inconsistent manner. This was further complicated by the IT procedures that were in place for deploying the ICM binary underlying FOCUS. The slow pace of package update and delivery did not match the rapid changes in development and was another challenge that we faced at this stage. Consequently, a modification to load updated functionality upon startup of the software became a requirement for the next version of FOCUS.

Conclusions from this phase are as follows: (a) Include a wide range of technical experts such as software developers, interface designers, and cheminformaticians, as well as modelers in the evaluation of software solutions. Try to understand the underlying software model and how far the design fits your future needs. Preferably review the source code of your final candidate. (b) Set up all functionality around good software development practices, such as bug tracking systems, assemble user acceptance and test teams, define specific deliverables, and agree on a rollout plan. Take into account local specifics of every site especially with respect to differing IT architectures. If necessary, take into account turnaround times of centralized installation procedures. (c) If your major developers work in different time zones to the software vendor, make sure they get a chance to learn the initial tool in close contact with the vendors. At an early stage, plan for sabbatical/job placements at the vendors' site if possible.

First Rewrite. After the initial problems were solved and we had a more robust version of the software, we decided to start a second development round in 2008. Two of the major aspects of this rewrite were to improve the user experience and to collect and implement the community feedback.

As mentioned, initial prototypes updates were distributed using customized standalone FOCUS sessions. As the added functionality changed continuously, these individual sessions quickly became outdated, and this caused an unacceptable situation with respect to user experience. We therefore developed a mechanism that allowed the loading of all FOCUS functionality from a centralized shared location at the startup of the ICM software. One of the major issues we faced at this stage was the creation of a central repository that could be synchronized across multiple sites due to different architectures of local IT systems. Especially, smaller sites that only recently joined NIBR showed very different directory hierarchies, used different authentication procedures, and were missing the relevant permission groups. Here, close contacts with the local IT support groups became of major importance to the success of this task.

Even though FOCUS was actively supported by global management and presented in various forums and meetings, initial take up by bench and theoretical medicinal chemists at different sites was slower than hoped by the implementation team. Major hurdles seemed to be an overload with new

