JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics

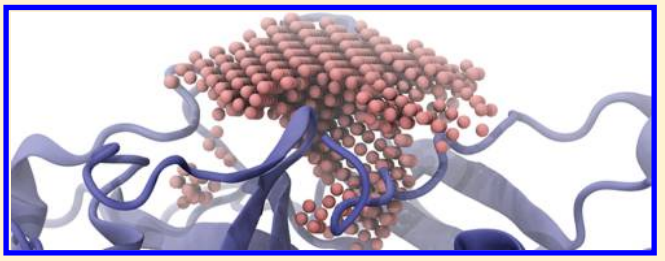
Jacob D. Durrant,[†,‡] Lane Votapka,[†] Jesper Sørensen,[†] and Rommie E. Amaro*[,†,‡]

[†]Department of Chemistry & Biochemistry, University of California San Diego, La Jolla, California 92093, United States

[‡]National Biomedical Computation Resource, Center for Research in Biological Systems, University of California San Diego, La Jolla, California 92093, United States

Supporting Information

ABSTRACT: Analysis of macromolecular/small-molecule binding pockets can provide important insights into molecular recognition and receptor dynamics. Since its release in 2011, the POVME (POcket Volume MEasurer) algorithm has been widely adopted as a simple-to-use tool for measuring and characterizing pocket volumes and shapes. We here present POVME 2.0, which is an order of magnitude faster, has improved accuracy, includes a graphical user interface, and can produce volumetric density maps for improved pocket analysis. To demonstrate the utility of the algorithm, we use it to analyze the binding pocket of RNA editing ligase 1 from the unicellular parasite *Trypanosoma brucei*, the etiological agent of African sleeping sickness. The POVME analysis characterizes the full dynamics of a potentially druggable transient binding pocket and so may guide future antitrypanosomal drug-discovery efforts. We are hopeful that this new version will be a useful tool for the computational- and medicinal-chemist community.



INTRODUCTION

Binding-pocket analysis is an active area of research that includes pocket detection and characterization, druggability prediction, and the study of binding-site flexibility.[1] The advent of the Protein Data Bank (PDB[2]) spurred the creation of a number of software packages aimed at facilitating the analysis of macromolecular pockets.[3−6] In recent years, additional programs have been developed with improved accuracy and increasingly advanced pocket-characterization algorithms,[7−16] as reviewed by Zheng et al.[17]

Pocket analysis is useful for studying receptor dynamics.[18−40] One can get a good sense of the full gamut of possible binding-pocket conformational states by obtaining multiple structures from X-ray crystallography, NMR spectroscopy, or molecular dynamics (MD) simulations and comparing pocket volumes and, in particular, shapes. These comparisons facilitate the identification of novel, pharmacologically relevant binding-pocket conformations, as well as transient binding pockets that are not evident when a limited number of static structures are considered.

Additionally, pocket analysis can also be applied to computer-aided drug discovery (CADD). Among the many complex factors that govern molecular recognition,[41,42] pocket volume and shape are perhaps the most straightforward. Simply put, a ligand will not generally bind to a receptor if it cannot physically fit within the confines of the binding pocket, and receptor/ligand shape complementarity plays a key role in molecular recognition.[43] Consequently, pocket characterization has been used to inform CADD efforts aimed at predicting ligand binding, whether through virtual screening, QSAR, or volumetric similarity searching.[44−46] Given the astounding variety of pocket geometries possible,[47] this characterization is no trivial task.

To address this challenge, both ligand- and receptor-centric approaches have been developed. Ligand-based methods such as OpenEye's Rapid Overlay of Chemical Structures (ROCS) algorithm[48] seek to identify novel small-molecule binders by querying a compound database for entries with three-dimensional shapes that are similar to that of a known template ligand,[49] as assessed by the degree of volume-overlap mismatch. These techniques perform comparably to more traditional virtual-screening methods[50] and have been used to successfully identify a number of experimentally validated ligands (see, for example, refs 51−53).

While ligand-based approaches will certainly continue to have high utility,[54] a more receptor-centric methodology is sometimes advantageous. Bound ligands often occupy only a portion of their respective pockets,[43,47] on average perhaps as little as a third of the total space available.[43] Analysis of ligand volume and shape alone cannot account for potential interactions with pocket regions that are not occupied by the template ligand itself. In contrast, receptor-based pocket analysis elucidates the volume and shape of the entire cavity, including regions not yet exploited by existing pharmacophores.

Receptor-centric techniques can also be used to select diverse pocket shapes for use in subsequent virtual-screening efforts. It is often helpful to dock a library of small molecules into

Received: May 1, 2014
Published: September 29, 2014

   dx.doi.org/10.1021/ct500381c | J. Chem. Theory Comput. 2014, 10, 5047−5056

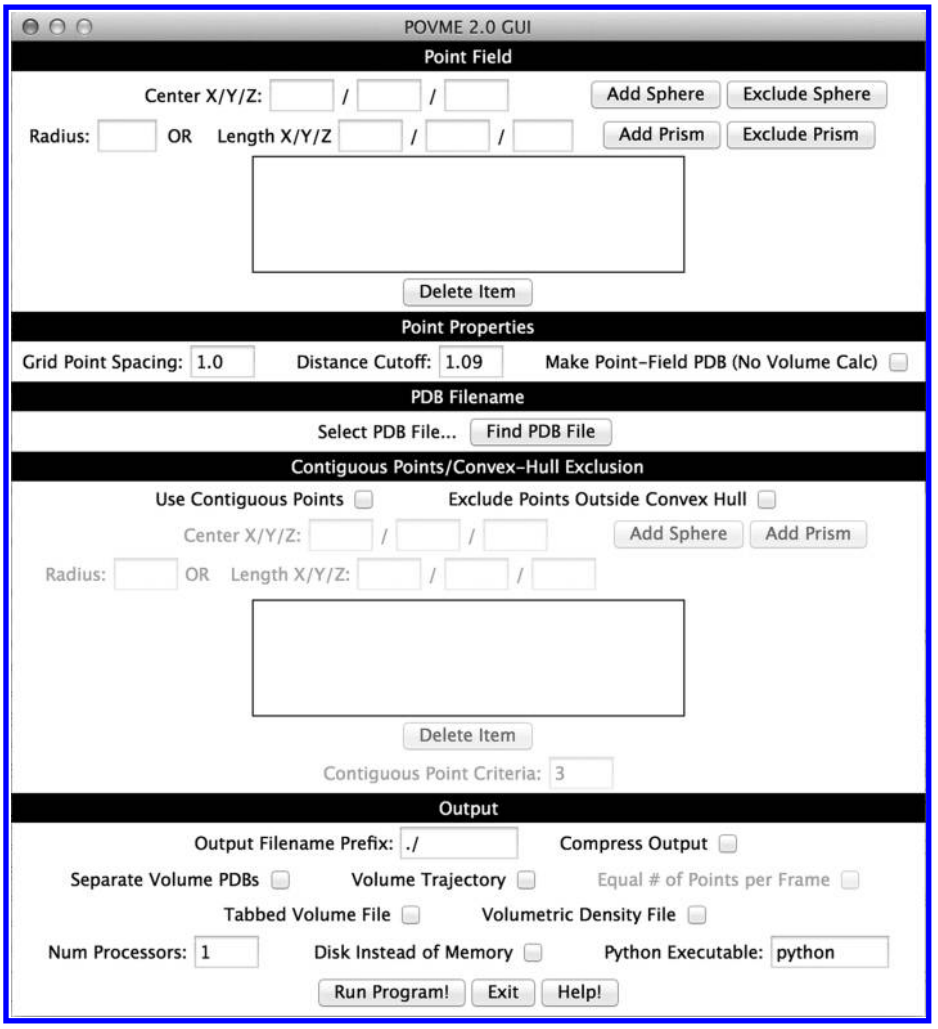


Figure 1. POVME 2.0 graphical user interface.

multiple receptor conformations in order to account for receptor flexibility. Carefully selecting conformations with unique pocket geometries has been shown to enhance hit rates and subsequent ligand diversity.[55−58]

To simplify binding-pocket characterization, we recently developed an algorithm called POVME (POcket Volume MEasurer).[7] POVME floods a pocket-encompassing region with equidistant points, removes those points that are near receptor atoms, and calculates the volume from the remaining points. The points can themselves be saved, providing a specific description of the pocket shape as well. Inspired by the fairly widespread adoption of our program (at least 43 citations as of June 2014), we have now created a second, much improved version. POVME 2.0 is over an order of magnitude faster than POVME 1.0, includes a graphical user interface (Figure 1) that greatly improves usability, can calculate volumetric density maps to facilitate analysis, and has improved accuracy.

