Journal of Molecular Graphics and Modelling 29 (2011) 773–776

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Short communication

POVME: An algorithm for measuring binding-pocket volumes

Jacob D. Durrant[a,*], César Augusto F. de Oliveira[b,c], J. Andrew McCammon[b,c]

[a] Department of Chemistry & Biochemistry, University of California San Diego, La Jolla, CA 92093-0365, United States
[b] Howard Hughes Medical Institute, University of California San Diego, La Jolla, CA 92093-0365, United States
[c] Department of Chemistry & Biochemistry, Department of Pharmacology, and NSF Center for Theoretical Biological Physics, University of California San Diego, La Jolla, CA 92093-0365, United States

ARTICLE INFO

Article history:
Received 21 September 2010
Received in revised form 21 October 2010
Accepted 26 October 2010
Available online 3 November 2010

Keywords:
POVME
Binding-pocket volume
Algorithm
Computer-aided drug design

ABSTRACT

Researchers engaged in computer-aided drug design often wish to measure the volume of a ligand-binding pocket in order to predict pharmacology. We have recently developed a simple algorithm, called POVME (POcket Volume MEasurer), for this purpose. POVME is Python implemented, fast, and freely available. To demonstrate its utility, we use the new algorithm to study three members of the matrix-metalloproteinase family of proteins. Despite the structural similarity of these proteins, differences in binding-pocket dynamics are easily identified.

1. Introduction

The volume of a binding site has great pharmacological significance, both as one of the many structural features comprising the pharmacophore and as one of the characterizations used in QSAR. When the volume of a target binding site is known, potential ligands that are too large to fit within that volume can be eliminated early in the lead-generation process, prior to virtual or high-throughput screening. Additionally, variations in pocket size when multiple structures of the same protein are considered can provide pharmacologically useful insights into protein dynamics.

Our lab has recently developed a simple algorithm called POVME (POcket Volume MEasurer) for measuring the volume of ligand-binding sites. POVME has been implemented in Python and so is editable, customizable, and platform independent. Additionally, the algorithm is fast, open source, and easy to use.

2. Materials and methods

2.1. The POVME algorithm

The POVME algorithm includes four steps. First, the user must select a region defined by overlapping spheres and right rectangular prisms that entirely encompasses the binding pocket but does not extend beyond the volumes of the outermost, solvent-exposed protein atoms (Fig. 1a). Additionally, volume can be subtracted from this region using exclusion spheres and prisms. For example, the region encompassing the matrix-metalloproteinase binding pocket shown in Fig. 1a was defined using six inclusion spheres of radius 3 Å or greater and 13 exclusion spheres of radius 2.5 Å or smaller. If multiple protein structures are to be measured (e.g., frames extracted from a molecular dynamics trajectory), all structures must be aligned so that the single defined region encompasses all of the binding pockets.

Second, the algorithm creates a single volume-grid file by filling the defined region with equispaced points (Fig. 1b). To define an encompassing region and to generate its associated volume-grid file, a computer program that facilitates molecular visualization may be helpful. One free option is Visual Molecular Dynamics (VMD) [1], a program provided by the Theoretical and Computational Biophysics Group at the University of Illinois at Urbana-Champaign. A useful iterative technique might be to define an initial sphere or prism using the coordinates of a protein atom lining the active site. POVME can then be used to generate an inter-

* Corresponding author at: Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, Mail Code 0365, La Jolla, CA 92093-0365, United States. Tel.: +1 858 822 0168; fax: +1 858 534 4974.
E-mail addresses: jdurrant@ucsd.edu (J.D. Durrant), cesar@mccammon.ucsd.edu (C.A.F. de Oliveira), jmccammon@mail.ucsd.edu (J.A. McCammon).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.10.007