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

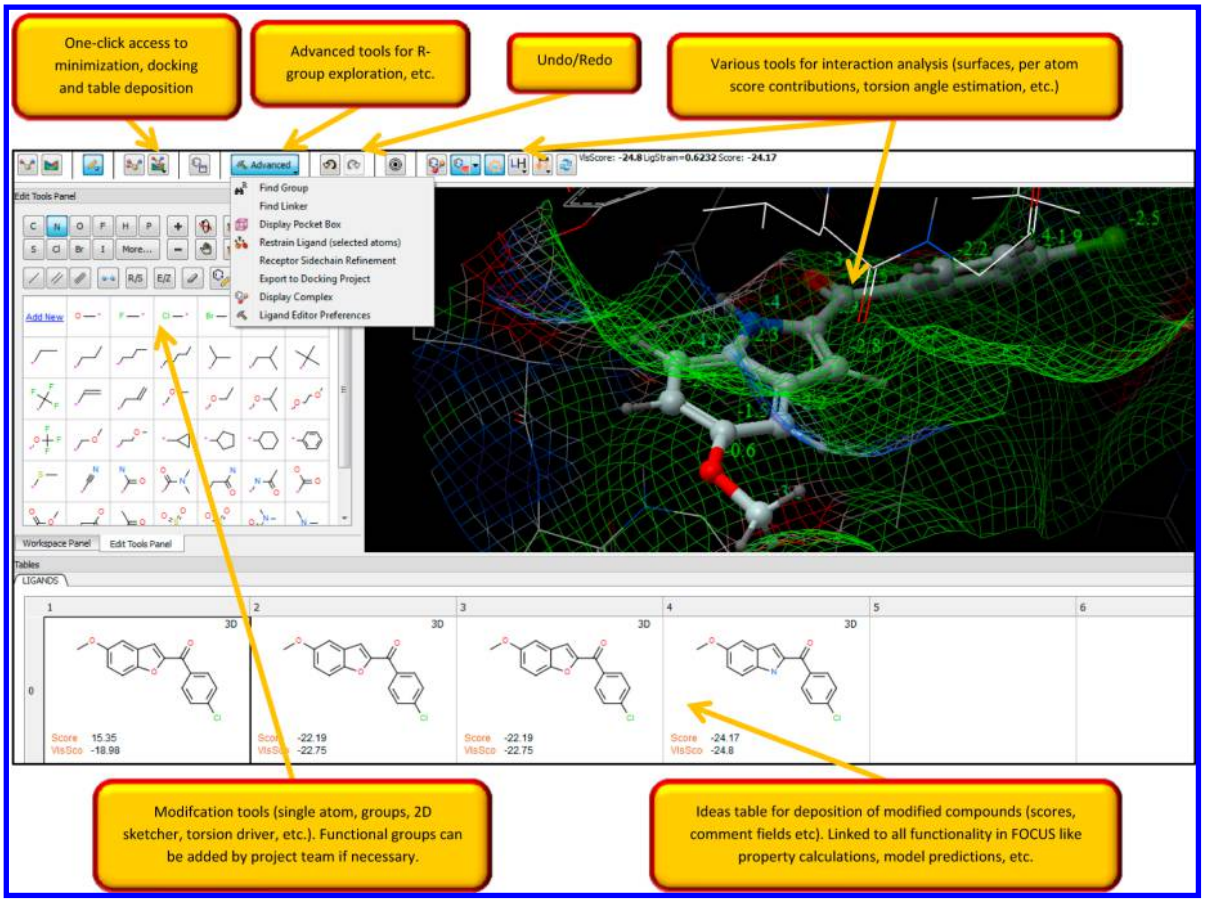


Figure 8. Screenshot of the 3D editor interface in Focus/ICM-Chemist. The two leftmost buttons on the uppermost row are used for setting up editor sessions (protein preparation). The next three buttons are used to minimize or redock compounds and then store in a predefined table. Buttons for undo/redo, hydrogen-bond display, various surface displays, per atom score contribution display, and torsional preferences of the ligand allow the user to visually analyze new ideas. The editor window (left) can be used in 2D and 3D mode and behaves like drawing tools known to chemists to avoid additional training needs.

software that appeared at the same time at NIBR and insufficient time spent on training, as well as little or no experience of how a tool like FOCUS could impact projects and improve the workflows that had been in place for years. Put differently, end users first had to understand the advantages that an idea like FOCUS (some people at NIBR call FOCUS a "lifestyle" change) could offer them. This was especially difficult at some sites because previous attempts to enable chemists to perform desktop modeling had not been successful. We tried to tackle this issue in two ways. First, we identified volunteers in the user community and worked with them to become and act as local knowledge experts. The idea behind this was that information provided by peers is often accepted as more trustworthy and less intimidating than top-down approaches. Second, a lot of time was spent on the training of individual project teams by one of the developers. This usually involved understanding the main needs of that project (e.g., library design, conformation analysis, protein alignments, etc.), and focusing on those aspects during the training. The fact that the project was driven from within the chemistry community and not by the informatics department helped because it was easier to understand the needs described by the users.

The training experience also provided us with two positive side effects. Most users suddenly had a personal connection with one of the developers and felt more comfortable asking questions or bringing up issues with the software. Additionally, a lot of user feedback was given on a one-to-one basis. That way, the most important requirements for the further development process of FOCUS were identified.

With the newly released version, called FOCUSEasy (Figure 7), usage patterns became more diverse. In addition to project specific functionality (e.g., a docking browser) users were now able to apply more standard tasks to tables in any session. This was greatly supported by the introduction of a connectivity of FOCUS to other external services, such as cLogP calculators.

At this stage medicinal chemists started to use FOCUS without direct modeling support; virtual library enumeration and profiling became a task that was most commonly applied. In addition, feedback showed that the possibility to align ligand structures in 3D was of high importance. A typical workflow that was applied frequently by medicinal chemists was the generation of a set of virtual molecules around a common core followed by the calculation of properties for filtering (e.g., cLogP and PSA) and finally the alignment of the filtered molecules to a compound with the same core cocrystallized in the protein pocket. The results from that were then discussed with the project modeler to give a prioritized list of molecules for synthesis.

Another challenge was that properties calculated with various in-house tools (that project scientists knew and trusted) resulted sometimes in (slightly) different reported values to those generated by the web service tools. For many users, this

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

was unexpected and education became important. For example, PSA as implemented in ICM correlated with an R^2 of ~0.99 with our in-house calculator based on tPSA.[31] However, there was an offset of roughly 20 A^2 on the absolute values. Sometimes this discrepancy resulted in confusion and clearly highlighted that centralized services would help—not only for FOCUS but for all tools in the NIBR environment. Until those were available, we were in the fortunate situation that we could switch some calculators to the SOAP web services provided for other workflow tools. Most other menu items and buttons were taken off the FOCUS GUI in order to avoid confusion (e.g., cLogP, number of hydrogen bond donors, etc.). The positive side effect of this experience also highlighted the need to train users about the expected accuracy of tools like cLogP and the changes that are to be expected between model versions.

3D Molecular Designer and Editor. One of the most important requirements from the initial concept of FOCUS was to have a tool to enable medicinal chemists to work directly in 3D and perform their own interactive structure-based design (e.g, the modification of known compounds inside the binding pocket of a target protein). On the basis of our experiences from the initial prototype (Figure 6c), a prototype workflow was created starting from a compound in a binding pocket that could be modified and then optimized or redocked using ICM's docking technology. This prototype helped refine our requirements, working closely with the vendor to develop a scientifically sound and robust tool. After multiple iterations of optimization of the tool, it became clear that to create a method that is both user-friendly and fast enough to be interactive on low-end computers we needed to back port the initial method into the ICM environment—the "3D editor" project. A combined team of Molsoft and NIBR scientists worked on this development project based on the ICM molecular mechanics, 3D chemical spreadsheet, and global optimization/docking technology[32,33] with special attention given to the simplicity of the user interface (Figure 8). The full undo/redo functionality was also implemented to make the environment more user-friendly.