POVME 2.0 has been tested on all major operating systems with various versions of python, *numpy*, and *scipy* (Table 1).[59−63] A copy of the program, which is released under the terms of the GNU General Public License, can be obtained from http://nbcr.ucsd.edu/POVME. We are hopeful that POVME will be a useful tool for the computational- and medicinal-chemist community.

Table 1. Operating-System Compatibility[a]

operating system	python version	*numpy* version	*scipy* version
Scientific Linux 6.2	2.6.6	1.6.2	0.11.0
OS X 10.9.1	2.7.5	1.6.2	0.11.0
Windows 7 Home Premium	2.7.6	1.8.0	0.13.3

[a]POVME 2.0 has been successfully tested on all major operating systems with various versions of python, *numpy*, and *scipy*.

MATERIALS AND METHODS

The POVME Algorithm. Successful POVME use includes three required and two optional steps. Trajectory alignment, the construction of a pocket-encompassing region, and the subsequent identification of the pocket-occupying space are required. Optionally, the user can also instruct POVME to eliminate subregions that fall outside the receptor's convex hull and/or are noncontiguous with the primary pocket. A detailed description of each of these steps follows.

1). Aligning the Trajectory. POVME accepts a multiframe PDB (Protein Data Bank) file as input. We expect that MD simulations will be the most common source of these files, but multiple crystal structures or NMR conformations can also be used. We have found that the computer program Visual Molecular Dynamics (VMD)[64] is useful for aligning trajectories and converting files to the PDB format, but other software

packages can also be used for this purpose. Alignment is necessary because the POVME algorithm assumes the pocket being measured does not translate or rotate in space. Different alignment methodologies can subtly alter how this requirement is met, as discussed in the Results and Discussion. A tutorial showing how to align and convert trajectory files using VMD is included in the Supporting Information (Text S1). We note also that single-frame PDB files can likewise serve as POVME input if the user wishes only to characterize a single pocket.

2). Defining a Region That Encompasses All Trajectory Binding Pockets. The user must next define "inclusion" (Figure 2A) and "exclusion" (Figure 2B) regions, respectively. Both of these regions are constructed from a combination of user-specified spheres and rectangular prisms. The required inclusion region should entirely encompass all the binding-pocket conformations of the trajectory. POVME includes a helper script called "POVME Pocket ID," described in greater detail below, to assist in identifying this region if necessary. The optional exclusion region defines portions of the inclusion region that should be ignored, perhaps because they are not truly associated with the pocket. To generate a grid of equidistant points that encompasses all the binding-pocket conformations of the trajectory (spaced 1.0 Å apart by default), POVME first floods the user-specified inclusion region with points and then removes any points also contained in the optional exclusion region (Figure 2C).

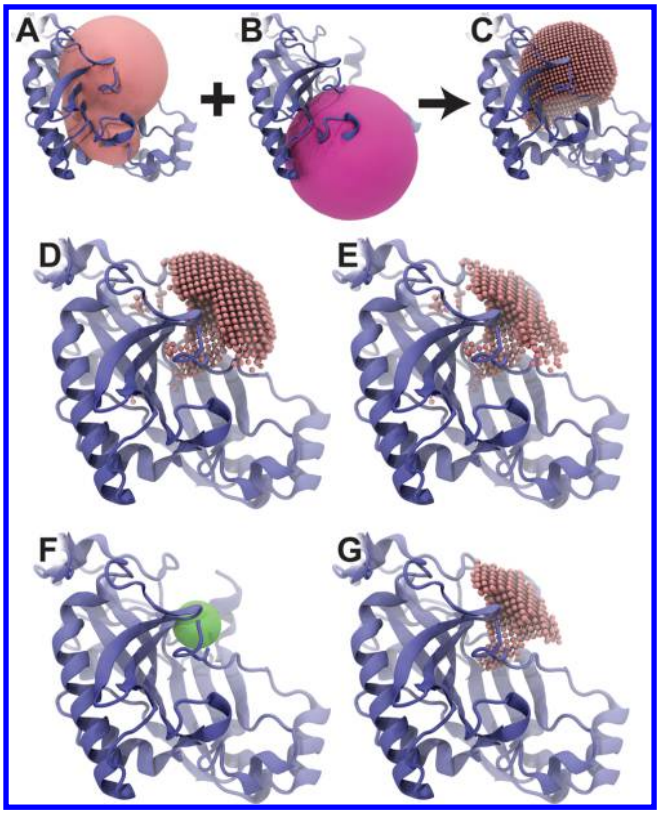


Figure 2. A graphical summary of the POVME 2.0 algorithm. A) The user defines an inclusion region. B) The user defines an exclusion region. C) The portion of the inclusion region that is not also in the exclusion region is flooded with equidistant points. D) Any of the points that are close to receptor atoms are deleted. E) Any points outside the convex hull are optionally deleted. F) The user can optionally define a contiguous-points region. G) All points that are not contiguous with that region are similarly deleted.

3). Removing Points That Are near Receptor Atoms. As the purpose of POVME is to measure the volume of the binding-pocket cavity, the program next removes any points that are close to receptor atoms, leaving only those points that are likely to be located within the binding pocket itself (Figure 2D). Specifically, the pairwise distances between all atoms and POVME points are calculated. Any point that is closer to a given atom than that atom's van der Waals radius, plus a user-specified tolerance (1.09 Å, the radius of a hydrogen atom, by default), is removed. As POVME is written in python, users can easily modify the default radii or add new radius values for novel constituent atom types. If the user wishes to ignore all receptor hydrogen atoms when calculating pocket volumes, the hydrogen radius specified in the script can be set to 0.0 Å.

4). Removing Points Outside the Receptor's Convex Hull. POVME 2.0 introduces an optional new feature for removing points that lie entirely outside the binding pocket. Specifically, the gift-wrapping algorithm is used in combination with the Akl-Toussaint heuristic[65] to define the convex hull of receptor atoms near the user-defined inclusion region. As the gift-wrapping algorithm runs in $O(n^2)$ time, where n is the number of atoms in the receptor structure, it is not necessarily the fastest algorithm for computing the convex hull. However, by coupling it with the Alk-Toussaint heuristic, the expected running time is lowered to $O(n)$. Ultimately, any points that fall outside the convex hull are removed (Figure 2E). This feature is particularly useful when the user defines an inclusion region that protrudes into the surrounding solvent-occupying space.

5). Removing Points That Are Not Contiguous with the Primary Pocket. Like the original POVME program, version 2.0 retains the optional ability to remove isolated patches of points that are not contiguous with the primary binding pocket. This feature requires that the user define a third region, again using spheres and rectangular prisms, that always falls within the primary binding-pocket region regardless of the trajectory frame considered (Figure 2F). All pocket-occupying points within or contiguous to this region are retained, but isolated patches of points that are not directly connected are deleted (Figure 2G).