The workflow can be described in three steps. First, expert modelers prepare the sessions (protein setup, ligand identification, and energy map definitions). Internal analysis showed that given the more forgiving nature of the sampling procedure implemented proper setup is of great importance to create reasonable results. End users can then modify the ligand and minimize or redock inside the binding pocket. Some do this at the terminal together with a computational chemist, while others start this way and increasingly over time begin to also work on their own. Some end users work entirely independently. Once the users have analyzed the result (H-bonds, per atom docking score contribution, torsion angles, etc.), they store interesting compounds in the chemical spreadsheet. Finally, modelers and chemists analyze and prioritize the preferred ideas and their chemical feasibility in detail. The advantage of medicinal chemists working independently is that they can quickly test their hypotheses in 3D in their own time at their own desks. While bench chemists working independently may seem contradictory to the idea of improved communication, it can actually lead to improved collaboration between the two groups. Chemists working hands-on in 3D develop improved intuition for the spatial interactions and constraints of their system, and discussions become richer in good ideas from the start.

Conclusions from this phase are the following: (a) Carefully control, or better avoid, deploying local versions of tools. Removing those across an entire organization can be a major undertaking. When deploying cross-site tools, a close interface with the respective infrastructure groups in IT is of major importance to generate an early deployment plan. (b) Be conscious of local customs and cultures. Word of mouth instead of a top down approach to spread information can be a very convincing argument for usage. In addition, it reduces the risk of cultural differences in communication. (c) Preparing training sessions tailored to individual project teams can result in a better uptake of the tool. Make sure the feedback from these sessions goes back to the developers, especially if they are not actively working in medicinal chemistry projects. (d) Check property calculators in vendor tools against internal web pages, and clearly annotate the differences. Remove or hide everything that provides differing results to avoid confusion. (e) The possibility to back port complex scientific tasks in collaboration with the software vendor can be of high importance. Make sure these things are possible and supported without major costs.

Second Rewrite. In 2009, the informatics landscape in NIBR started to change. Especially due to the introduction of the Cheminformatics (CIx) Framework initiative[34] at NIBR, we were now able to interface with many more tools (descriptor and property calculations, name to structure conversions, compound lookups, etc. (Figure 9)).

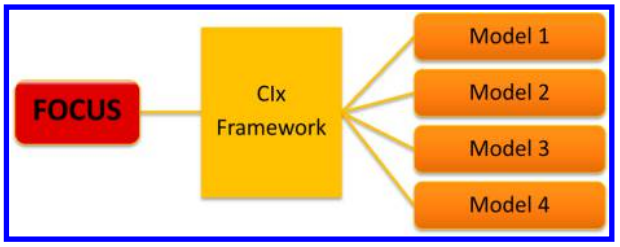


Figure 9. CIx Framework provides a RESTful web service interface as a unified access point to cheminformatics and computational services. This separates specialized software or data sources from end user tools. A single interface facilitates integration of new functionality and greatly reduces the maintenance burden.

Up to this point, most of the code base was implemented with 50% time of a local modeler. However, with many more requests for functionality and in order to be able to make use of the CIx Framework (refactoring the web service integration), the FOCUS team was restructured by management with a new project leader who could codevelop and resource the project to critical mass, which helped FOCUS to move from a mostly development phase to maintenance mode.

An unfortunate consequence of our efforts to improve FOCUS capability was that user feedback now stated that FOCUS had become too complex and contained too much functionality. This was especially true because a lot of functionality was hidden inside GUI windows of workflows (Figure 7). Users had to first find the function they needed by testing various interfaces. We approached the problem in multiple ways. First, the rearranging of closely related functionality into so-called modules (Figure 3) helped the end user find what they were searching for more easily. The usage of dynamically created HTML pages provided a familiar interface, and for infrequent users, a short help text was provided along with every function. Second, we moved many small and specialized tools used mainly by specialists such as

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

modelers, structural biologists, or analytical scientists, from the main user interface into specialized modules that could be enabled on demand. The basic infrastructure and GUI concept introduced at this point proved successful and remained unchanged since 2009 with the exception of small cosmetic changes. The necessary review and cleanup of the code base was also used to move the code into a central code repository that is now used for an automated global deployment process using Subversion and Jenkins. As part of the review process, we also worked with the software vendor to support new objects in the ICM scripting language to simplify code and make it more efficient. Also, at this stage we added code in every function to log usage and send the statistics to a central server. These were added to prepare for maintenance mode and identify power users and infrequently used functionality (Figure 10)

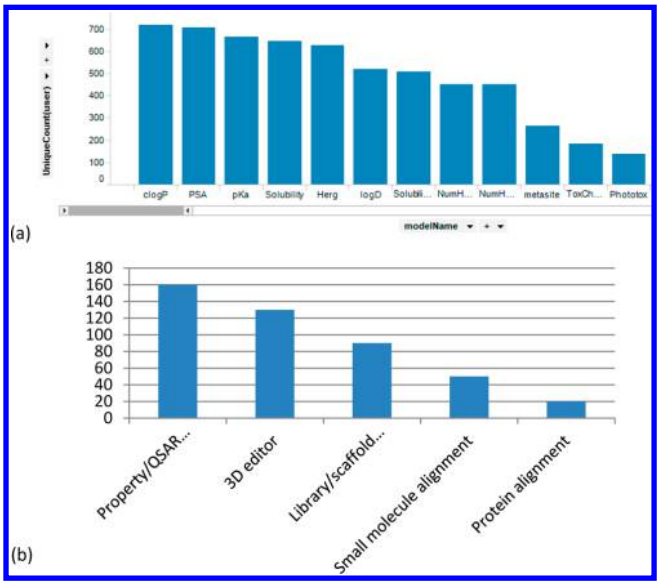


Figure 10. (a) Usage statistics of the models called from FOCUS per unique user during 2014. It is shown that, for example, Metasite functionality was called comparatively infrequently. This information is used to handle licenses or prioritize functionality in the interface. (b) Average unique users in a 3 month period summarized by task (top 5). It is shown that apart from property calculations the 3D editor is the most frequently used tool.

On the basis of these usage statistics, common workflows are now easier to identify. Apart from the more usual ones like library enumeration followed by property calculations and filtering or 3D editing of a set of analogues in a binding pocket, more extensive workflows (Figure 11) can be identified.

We also worked on specific documentation for the functionality available in FOCUS. For this, we had to revise our initial approach to provide simple and short documentation because we realized that many things obvious to one type of chemists are not understandable to the other. Examples for this are library enumeration tasks or small molecule alignment. Hence, we created a complete set of documentation around specific tasks, modules, and typical problems appearing in medicinal chemistry projects, as well as frequently asked questions. Still, some of the feedback we received around training—mostly with respect to frequency of training courses—was not satisfactory. Fortunately, we were able to use a professional writer to create training documents and also have the training prepared and performed by the scientific training group in the NIBR Informatics department.

Conclusions from this phase are the following: (a) Changes in the corporate environment make review of tool usage necessary. Management needs to take these reviews into account because additional resources might be necessary even at later stages of a software project. (b) Tools tend to become too complex and contain too much functionality. A modular setup can help reduce this problem by hiding more specialized functionality; however, a critical review is necessary from time to time. (c) Providing a simple interface that is easy to remember for infrequent users is a critical problem to solve. Provision of workflow style interfaces helps. (d) Changing the structure of a long-term project team, especially at the project leader level, can result in more effective implementations. Make sure the leader is technically involved especially if the change happens in later project stages. (e) Documentation and training are very important. They cannot be done as an afterthought. Plan for professional support in these fields. If not available, plan the resources in the team to provide this support.

From Development to Maintenance. Today, there are about 1500 registered FOCUS users. This is by far many more users than the NIBR global chemistry community. Many users come from other scientific functions such as structural biologists, bioinformaticians, and screeners. While this number may sound impressive at first, it is more important to look at active users. Figure 12 shows the number of users during each week over the last five years. Since 2009, the user community slowly increased from 150 to over 200 unique users per week (a large portion of the medicinal chemistry community; Figure 12). It is also worthy to note that 50% of these users access FOCUS on two or more days per week. This clearly reflects that FOCUS has become an integral part of the average day workflow for many medicinal chemists at NIBR. Based on nonuser feedback, those that do not work with FOCUS tend to either rely fully on modelers to do *in-silico* work, use other interfaces that make use of the same web services to perform 2D work (mostly Spotfire[2] or Stardrop[5]), or only visually inspect protein structures with the help of, for example, PYMOL[1].

FOCUS is a mature application at NIBR. In recent times, we have mostly worked together with Molsoft to implement new functionality into the scripting language and much less on the actual functionality of FOCUS. In contrast to the early days of the FOCUS implementation, at this later stage, the time difference between the sites could be very useful rather than a hindrance. After suggesting new functionality, the time shift allowed both ends to think and specify potential implementation details. Recently, other aspects of the ICM suite to further improve communication in multiple ways were evaluated. Here, the ActiveICM plugin[35,36] for web pages or PowerPoint and the iMolView[37] application for mobile devices could be extended to FOCUS versions of the tools.

A very good indication of positive user acceptance at this stage was also the frequency with which end users identified FOCUS as a preferred platform to deploy new, specialized workflows. For example, the peptide group requested library enumeration tools to work with macrocycles. In addition, usage statistics showed that the amount of data calculated for QSAR/property models increased dramatically. On the basis of that, changes to the web service code had to be implemented to cope with large data submissions (e.g., subset submission, autoupdate missing data only).