POVME Output. By default, POVME writes a number of files to the disk. The calculated pocket volumes, as well as user-defined parameters and progress messages, are saved to a simple text-based log file. POVME can also be instructed to save the volume measurements to a second file in a simple tabular format that can be easily pasted into popular spreadsheet programs. Pocket-occupying points are equidistant (1.0 Å by default), so each point is associated with an identical cubical volume (e.g., 1.0 $Å^3$). The volume of a whole pocket is calculated by simply summing the individual volumes associated with each unique point.

POVME also optionally saves the pocket-occupying points of each frame to PDB file(s) on the disk. The user can instruct the program to save these points to separate files and/or to a single PDB trajectory. Some visualization programs (e.g., VMD) are only compatible with trajectories that have the same number of atoms in each frame. POVME can optionally write extra points to the origin (0.0, 0.0, 0.0) on a frame-by-frame basis to satisfy this requirement. As these POVME frames are formatted similarly to those produced by SiteMap, they are also compatible with the pocket-shape volumetric-overlap clustering tools produced by Schrödinger.[13,14,55,56]

Finally, POVME also optionally saves a volumetric density map in the Data Explorer (DX) format, similar to the

MDpocket algorithm.[12] A volumetric density value is associated with each of the pocket-occupying points by calculating the fraction of all trajectory pockets that include the given point. If the density map is displayed as an isosurface, the value of the isosurface expresses the fraction of time (e.g., over the course of the simulation) that the pocket included the displayed volume.

The POVME Pocket ID Algorithm. The POVME distribution file also includes POVME Pocket ID, a simple script that identifies binding pockets and generates appropriate POVME inclusion regions. After loading the heavy atoms from a PDB file, the algorithm identifies pockets by 1) covering the entire protein in a low-resolution 3D grid of equidistant points (spaced 4.0 Å apart by default), 2) removing points that come within a user-specified distance of any protein atom (3.0 Å by default), and 3) removing points that fall outside the convex hull defined by the protein alpha carbons. The remaining points tend to congregate in binding pockets.

These points are then replaced with smaller higher-resolution 3D grids of equidistant points (spaced 1.0 Å apart by default). These higher-resolution points are subjected to the same protocol (i.e., points are removed if they are too close to protein atoms or fall outside the convex hull), thus providing a more detailed description of binding-pocket geometries. As stray, isolated points often remain, the algorithm iteratively removes points that have fewer than a user-specified number of neighbors (4 by default) until no such points remain.

Finally, stretches of contiguous high-resolution points are grouped together, thus partitioning the points according to their associated pockets. The points of each pocket are further divided into a user-defined number of clusters (5 by default) using the k-means clustering algorithm, and encompassing spheres are generated for each cluster.

The script outputs a separate PDB file for the points of each pocket. Within each of these files, each cluster of points is assigned a unique PDB chain id. The PDB header describes the inclusion-region sphere associated with each chain/cluster, formatted for use in a POVME input file. Once the user has determined via visualization which of the identified pockets is of greatest interest, he or she can select which calculated spheres are required to entirely encompass that pocket.

Test System: RNA Editing Ligase 1. The 1XDN crystal structure,[66] which includes enzyme residues 52−365 as well as an active-site ATP molecule and magnesium ion, was obtained from the Protein Data Bank.[2] Selenomethionine residues were replaced with methionine. All crystallographic water molecules were retained. The AMBER LEaP module was used to submerge the protein in a rectangular box of water molecules that extended 10 Å beyond the system atoms in all three Cartesian dimensions. Monovalent ions were added to neutralize the system and to bring it to a 0.1 M salt concentration. The protein and water atoms were parameterized using the Amber99SB force field[67] and the TIP4P-ew water model,[68] respectively. Additionally, the parameters for ATP, magnesium, and monovalent ions developed by Meagher et al.,[69] Allner et al.,[70] and Joung and Cheatham[71] were used, respectively.

The REL1 system was subjected to five 5000-step energy minimizations using the NAMD molecular-dynamics simulation package[72,73] to gradually introduce full flexibility. We first allowed only hydrogen atoms to move, second released all water molecules, third released ions and ATP, fourth released the protein amino-acid side chains, and fifth removed all constraints. The system was then heated from 0 to 310 K in an NVT ensemble for 500 ps, with the protein backbone restrained. Equilibration was achieved in two segments, each consisting of a 250 ps simulation in the NPT ensemble. In the first segment, the protein backbone was restrained; in the second, no restraints were applied.

Five production simulations were performed, starting from the fully equilibrated structure. A total of 650 ns were simulated (one simulation of 250 ns and four of 100 ns). Different random seeds were used for each productive simulation to generate different starting velocities.

To study the flexibility of the REL1 active site, we extracted 6,500 frames from the simulations, evenly spaced 100 ps apart. All waters, counterions, ATP molecules, and magnesium ions were removed. VMD's RMSD Trajectory Tool[64] was used to align the extracted frames. In order to determine how differing alignment methodologies would impact the POVME analysis, we used several different protocols. The extracted frames were concatenated and aligned by 1) the atoms of the bound ATP ligand; 2) the atoms of the active-site residues (e.g., any residue within 5 Å of the crystallographic ligand); 3) the alpha-carbon atoms (C_α) of the active-site residues; and 4) the C_α of the entire protein. Each of these four aligned trajectories was saved as a separate multiframe PDB file (Text S1).

Separate POVME analyses were performed for each aligned trajectory. In each case, we characterized the combined ATP/transient pockets using an inclusion region defined by 10 carefully positioned spheres, chosen by visualizing the system in VMD. This region was filled with equidistant points spaced 1.0 Å apart. No exclusion regions were required. Points that were not contiguous with those contained within a small sphere centered at the opening of the ATP-binding pocket were discarded. The new convex-hull feature was enabled.

To benchmark POVME 1.0 and POVME 2.0, we further considered the REL1 trajectory aligned by all C_α. Additional analyses of this trajectory were performed using POVME 1.0 and POVME 2.0 with the new convex-hull feature disabled.

RESULTS AND DISCUSSION

As pocket volume and shape play critical roles in determining small-molecule binding, they are often the focus of computer-docking campaigns, QSAR studies, and molecular-dynamics analyses. We previously created a novel algorithm for characterizing macromolecular pockets called POVME (POcket Volume MEasurer) that has been widely adopted.[7] We here present a much-improved version of the algorithm, POVME 2.0.

POVME 2.0 has four primary advantages over previous versions. First, it is an order of magnitude faster because it relies on the *numpy* and *scipy* python modules to perform matrix-based calculations at nearly the speed of compiled C programs.[59−63] Additionally, the user can instruct POVME 2.0 to take advantage of multiple processors to further improve the speed of the calculation.

Second, POVME 2.0 comes with an optional graphical user interface (GUI) to facilitate usability (Figure 1). The GUI requires that Tkinter,[74] a python binding to the Tk GUI toolkit,[75] be installed. Fortunately, Tkinter is included in the standard Windows and OS X python distributions, as well as many Linux distributions.

Third, POVME 2.0 includes a new convex-hull-clipping option that improves the accuracy of the volume calculation. Portions of the binding pocket that fall outside the convex hull of nearby receptor atoms are discarded; consequently, only

portions of the pocket that are truly interior to the protein surface are considered.

Fourth, unlike the original version, POVME 2.0 can analyze entire trajectories in addition to single protein conformations. With POVME 1.0, users were required to save each trajectory frame as a separate PDB file in order to study changes in pocket volume and shape over the course of a MD trajectory. In contrast, POVME 2.0 can read multiframe trajectory files without requiring that each frame be saved separately. When analyzing MD trajectories, POVME outputs both frame-by-frame and whole-trajectory analyses. For frame-by-frame analysis, POVME saves the individual pocket shapes in the PDB format. For whole-trajectory analysis, POVME creates a volumetric density map showing the frequency with which different regions of the protein are included in the pocket over the course of the trajectory.