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

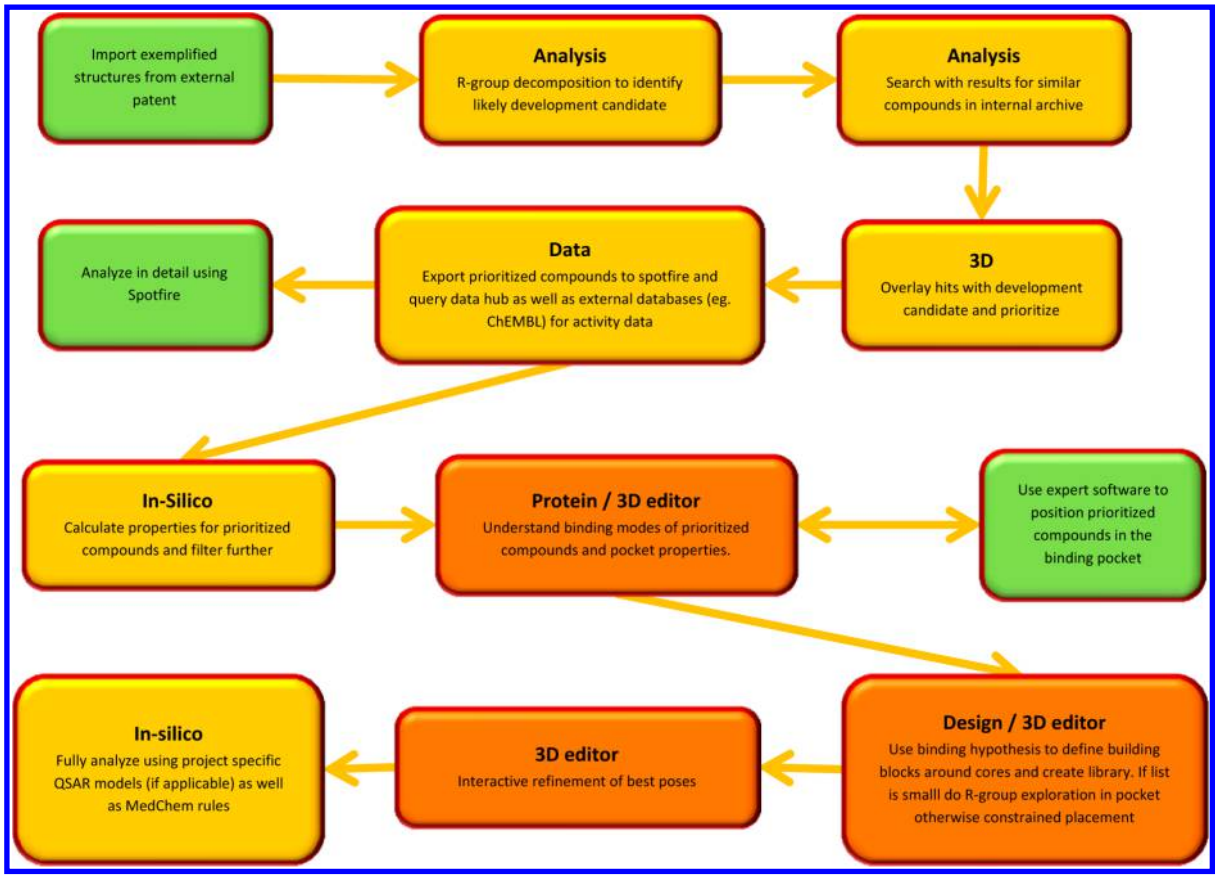


Figure 11. More extensive example workflow going from patent information to novel compounds for synthesis. FOCUS modules used are highlighted in bold. Green fields highlight interactions with outside software (e.g., Spotfire, SciFinder, etc.). Yellow fields highlight tasks that can be performed by medicinal chemists inside FOCUS. Orange fields highlight tasks where medicinal chemists and modelers work together interactively in FOCUS.

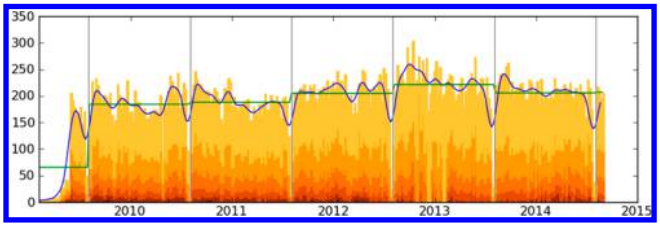


Figure 12. Number of unique users per week since start of the maintenance phase. The *Y*-axis shows the number of unique users per week. Color shading groups users by usage frequency per week, from 1 day in a week (light yellow) up to 5 or more days in a week (brown). The blue line is a running average, and the green line the yearly average.

Currently, FOCUS sessions are an integral part of the design and communication strategies of most medicinal chemistry projects. Those sessions are usually shared between all team members in specific FOCUS folders either on SharePoint project pages, shared project drives, or central locations on Linux folders. Information from FOCUS sessions can be exported directly into other commonly used software tools like Spotfire or intranet web pages. If a direct link to another software is not available, exports to data using common file format such as csv or sdf files is another possibility to transfer information between FOCUS and other tools.

Conclusions from this phase are as follows: (a) Continue to grow the capabilities as new scientific methods become available and workflow expectations change, but beware of complexity that will overburden the user experience. Make existing functionality better over time. (b) Improvement to the software can be invisible to the end user; prepare the ground for future functionality early on. Modular setup and good software management are necessary. (c) Analyze usage patterns with the goal to make the application more robust and easier to maintain. (d) Include solutions for agile deployment of new science to the end users to stay on top of evolving technologies. Once the experimental methods have proved their worth, they can be more easily transitioned into the application.

■ CONCLUSIONS

We have presented the development of FOCUS, a platform to produce, visualize, and share information on all aspects of a medicinal chemistry project. A hybrid approach was used, combining internally developed libraries on top of specialized vendor software for cheminformatics, 3D graphics, modeling, and structural design. This allows us to leverage the external support for the underlying code base but still provide improvements quickly using agile development practices. FOCUS was created and evolved over roughly eight years. Reoccurring themes during that period are that only continuous interaction with the user community, training, software review, and adaptation of the tool results in acceptance from the user community. The quick turnaround from user requests for new features to actual deployment (generally hours to days) distinguishes the hybrid approach used from vendor solutions. This ability to quickly respond to user needs without falling

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

into the trap of rushed actions and adding new interfaces to available services is of high importance. Only if this point is kept in mind can such a strategy be useful in the ever-changing environment of a large, complex research organization.

The key challenge we faced was introducing a new way of facilitating drug discovery via FOCUS into a diverse global multicultural environment, outside of a traditional IT development process. Adoption of new paradigms in large organizations is often difficult due to different local customs and cultures, management expectations, and IT practices and requires commitment from a dedicated team. We relied on early adopters not only to help test and guide development but to educate and spread the word about the capabilities in FOCUS for the larger community. Constant feedback and experience from the users resulted in a rapid evolution of FOCUS over time, although this pace was a challenge on a global scale.

Still, one of the most important messages is that even when using a hybrid approach, like we did, the creation and acceptance of such a tool is only possible if supported over a long time by a dedicated team and that the overall cost and effort is substantial. However, if accepted, it acts not just as another tool but facilitates drug discovery through communication and cooperation between the theoretical and applied medicinal chemists. Computational methods that were hard to access before are now more easily available to nonexperts. This experience ultimately broadens the scientists' thinking (be it 2D or 3D) and strengthens the knowledge they already have.

Today, aside from pure visualization and communication of ideas, the most common usage patterns observed for medicinal chemists are property calculations, 3D editing of a ligand in a binding pocket, library enumerations, and small molecule alignments. It is important to note that those are frequently combined into workflows—some of them being rather complex and applied to topics such as patent analysis, scaffold hopping or public data analysis. In addition, specialized teams have developed customized workflows that are not currently covered by other tools available to them.

Finally a word of caution—expectations of usage levels should be kept at a reasonable level. Even with a hypothetically perfect tool, one cannot expect 100% usage. Individual scientists will still work best if they are allowed to do their research in their style. Some will prefer a 2D rendering, while others enjoy working with more complex 3D structures. For this, it is important to design the system to be flexible enough to accommodate those individual styles. Our experience is that providing tools like FOCUS can help drug discovery scientists to explore and reevaluate new ways in how to achieve their goals.

■ ASSOCIATED CONTENT

Supporting Information

A full listing of the FOCUS modules and a description of the methods and information on where they are executed. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: nikolaus.stiefl@novartis.com.

Present Addresses

Peter Hunt: Optibrium Ltd., 7221 Cambridge Research Park, Beach Drive, Cambridge CB25 9TL, United Kingdom

Katrin Spiegel: 3 Impasse Franchemont, 75011 Paris, France

Scott Biller: Agios Pharmaceuticals, 38 Sidney Street, Cambridge, Massachusetts 02139, United States

Christoph Buenemann: Baxter IT BioScience Global LIMS, Donau-City-Straße 7, DC Tower, A-1220 Wien, Austria

Takanori Kanazawa: Eli Lilly Japan K.K., Sannomiya Plaza Bldg., 7-1-5, Isogamidori, Chuo-ku, Kobe 651-0086, Japan

Ruben Tommasi: AstraZeneca, 35 Gatehouse Drive, Waltham, Massachusetts 02451, United States

John van Drie: Van Drie Research LLC, 109 Millpond, North Andover, Massachusetts 01845, United States

Author Contributions

We want to emphasize that FOCUS was developed as a team effort over a long period. During that time, team members at NIBR and Molsoft made key contributions as needed. Additionally, Eric Martin, who was not part of the core FOCUS team, acted as a consultant in the design of the 3D editor project. Most of the FOCUS code base was implemented by Peter Gedeck and Nikolaus Stiefl. Donovan Chin initiated the project and created the first prototype of the 3D editor. Ruben Abagyan is the ICM project coordinator and lead architect responsible for strategic developments at Molsoft. Brian Vash created the interface to the in-house protein structural database. John van Drie and Scott Biller provided a lot of input on the initial architecture and strategy of the project. In addition, Scott Biller and Ruben Tommasi were instrumental in the strategic adoption of FOCUS as a front end to parallel synthesis.