Test Case: *Trypanosoma brucei* RNA Editing Ligase 1 (REL1). To demonstrate the utility of this new POVME implementation, we used it to analyze an MD simulation of RNA editing ligase 1 (REL1) from the parasite *Trypanosoma brucei*, the etiological agent of African sleeping sickness. REL1 is a critical component of the *T. brucei* editosome, which edits transcriptional RNA prior to translation. This extensive RNA-editing process is essential for trypanosomatid survival, and REL1 has been shown to be a viable drug target.[76,77] Indeed, REL1 inhibitors have been identified that kill the whole-cell parasite.[78]

Previous studies of related crystal structures have hinted at the existence of a transient subpocket connected to the distal portion of the primary ATP-binding site that may provide unique opportunities for drug discovery.[79] Compounds that bind to the REL1 primary site may also target other ATP-binding proteins with structurally similar pockets; however, compounds that bind to the unique transient pocket may prove more target specific.

To better characterize the dynamics of the REL1 pockets, we examined 6,500 combined ATP-transient pockets extracted from 650 ns of MD simulations. We first aligned the trajectory to ensure that the binding pocket was consistently in the same location. As with other pocket-analysis programs,[12,17,80] simulation-trajectory alignment impacts the calculation of the average volumetric density maps. To demonstrate this sensitivity, we performed four separate POVME analyses, aligning the REL1 trajectory by 1) all ATP-ligand atoms; 2) all the atoms of the active-site residues; 3) the alpha-carbon (C_α) atoms of the active-site residues; and 4) the C_α atoms of the entire protein. Volumetric density maps were calculated for each of these aligned trajectories and were visualized superimposed on the receptor structure using VMD. When displayed as an isosurface, these density maps show the fraction of frames with measured pockets that included the displayed volume.

For the purposes of comparison, we judged the utility of each alignment protocol by how consistently the associated POVME analysis captured the ATP-binding-pocket region over the course of the entire trajectory. As our simulations included a bound ATP ligand, the ATP-binding subpocket should always be open (i.e., the region of the volumetric map corresponding to ATP in our simulations should have a high density, in excess of 95%).

When the trajectory was aligned by all active-site C_α atoms, the POVME-identified pocket consistently included the ATP-binding region (Figure 3B). We also found that aligning by all active-site atoms or even the atoms of the bound ligand itself led to similar POVME results (Figure S1). In contrast, the pocket analysis was less than optimal when the trajectory was aligned by the C_α atoms of the whole receptor (Figure 3C), likely because substantial protein motions distant from the active site led to poor binding-pocket alignment. Consequently, the transient pocket was identified as open only half as often when the trajectory was aligned by all C_α atoms vs active-site C_α atoms.

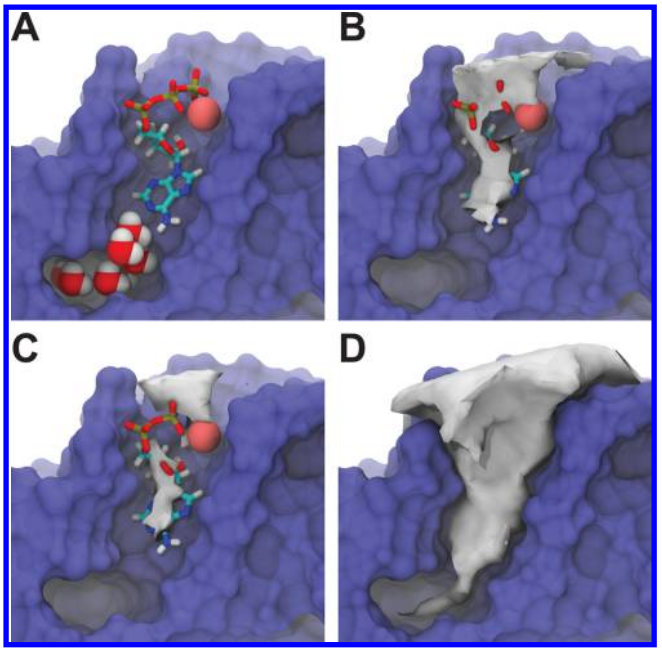


Figure 3. Volumetric density maps of the REL1 active site. Some regions of the protein have been removed to facilitate visualization. A) The crystallographic pose of the bound ATP molecule. Crystallographic water molecules indicate the location of a secondary binding pocket that is transiently accessible from the ATP-binding pocket. B) The region of the binding pocket identified as "open" at least 95% of the time when the trajectory was aligned by the active-site C_α atoms. C) The same region when the trajectory was aligned by the C_α atoms of the whole protein. D) The region of the binding pocket identified as "open" at least 25% of the time when the active-site-C_α alignment was again used.

While the best protocol to use is likely system dependent, based on these REL1 results we concur with others in recommending that trajectories be aligned by active-site C_α atoms.[80] This alignment 1) consistently identified the ATP-binding pocket; 2) does not rely on the presence of a bound ligand and so can be applied to *apo* systems as well; and 3) requires fewer atoms. When the binding pocket is partly composed of flexible loops, aligning by pocket C_α atoms that belong to stable secondary-structure elements may be more appropriate.

When the active-site-C_α alignment was used, POVME analysis revealed the full dynamics of the transient REL1 pocket, as indicated by the density maps in Figures 3B and 3D at isovalues of 95% and 25%, respectively. As expected given that we simulated the *holo* protein, the primary ATP-binding pocket was persistently open throughout the entire simulation (Figure 3D, 95% isovalue). The intermittent transient pocket was open at least 25% of the time (Figure 3D), suggesting a persistence sufficient to support our hypothesis of druggability.

Benchmarking/Insights into Use. To verify that the results of POVME 2.0 are comparable to those of the previous

version, we similarly analyzed a REL1 trajectory using POVME 1.0. When the convex-hull algorithm was disabled, both POVME 1.0 and 2.0 gave nearly identical volume measurements (Figure 4 graph, in black). When the new convex-hull

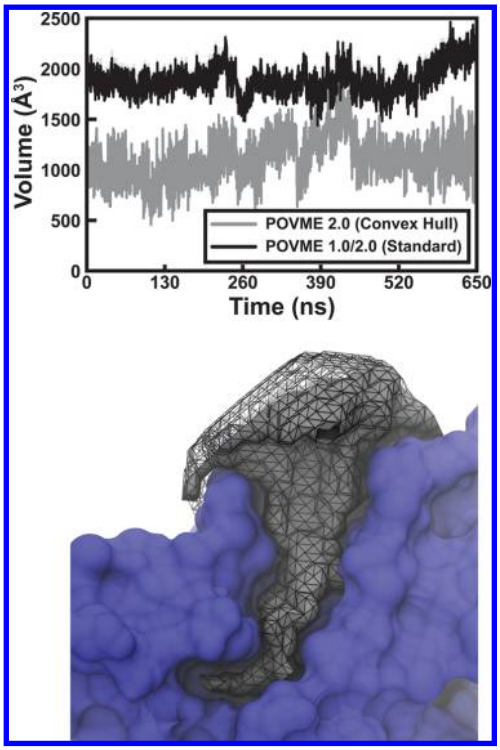


Figure 4. POVME 1.0 and 2.0 benchmarks. The graph shows benchmark REL1 pocket volumes as a function of simulation time. POVME 1.0 and 2.0 give nearly the same volume measurements (in black). When the POVME 2.0 convex-hull option is enabled, the volumes are smaller (in gray). The bottom panel, generated using the 1XDN crystal structure, illustrates the difference. When the convex-hull option is enabled, the region of the binding pocket is more accurately captured (solid gray) than when it is deactivated (black wireframe). Some portions of the protein have been removed to facilitate visualization.

feature was enabled, POVME-calculated volumes were lower, as expected (Figure 4 graph, in gray). To verify that the volumes calculated both with and without the convex-hull feature were correlated, we performed a linear regression. A two-tailed *t*-test suggested that the correlation was statistically significant (Pearson correlation coefficient: 0.19; *p*-value: 0.0).