Notes

The authors declare the following competing financial interest(s): Eugene Raush, Max Totrov, and Ruben Abagyan have significant financial interest in Molsoft or its products.

■ ACKNOWLEDGMENTS

The FOCUS team thanks all early adopters and supporters, specifically Kenji Namoto, David Wayne Hughes, Paul Butler, Gregory Landrum, Sandra Mueller, Joerg Muehlbacher, Finton Sirockin, Simona Cotesta, Marc Litherland, Rajeshri Karki, Yipin Lu, Robert Pulz, Catherine Leblanc, Sabina Pecchi, Gabriel Gamber, Jay Knowles, Michael Dechantstreiter, and Thomas Dice. We also thank Jose Duca for supporting the project during the last three years. On the Molsoft side, we thank Andrew Orry, Crystal Crawford, Levon Budagyan, Polo Lam, Irina Kufareva, and Elena Arnautova for many contributions.

■ REFERENCES

(1) PyMOL Molecular Graphics System, version 1.5.0.4 Schrödinger, LLC, 2013.

(2) Spotfire, Tibco Spotfire. http://spotfire.tibco.com/ (accessed September 23, 2014).

(3) *Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group, Inc., Montreal, QC, Canada, 2014.

(4) Maestro, Schrödinger, LLC. http://www.schrodinger.com/Maestro/ (accessed September 23, 2014).

(5) Stardrop, Optibrium, Ltd. http://www.optibrium.com/stardrop/stardrop-features.php (accessed February 28, 2015).

(6) PerkinElmer, Chemdraw. http://www.cambridgesoft.com/Ensemble_for_Chemistry/ChemDraw/ (accessed September 23, 2014).

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908

(7) Instant JChem 14.7.7, ChemAxon. http://www.chemaxon.com (accessed 2014).
(8) Sander, T.; Freyss, J.; von Korff, M.; Reich, J. R.; Rufener, C. OSIRIS, an Entirely in-House Developed Drug Discovery Informatics System. *J. Chem. Inf. Model.* **2009**, *49*, 232−246.
(9) Agrafiotis, A. K.; Alex, S.; Dai, H.; Derkinderen, A.; Farnum, M.; Gates, P.; Izrailev, S.; Jaeger, E. P.; Konstant, P.; Leung, A.; Lobanov, V. S.; Marichal, P.; Martin, D.; Rassokhin, D. N.; Shemanarev, M.; Skalkin, A.; Stong, J.; Tabruyn, T.; Vermeiren, M.; Wan, J.; Xu, X. Y.; Yao, X. Advanced Biological and Chemical Discovery (ABCD): Centralizing Discovery Knowledge in an Inherently Decentralized World. *J. Chem. Inf. Model.* **2007**, *47*, 1999−2014.
(10) Certara. D360: Bring Your Data to Life. http://www.certara.com/products/sci-info/d360 (accessed September 23, 2014).
(11) Biovia. Virtual Biosphere and Materials. http://www.3ds.com/products-services/biovia (accessed September 23, 2014).
(12) OpenEye. VIDA. http://www.eyesopen.com/vida (accessed September 23, 2014).
(13) Gerebtzoff, G. LUCID: Supervised Multidimensional Optimization of Compounds Using Matched Molecular Pairs, RDKit User Group Meeting, Cambridge, UK, 2013. https://github.com/rdkit/UGM_2013/blob/master/Presentations/Gerebtzoff.LUCID.pdf (accessed 2015).
(14) Chin, D. N.; Chuaqui, C.; Singh, J. Integration of Virtual Screening into the Drug Discovery Process. *Mini-Rev. Med. Chem.* **2004**, *4*, 1053−1065.
(15) Molsoft. ICM Chemist Pro. http://www.molsoft.com/icm-chemist-pro.html (accessed September 23, 2014).
(16) Landrum, G. A.; Stiefl, N. Is That a Scientific Publication or an Advertisement? Reproducibility, Source Code and Data in the Computational Chemistry Literature. *Future Med. Chem.* **2012**, *4*, 1885−1887.
(17) Guha, R. Papers About Systems You Can't Use or Buy, February 1, 2009. http://rguha.wordpress.com/2009/02/01/papers-about-systems-you-cant-use-or-buy/ (accessed 2105).
(18) Ritchie, T. J.; Lewis, R.; Ertl, P. The Graphical Representation of ADME-Related Molecule Properties for Medicinal Chemists. *Drug Discovery Today* **2011**, *16*, 65−72.
(19) PDB ID: 3LFB, 3LFC. Getlik, M.; Gruetter, C.; Simard, J. R.; Van Otterlo, W.; Robubi, A.; Aust, B.; Rauh, D. Development of Novel Thiazole-Urea Compounds Which Stabilize the Inactive Conformation of p38 Alpha. Paper in process. DOI: 10.2210/pdb3lfb/pdb and http://www.rcsb.org/pdb/explore/explore.do?structureId=3lfb; DOI: 10.2210/pdb3lfc/pdb and http://www.rcsb.org/pdb/explore/explore.do?structureId=3lfc.
(20) Totrov, M.; Abagyan, R. Protein−Ligand Docking as an Energy Optimization Problem. In *Drug-Receptor Thermodynamics: Introduction and Applications*, Raffa, R. B., Ed.; Wiley: New York, 2001, pp 603−624.
(21) Fernandez-Recio, J.; Totrov, M.; Abagyan, R. Soft Protein−Protein Docking in Internal Coordinates. *Protein Sci.* **2002**, *11*, 280−291.
(22) Fernandez-Recio, J.; Totrov, M.; Abagyan, R. Screened Charge Electrostatic Model in Protein−Protein Docking Simulations. *Pac. Symp. Biocomput.* **2002**, 552−563.
(23) Totrov, M.; Abagyan, R. Rapid Boundary Element Solvation Electrostatics Calculations in Folding Simulations: Successful Folding of a 23-Residue Peptide. *Biopolymers* **2001**, *60*, 124−133.
(24) Abagyan, R.; Totrov, M. Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *J. Mol. Biol.* **1994**, *235*, 983−1002.
(25) Bottegoni, G.; Kufareva, I.; Totrov, M.; Abagyan, R. Four-Dimensional Docking: A Fast and Accurate Account of Discrete Receptor Flexibility in Ligand Docking. *J. Med. Chem.* **2009**, *52*, 397−406.
(26) Schapira, M.; Totrov, M.; Abagyan, R. Prediction of the Binding Energy for Small Molecules, Peptides and Proteins. *J. Mol. Recognit.* **1999**, *12*, 177−190.
(27) PDB ID: 3AD4. Ozawa, M.; Ozawa, T.; Tsuji, E.; Okazaki, K.; Takeda, K. Ab Initio Fragment Molecular Orbital Study of Ligand Binding to Leukocyte-Specific Protein Tyrosine (LCK) Kinase. Paper in process. DOI: 10.2210/pdb3ad4/pdb and http://www.rcsb.org/pdb/explore/explore.do?structureId=3ad4.
(28) Lee, B.; Richards, F. M. The Interpretation of Protein Structures: Estimation of Static Accessibility. *J. Mol. Biol.* **1971**, *55*, 379−400.
(29) Apache Subversion, Subversion. http://subversion.apache.org/ (accessed September 23, 2014).
(30) Jenkins − An Extendable Open Source Continuous Integration Server. http://jenkins-ci.org/ (accessed September 23, 2014).
(31) Ertl, P.; Rohde, B.; Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *J. Med. Chem.* **2000**, *43*, 3714−3717.
(32) Abagyan, R. A.; Totrov, M. M.; Kuznetsov, D. N. ICM − A New Method for Protein Modelling and Design. Applications to Docking and Structure Prediction From the Distorted Native Conformation. *J. Comput. Chem.* **1994**, *15*, 488−506.
(33) Totrov, M.; Abagyan, R. Flexible Protein−Ligand Docking by Global Energy Optimization in Internal Coordinates. *Proteins* **1997**, *Suppl 1*, 215−220.
(34) Landrum G. A. Open-source from/in the enterprise: The RDKit. Slide 27/33 from http://www.slideshare.net/GregLandrum1/rd-kitopen-sourceintheenterprise (accessed 2015).
(35) Abagyan, R.; Lee, W. H.; Raush, E.; Budagyan, L.; Totrov, M.; Sundstrom, M.; Marsden, B. D. Disseminating Structural Genomics Data to the Public: From a Data Dump to an Animated Story. *Trends Biochem. Sci.* **2006**, *31*, 76−78.
(36) Raush, E.; Totrov, M.; Marsden, B. D.; Abagyan, R. A New Method for Publishing Three-Dimensional Content. *PLoS One* **2009**, *4*, e7394.
(37) Yiu, C. P.; Chen, Y. W. High-Quality Macromolecular Graphics on Mobile Devices: A Quick Starter's Guide. *Methods Mol. Biol.* **2014**, *1091*, 343−352.

DOI: 10.1021/ci500598e
J. Chem. Inf. Model. 2015, 55, 896−908