Although the convex-hull feature does add computational expense, in some contexts it may lead to more accurate characterizations. If the user's ultimate goal is to calculate pocket volumes in absolute terms, clearly discarding volumes that lie outside the convex hull leads to improved accuracy (Figure 4, bottom panel). In contrast, if the goal is to determine how frequently different pocket regions are open vs closed by calculating volumetric density maps, the convex-hull feature, which discards only volumes that fall outside the pocket, will have little impact.

Finally, if the goal is to compare *changes* in the total pocket volume over the course of a trajectory, the convex-hull feature should be used with caution. Arbitrary movements of surface side chains near pocket openings can subtly affect the specific convex hull calculated for each frame. If the POVME inclusion region includes many points that lie outside the convex hull, and if the system being studied has highly flexible pocket-adjacent surface residues, using the convex-hull feature can introduce substantial "noise" into the total-volume calculations. This noise can make it difficult to detect subtle changes in the total volume caused by actual pocket dynamics, especially if the pocket itself is relatively stable. To overcome this challenge, we recommend carefully defining the POVME inclusion region to encompass only the pocket itself. If the convex-hull feature is used at all, rely on it only as a secondary refinement to better account for those rare frames when the pocket opening "pulls back," causing a few POVME points that are typically located within the pocket to barely fall outside the pocket boundaries.

Software Comparison. *Pocket Identification.* Many approaches have been developed to identify and measure protein pockets (see recent reviews[81,82]). These approaches are commonly divided into geometry- and energy-based detection algorithms.[17,81,82] We here limit our comparison to free software packages that natively process ensembles of structures, whether derived from multiple crystal structures or MD simulations, without requiring custom scripting. Software packages with this native support, including POVME 2.0, EPOS^{BP},[9] MDpocket,[12] PocketAnalyzer^{PCA},[80] and *trj_cavity*,[83] can output ensemble-dependent features such as the frequency of transient subpocket opening and the pocket size over multiple frames.

For all these programs, pocket identification is the initial step. POVME 2.0 pocket-defining regions can be generated automatically using the new grid-based POVME Pocket ID helper script described above. PocketAnalyzer^{PCA} and *trj_cavity* use similar grid-based detection algorithms. PocketAnalyzer^{PCA} implements a variant of the LIGSITE algorithm,[84,85] and *trj_cavity* employs a novel neighbor-search method. In contrast, MDpocket uses fpocket to detect and calculate pocket properties. Fpocket calls the Qhull algorithm[86] to perform a Voronoi tessellation of the receptor; the coordinates of each Voronoi vertex, together with the associated atomic and vertex neighbors, constitute an "α-sphere" with a distinct radius. The collection of all α-spheres is then used to locate protein pockets. Finally, EPOS^{BP} relies on the PASS algorithm, which first covers the entire protein surface with spheres and then removes spheres with low burial counts. As a second step, the algorithm covers the remaining spheres from the first step with an additional layer of spheres. Pockets are enumerated once repeating this process iteratively has successfully filled all cavities. EPOS^{BP} also maps the identified pockets to a set of pocket-lining atoms (PLAs), avoiding the need to align the trajectory.

Volume Measurements. Pocket volumetric analysis follows the initial identification step. In order to facilitate comparison, we repeated the trajectory-based analysis of the REL1 ATP-binding pocket using several programs (active-site-C_α alignment). For each program, a grid resolution of 1.0 Å was used where appropriate. The remaining parameters were set to their default values.

Given the set of POVME inclusion and exclusion spheres chosen, POVME 2.0 without the convex-hull algorithm generally calculated pocket volumes that were larger than those determined by other codes. In contrast, when the convex-hull algorithm was used to ignore regions of space outside the confines of the pocket, the POVME-calculated volumes decreased substantially (see Table 2 and Figure S2).

If accuracy is judged by consistency with other programs, using POVME 2.0 together with the convex-hull algorithm

Table 2. Program Comparisons[a]

program	average volume ± SD	run time (1 thread)	run time (24 threads)
POVME 1.0		10500	
POVME 2.0	2071.3 ± 129.1	16	2
POVME 2.0/convex hull	1021.8 ± 154.8	92	8
trj_cavity	814.4 ± 254.6	16	
MDpocket	523.2 ± 60.5	11	
PocketAnalyzer^{PCA}	811.5 ± 100.2	96*	
EPOS^{BP} without clustering		43	
EPOS^{BP} with clustering		259	

[a]The average pocket volume (in Å^3), plus or minus the standard deviation, measured over the course of a REL1 trajectory using several pocket-analysis programs. Note that the POVME 1.0 results were, for all intents and purposes, identical to the POVME 2.0 results with the convex-hull feature disabled. Additionally, EPOS^{BP} volume measurements are not included because that program does not output volume-per-frame data. The total run times for each program are given in minutes. A PCA calculation accounted for approximately 6 min of the PocketAnalyzer^{PCA} run time (marked with an asterisk). As POVME 2.0 is designed to use multiple processors, the run times for parallel POVME 2.0 calculations are also shown.

leads to substantial improvements when the user wishes to measure pocket volume in absolute terms. Nevertheless, the POVME-calculated volumes are still somewhat larger than those calculated by other algorithms, even with the convex-hull algorithm enabled. This discrepancy does not necessarily mean that POVME volumes are overestimated, as determining the accuracy of these varied methods is not straightforward. Binding pockets are not uniquely defined; they often have an opening toward the solvent, and no commonly agreed upon criteria exists for delineating the exact boundary between the pocket and solvent-filled spaces.

The tunable parameters of each algorithm may also contribute to the differences in calculated volumes. Although it is useful to allow the user to adjust key parameters as required to best analyze a specific pocket of interest, doing so often presents a problem when the parameters are not intuitive. With the end user in mind, POVME 2.0 requires only a small number of parameters with straightforward physical interpretations. We believe this is a distinct advantage of our code. While useful, parameters such as degree of buriedness and minimal cluster size (PocketAnalyzer^{PCA}); PASS input (EPOS^{BP}); and allowed range of α-sphere radii and number of α-spheres per pocket (MDpocket) are less intuitive.

PocketAnalyzer^{PCA} and *trj_cavity* volumes are somewhat smaller than the POVME volumes obtained when the convex-hull feature is enabled; in contrast, MDpocket volumes are substantially smaller and have only minor fluctuations (Table 2 and Figure S2), possibly because MDpocket uses α-spheres rather than equidistant points. To determine if the volumes calculated using these four methods were even correlated, we first reduced the noise in the data using a simple moving average with a 40-point sample window. We then performed linear regressions to explore the relationships between these averaged POVME volumes and the averaged volumes obtained using PocketAnalyzer^{PCA}, *trj_cavity*, and MDpocket, respectively. A two-tailed *t*-test was applied to the Pearson's correlation coefficients to judge significance. The *p*-values were 0.46, 0.00, and 0.00 for POVME vs PocketAnalyzer^{PCA}, *trj_cavity*, and MDpocket, respectively, suggesting that the correlations between POVME volumes and both *trj_cavity* and MDpocket volumes are statistically significant.

Future efforts will focus on expanding POVME analysis beyond simple volume and shape calculations. We intend to implement other pocket metrics (e.g., pocket surface-area calculations) and analysis tools (e.g., volume-based pharmacophore modeling).

Execution Time. The trajectory-based pocket calculations were also used to compare the execution times of each program (Table 2). All calculations were run on the same machine (2 Intel Xeon X5690 3.47Ghz processors, 6 CPU cores per processor, 2 threads per core = 24 threads total). On a single thread, POVME 2.0 is several orders of magnitude faster than its predecessor. When the convex-hull feature is disabled, POVME 2.0 is also faster than EPOS^{BP} and PocketAnalyzer^{PCA}, comparable to *trj_cavity,* and only slightly slower than MDpocket. We are exploring alternative methods for optimizing the optional convex-hull feature. Fortunately, the current POVME implementation is highly parallelized; when 24 threads are used rather than one, POVME 2.0 is substantially faster than all other packages tested, even with the convex-hull feature enabled.

Program Input and Output. POVME 2.0, *trj_cavity*, MDpocket, PocketAnalyzer^{PCA}, and EPOS^{BP} all output simple PDB file(s) that are easily visualized with other software packages. POVME 2.0 and MDpocket also output volumetric density maps in the DX format to make exploration of potentially druggable transient subpockets with varying opening frequencies simpler to interpret. Finally, EPOS^{BP} and PocketAnalyzer^{PCA} both have built-in features to cluster pocket shapes. Similarly, POVME output is easily clustered using utilities like the volume-overlap tool in Schrödinger's Maestro suite.[55,56]

POVME 2.0 accepts a multiframe PDB file as input. As the PDB format is general purpose and universally accepted, POVME is independent of any specific simulation package or force field. Nevertheless, PDB trajectories can become quite large, so being able to read binary formats specific to selected simulation software packages would also be useful. Indeed, *trj_cavity* can be compiled to read binary trajectories saved in the Gromacs format. Future implementations of POVME may include support for additional formats, though we do wish to avoid limiting our software to a specific simulation engine. We note that POVME trajectories do not include water molecules, substantially cutting down on the file size even when the PDB format is used. Additionally, freely available software can easily convert binary formats to PDB. A helpful tutorial is included in the Supporting Information that describes how to perform this conversion using VMD[64] (Text S1).

CONCLUSION

POVME 2.0 is a much improved version of our popular algorithm for characterizing the volumes and shapes of macromolecular (e.g., protein) binding pockets. Version 2.0 implements a number of enhancements, including speed improvements due to *numpy/scipy* integration and the optional use of multiple processors; better accuracy due to an optional convex-hull implementation; additional volumetric-analysis tools (i.e., volumetric density maps); and a graphical user interface that improves usability.

Although pocket-shape and volumetric analyses are not novel, factors such as the high computational cost of most

algorithms have discouraged widespread adoption. POVME 2.0 significantly reduces the amount of time required, allowing users to more rapidly analyze large ensembles of pocket shapes derived from multiple experimental structures or simulation methods, such as MD. The added volumetric-density-map analysis feature provides a pocket-centric view of receptor flexibility with potentially useful drug-discovery applications. Indeed, others have shown that docking into structurally distinct binding pockets can lead to enhanced hit rates and chemical diversity.[55−58]

To demonstrate how POVME 2.0 can provide pharmacologically relevant information about pocket flexibility, we used it to analyze the dynamics of an essential, ATP-binding component of the *T. brucei* editosome, REL1.[76,77] Given that ATP-binding pockets are ubiquitous, small-molecule inhibitors that bind exclusively to the primary REL1 pocket may also bind to the ATP pockets of critical human enzymes, leading to undesirable side effects. Consequently, we considered a unique secondary binding pocket that is transiently accessible from the primary REL1 ATP-binding pocket. POVME suggests this transient pocket assumes an open conformation roughly 25% of the time. Identifying less promiscuous REL1 inhibitors that exploit this unique pocket is an important component of our ongoing efforts to target this crucial enzyme.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information contains two files. The first contains Figures S1 and S2. Figure S1 shows the POVME volumetric density maps generated when the REL1 trajectory was aligned by all active-site atoms and the atoms of the bound ligand. Figure S2 shows the pocket volumes calculated over the course of a REL1 molecular dynamics simulation, using several different software packages. The second file, Text S1, contains a tutorial that shows how VMD can be used to align a trajectory. The same tutorial also shows how to save a trajectory in the multiframe PDB format for subsequent POVME analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 858.534.9629. Fax: 858.534.9645. E-mail: ramaro@ucsd.edu. Corresponding author address: Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, Mail Code 0340, La Jolla, CA 92093-0340.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Dr. Victoria A. Feher and Mr. Jeffrey R. Wagner for helpful discussions, as well as Mr. Cam Farnell for developing Rapyd-Tk, a python-based program used to develop the POVME 2.0 graphical user interface. This work was funded by the National Institutes of Health through the NIH Director's New Innovator Award Program DP2-OD007237 and the NSF XSEDE Supercomputer resources grant RAC CHE060073N to R.E.A. L.V. is funded by the National Science Foundation's Graduate Research Fellowship Program. J.S. thanks the Alfred Benzon Foundation for generous postdoctoral funding. In addition, this work was also supported by the National Biomedical Computation Resource (NBCR), P41 GM103426.

■ REFERENCES

(1) Perot, S.; Sperandio, O.; Miteva, M. A.; Camproux, A. C.; Villoutreix, B. O. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discovery Today* **2010**, *15*, 656−667.

(2) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(3) Levitt, D. G.; Banaszak, L. J. Pocket - a Computer-Graphics Method for Identifying and Displaying Protein Cavities and Their Surrounding Amino-Acids. *J. Mol. Graphics* **1992**, *10*, 229−234.

(4) Smart, O. S.; Goodfellow, J. M.; Wallace, B. A. The Pore Dimensions of Gramicidin-A. *Biophys. J.* **1993**, *65*, 2455−2460.

(5) Kleywegt, G. J.; Jones, T. A. Detection, Delineation, Measurement and Display of Cavities in Macromolecular Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 178−185.

(6) Laskowski, R. A. Surfnet - a Program for Visualizing Molecular-Surfaces, Cavities, and Intermolecular Interactions. *J. Mol. Graphics* **1995**, *13*, 323−330.

(7) Durrant, J. D.; de Oliveira, C. A.; McCammon, J. A. POVME: An algorithm for measuring binding-pocket volumes. *J. Mol. Graphics Modell.* **2011**, *29*, 773−776.

(8) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8*, e1002708.

(9) Eyrisch, S.; Helms, V. Transient pockets on protein surfaces involved in protein-protein interaction. *J. Med. Chem.* **2007**, *50*, 3457−3464.

(10) Brady, G. P.; Stouten, P. F. W. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 383−401.

(11) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinf.* **2009**, *10*, 168.

(12) Schmidtke, P.; Bidon-Chanal, A.; Luque, F. J.; Barril, X. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* **2011**, *27*, 3276−3285.

(13) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377−389.

(14) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146−148.

(15) Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, *25*, 621−627.

(16) Votapka, L.; Amaro, R. E. Multistructural hot spot characterization with FTProd. *Bioinformatics* **2013**, *29*, 393−394.

(17) Zheng, X. L.; Gan, L. F.; Wang, E. K.; Wang, J. Pocket-Based Drug Design: Exploring Pocket Space. *AAPS J.* **2013**, *15*, 228−241.

(18) Amaro, R. E.; Swift, R. V.; Votapka, L.; Li, W. W.; Walker, R. C.; Bush, R. M. Mechanism of 150-cavity formation in influenza neuraminidase. *Nat. Commun.* **2011**, *2*, 388.

(19) Baron, R.; Vellore, N. A. LSD1/CoREST is an allosteric nanoscale clamp regulated by H3-histone-tail molecular recognition. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 12509−12514.

(20) Fuchs, J. E.; Huber, R. G.; Von Grafenstein, S.; Wallnoefer, H. G.; Spitzer, G. M.; Fuchs, D.; Liedl, K. R. Dynamic Regulation of Phenylalanine Hydroxylase by Simulated Redox Manipulation. *PLoS One* **2012**, *7*, e53005.

(21) Sinko, W.; de Oliveira, C.; Williams, S.; Van Wynsberghe, A.; Durrant, J. D.; Cao, R.; Oldfield, E.; McCammon, J. A. Applying Molecular Dynamics Simulations to Identify Rarely Sampled Ligand-bound Conformational States of Undecaprenyl Pyrophosphate

Synthase, an Antibacterial Target. *Chem. Biol. Drug Des.* **2011**, *77*, 412−420.
(22) Lindert, S.; McCammon, J. A. Dynamics of Plasmodium falciparum enoyl-ACP reductase and implications on drug discovery. *Protein Sci.* **2012**, *21*, 1734−1745.
(23) Boechi, L.; de Oliveira, C. A.; Da Fonseca, I.; Kizjakina, K.; Sobrado, P.; Tanner, J. J.; McCammon, J. A. Substrate-dependent dynamics of UDP-galactopyranose mutase: Implications for drug design. *Protein Sci.* **2013**, *22*, 1490−1501.
(24) Wu, Y.; Qin, G. R.; Gao, F.; Liu, Y.; Vavricka, C. J.; Qi, J. X.; Jiang, H. L.; Yu, K. Q.; Gao, G. F. Induced opening of influenza virus neuraminidase N2 150-loop suggests an important role in inhibitor binding. *Sci. Rep.* **2013**, *3*, 1551.
(25) Han, N. Y.; Mu, Y. G. Plasticity of 150-Loop in Influenza Neuraminidase Explored by Hamiltonian Replica Exchange Molecular Dynamics Simulations. *PLoS One* **2013**, *8*, e60995.
(26) Schultes, S.; Nijmeijer, S.; Engelhardt, H.; Kooistra, A. J.; Vischer, H. F.; de Esch, I. J. P.; Haaksma, E. E. J.; Leurs, R.; de Graaf, C. Mapping histamine H-4 receptor-ligand binding modes. *MedChemComm* **2013**, *4*, 193−204.
(27) Li, P.; Chen, Z.; Xu, H.; Sun, H.; Li, H.; Liu, H.; Yang, H.; Gao, Z.; Jiang, H.; Li, M. The gating charge pathway of an epilepsy-associated potassium channel accommodates chemical ligands. *Cell Res.* **2013**, *23*, 1106−1118.
(28) Kekenes-Huskey, P. M.; Metzger, V. T.; Grant, B. J.; McCammon, J. A. Calcium binding and allosteric signaling mechanisms for the sarcoplasmic reticulum Ca2+ATPase. *Protein Sci.* **2012**, *21*, 1429−1443.
(29) Bung, N.; Pradhan, M.; Srinivasan, H.; Bulusu, G. Structural Insights into E. coli Porphobilinogen Deaminase during Synthesis and Exit of 1-Hydroxymethylbilane. *PLoS Comput. Biol.* **2014**, *10*, e1003484.
(30) Torres, R.; Swift, R. V.; Chim, N.; Wheatley, N.; Lan, B. S.; Atwood, B. R.; Pujol, C.; Sankaran, B.; Bliska, J. B.; Amaro, R. E.; Goulding, C. W. Biochemical, Structural and Molecular Dynamics Analyses of the Potential Virulence Factor RipA from Yersinia pestis. *PLoS One* **2011**, *6*, e25084.
(31) Grant, B. J.; Lukman, S.; Hocker, H. J.; Sayyah, J.; Brown, J. H.; McCammon, J. A.; Gorfe, A. A. Novel Allosteric Sites on Ras for Lead Generation. *PLoS One* **2011**, *6*, e25711.
(32) Mowrey, D. D.; Liu, Q.; Bondarenko, V.; Chen, Q.; Seyoum, E.; Xu, Y.; Wu, J.; Tang, P. Insights into Distinct Modulation of alpha 7 and alpha 7 beta 2 Nicotinic Acetylcholine Receptors by the Volatile Anesthetic Isoflurane. *J. Biol. Chem.* **2013**, *288*, 35793−35800.
(33) Yi-Xin, A.; Jun-Rui, L.; Chun-Wei, X.; Jiang-Bei, M.; Xu-Yun, Y.; He, Z. Simulated Mechanism of Triclosan in Modulating the Active Site and Loop of FabI by Computer. *Acta Phys.-Chim. Sin.* **2014**, *30*, 559−568.
(34) Blachly, P. G.; de Oliveira, C. A. F.; Williams, S. L.; McCammon, J. A. Utilizing a Dynamical Description of IspH to Aid in the Development of Novel Antimicrobial Drugs. *PLoS Comput. Biol.* **2013**, *9*, e1003395.
(35) Demir, O.; Amaro, R. E. Elements of Nucleotide Specificity in the Trypanosoma brucei Mitochondrial RNA Editing Enzyme RET2. *J. Chem. Inf. Model.* **2012**, *52*, 1308−1318.
(36) Mowrey, D.; Cheng, M. H.; Liu, L. T.; Willenbring, D.; Lu, X. H.; Wymore, T.; Xu, Y.; Tang, P. Asymmetric Ligand Binding Facilitates Conformational Transitions in Pentameric Ligand-Gated Ion Channels. *J. Am. Chem. Soc.* **2013**, *135*, 2172−2180.
(37) Bustamante, J. P.; Abbruzzetti, S.; Marcelli, A.; Gauto, D.; Boechi, L.; Bonamore, A.; Boffi, A.; Bruno, S.; Feis, A.; Foggi, P.; Estrin, D. A.; Viappiani, C. Ligand Uptake Modulation by Internal Water Molecules and Hydrophobic Cavities in Hemoglobins. *J. Phys. Chem. B* **2014**, *118*, 1234−1245.
(38) Selvam, B.; Porter, S. L.; Tikhonova, I. G. Addressing Selective Polypharmacology of Antipsychotic Drugs Targeting the Bioaminergic Receptors through Receptor Dynamic Conformational Ensembles. *J. Chem. Inf. Model.* **2013**, *53*, 1761−1774.
(39) Weinreb, V.; Li, L.; Chandrasekaran, S. N.; Koehl, P.; Delarue, M.; Carter, C. W., Jr. Enhanced Amino Acid Selection in Fully Evolved Tryptophanyl-tRNA Synthetase, Relative to Its Urzyme, Requires Domain Motion Sensed by the D1 Switch, a Remote Dynamic Packing Motif. *J. Biol. Chem.* **2014**, *289*, 4367−4376.
(40) Li, J. N.; Jonsson, A. L.; Beuming, T.; Shelley, J. C.; Voth, G. A. Ligand-Dependent Activation and Deactivation of the Human Adenosine A(2A) Receptor. *J. Am. Chem. Soc.* **2013**, *135*, 8749−8759.
(41) Baron, R.; McCammon, J. A. Molecular Recognition and Ligand Association. *Annu. Rev. Phys. Chem.* **2013**, *64*, 151−175.
(42) Ariga, K.; Ito, H.; Hill, J. P.; Tsukube, H. Molecular recognition: from solution science to nano/materials technology. *Chem. Soc. Rev.* **2012**, *41*, 5800−5835.
(43) Kahraman, A.; Morris, R. J.; Laskowski, R. A.; Thornton, J. M. Shape variation in protein binding pockets and their ligands. *J. Mol. Biol.* **2007**, *368*, 283−301.
(44) Seddon, G.; Lounnas, V.; McGuire, R.; van den Bergh, T.; Bywater, R. P.; Oliveira, L.; Vriend, G. Drug design for ever, from hype to hope. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 137−150.
(45) Meng, X. Y.; Zhang, H. X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput.-Aided Drug Des.* **2011**, *7*, 146−157.
(46) Golbraikh, A.; Wang, X. S.; Zhu, H.; Tropsha, A. Predictive QSAR Modeling: Methods and Applications in Drug Discovery and Chemical Risk Assessment. In *Handbook of Computational Chemistry*; Leszczynski, J., Ed.; Springer: Dordrecht, Netherlands, 2012; pp 1309−1342.
(47) Liang, J.; Edelsbrunner, H.; Woodward, C. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci.* **1998**, *7*, 1884−1897.
(48) Rush, T. S.; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48*, 1489−1495.
(49) Wirth, M.; Volkamer, A.; Zoete, V.; Rippmann, F.; Michielin, O.; Rarey, M.; Sauer, W. H. B. Protein pocket and ligand shape comparison and its application in virtual screening. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 511−524.
(50) Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74−82.
(51) Distinto, S.; Esposito, F.; Kirchmair, J.; Cardia, M. C.; Gaspari, M.; Maccioni, E.; Alcaro, S.; Markt, P.; Wolber, G.; Zinzula, L.; Tramontano, E. Identification of HIV-1 reverse transcriptase dual inhibitors by a combined shape-, 2D-fingerprint- and pharmacophore-based virtual screening approach. *Eur. J. Med. Chem.* **2012**, *50*, 216−229.
(52) LaLonde, J. M.; Elban, M. A.; Courter, J. R.; Sugawara, A.; Soeta, T.; Madani, N.; Princiotto, A. M.; Do Kwon, Y.; Kwong, P. D.; Schon, A.; Freire, E.; Sodroski, J.; Smith, A. B. Design, synthesis and biological evaluation of small molecule inhibitors of CD4-gp120 binding based on virtual screening. *Bioorg. Med. Chem.* **2011**, *19*, 91−101.
(53) Tuccinardi, T.; Ortore, G.; Santos, M. A.; Marques, S. M.; Nuti, E.; Rossello, A.; Martinelli, A. Multitemplate Alignment Method for the Development of a Reliable 3D-QSAR Model for the Analysis of MMP3 Inhibitors. *J. Chem. Inf. Model.* **2009**, *49*, 1715−1724.
(54) Nicholls, A.; McGaughey, G. B.; Sheridan, R. P.; Good, A. C.; Warren, G.; Mathieu, M.; Muchmore, S. W.; Brown, S. P.; Grant, J. A.; Haigh, J. A.; Nevins, N.; Jain, A. N.; Kelley, B. Molecular Shape and Medicinal Chemistry: A Perspective. *J. Med. Chem.* **2010**, *53*, 3862−3886.
(55) Osguthorpe, D. J.; Sherman, W.; Hagler, A. T. Exploring Protein Flexibility: Incorporating Structural Ensembles From Crystal Structures and Simulation into Virtual Screening Protocols. *J. Phys. Chem. B* **2012**, *116*, 6952−6959.
(56) Osguthorpe, D. J.; Sherman, W.; Hagler, A. T. Generation of Receptor Structural Ensembles for Virtual Screening Using Binding Site Shape Analysis and Clustering. *Chem. Biol. Drug Des.* **2012**, *80*, 182−193.

(57) Ben Nasr, N.; Guillemain, H.; Lagarde, N.; Zagury, J. F.; Montes, M. Multiple Structures for Virtual Ligand Screening: Defining Binding Site Properties-Based Criteria to Optimize the Selection of the Query. *J. Chem. Inf. Model.* **2013**, *53*, 293−311.
(58) Nichols, S. E.; Swift, R. V.; Amaro, R. E. Rational Prediction with Molecular Dynamics for Hit Identification. *Curr. Top. Med. Chem.* **2012**, *12*, 2002−2012.
(59) Ascher, D.; Dubois, P. F.; Hinsen, K.; James, J. H.; Oliphant, T. *Numerical Python*; UCRL-MA-128569 ed.; Lawrence Livermore National Laboratory: Livermore, CA, 1999.
(60) Dubois, P. F. Extending Python with Fortran. *Comput. Sci. Eng.* **1999**, *1*, 66−73.
(61) Jones, E.; Oliphant, T.; Peterson, P. *Others SciPy: Open Source Scientific Tools for Python*, 0.11.0; 2001.
(62) Oliphant, T. E. *Guide to NumPy*; Brigham Young University: Provo, UT, 2006.
(63) Peterson, P. F2PY: a tool for connecting Fortran and Python programs. *Int. J. Comput. Sci. Eng.* **2009**, *4*, 296−305.
(64) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(65) Akl, S. G.; Toussaint, G. T. In *Efficient convex hull algorithms for pattern recognition applications*, Proc. 4th. Int. Joint Conf. on Pattern Recognition (Kyoto, Japan), 1978; pp 483−487.
(66) Deng, J.; Schnaufer, A.; Salavati, R.; Stuart, K. D.; Hol, W. G. High resolution crystal structure of a key editosome enzyme from Trypanosoma brucei: RNA editing ligase 1. *J. Mol. Biol.* **2004**, *343*, 601−613.
(67) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712−725.
(68) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665−9678.
(69) Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.* **2003**, *24*, 1016−1025.
(70) Allner, O.; Nilsson, L.; Villa, A. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 1493−1502.
(71) Joung, I. S.; Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(72) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **1999**, *151*, 283−312.
(73) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(74) Shipman, J. W. *Tkinter reference: a GUI for Python*; New Mexico Tech Computer Center: Socorro, NM, 2010.
(75) Welch, B. B.; Jones, K. *Practical programming in Tcl/Tk*, 4th ed.; Prentice Hall/PTR: Upper Saddle River, NJ, 2003.
(76) Schnaufer, A.; Panigrahi, A. K.; Panicucci, B.; Igo, R. P., Jr.; Salavati, R.; Stuart, K. An RNA Ligase Essential for RNA Editing and Survival of the Bloodstream Form of Trypanosoma brucei. *Science* **2001**, *291*, 2159−2162.
(77) Rusche, L. N.; Huang, C. E.; Piller, K. J.; Hemann, M.; Wirtz, E.; Sollner-Webb, B. The two RNA ligases of the Trypanosoma brucei RNA editing complex: cloning the essential band IV gene and identifying the band V gene. *Mol. Cell. Biol.* **2001**, *21*, 979−989.
(78) Durrant, J. D.; Hall, L.; Swift, R. V.; Landon, M.; Schnaufer, A.; Amaro, R. E. Novel Naphthalene-Based Inhibitors of Trypanosoma brucei RNA Editing Ligase 1. *PLoS Neglected Trop. Dis.* **2010**, *4*, e803.
(79) Durrant, J. D.; Friedman, A. J.; McCammon, J. A. CrystalDock: a novel approach to fragment-based drug design. *J. Chem. Inf. Model.* **2011**, *51*, 2573−2580.
(80) Craig, I. R.; Pfleger, C.; Gohlke, H.; Essex, J. W.; Spiegel, K. Pocket-Space Maps To Identify Novel Binding-Site Conformations in Proteins. *J. Chem. Inf. Model.* **2011**, *51*, 2666−2679.
(81) Ghersi, D.; Sanchez, R. Beyond structural genomics: computational approaches for the identification of ligand binding sites in protein structures. *J. Struct. Funct. Genomics* **2011**, *12*, 109−117.
(82) Perot, S.; Sperandio, O.; Miteva, M. A.; Camproux, A. C.; Villoutreix, B. O. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discovery Today* **2010**, *15*, 656−667.
(83) Paramo, T.; East, A.; Garzon, D.; Ulmschneider, M. B.; Bond, P. J. Efficient Characterization of Protein Cavities within Molecular Simulation Trajectories: trj_cavity. *J. Chem. Theory Comput.* **2014**, *10*, 2151−2164.
(84) Hendlich, M.; Rippmann, F.; Barnickel, G. LIGSITE: Automatic and efficient detection of potential small molecule-binding sites in proteins. *J. Mol. Graphics Modell.* **1997**, *15*, 359−363.
(85) Stahl, M.; Taroni, C.; Schneider, G. Mapping of protein surface cavities and prediction of enzyme class by a self-organizing neural network. *Protein Eng.* **2000**, *13*, 83−88.
(86) Barber, C. B.; Dobkin, D. P.; Huhdanpaa, H. The Quickhull algorithm for convex hulls. *ACM Trans. Math. Software* **1996**, *22*, 469−483.