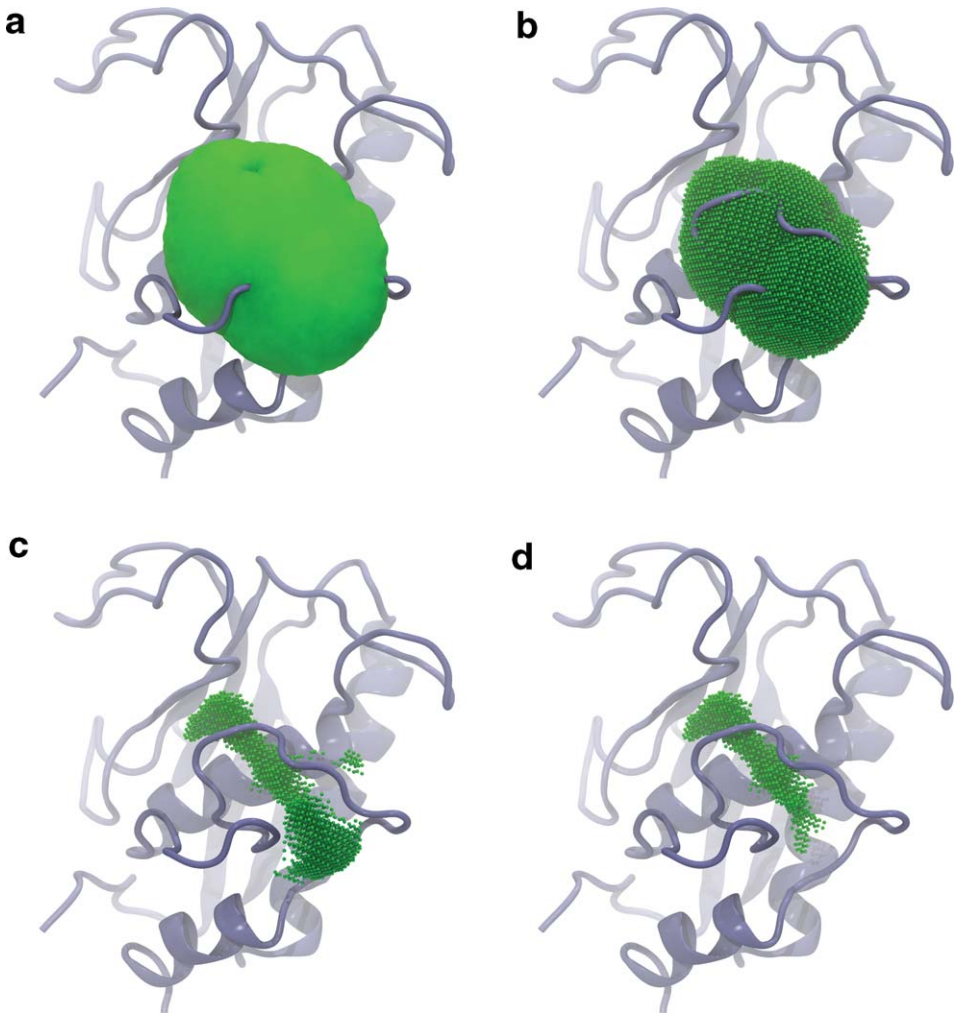

Fig. 1. The POVME algorithm measures the volume of ligand-binding pockets. (a) First, a region is defined that encompasses the binding pocket but is contained within the boundaries of the protein. (b) The specified volume is filled with equispaced points. (c) Those points near protein atoms are removed, so that only points in the binding pocket remain. (d) Optionally, remaining points that are isolated or discontiguous can also be removed. The volume is calculated from the remaining points.

mediate volume-grid file, which is saved in the PDB format and so can itself be loaded into VMD together with the protein structure. The coordinates of selected points from the intermediate volume-grid file or the coordinates of additional protein atoms can then be used to define additional inclusion and/or exclusion spheres and prisms, as needed. By iteratively defining new spheres and prisms and running POVME to generate intermediate volume-grid files, a region encompassing the protein binding pocket can be easily constructed.

After the binding-pocket region and final volume-grid file have been defined and generated, volume-grid points near protein atoms are systematically deleted, leaving only those points present in the binding pocket (Fig. 1c). As an optional step, remaining points that are isolated or discontiguous can be removed (Fig. 1d). For example, suppose a tunnel-like binding pocket collapses in the middle, so that the points used to define the volume of that binding pocket are divided into two discontinuous regions. The sum of these two discontiguous volumes may misrepresent the actual volume of the pocket; rather, it may be useful to remove the region distal to the collapse and consider only the volume of the proximal region.

Finally, because each of the points is equispaced, each represents a fixed volume. The volume of the binding pocket can therefore be calculated based on the number of points remaining.

2.2. Molecular dynamics simulations

To study the pocket-volume dynamics of three matrix metalloproteinases, we performed a molecular dynamics simulation of each, generally following a protocol described previously [2]. In brief, crystal structures of MMP-2, MMP-3, and MMP-9 (PDB IDs 1QIB, 1QIA, and 1L6J, respectively) [3–5] were prepared and subjected to 1500 steps of minimization. Each of the minimized structures was then solvated and neutralized, and the water molecules were subjected to another 2500 steps of minimization using constant-volume periodic boundaries. To achieve the correct system density, water molecules were subjected to 50 ps of simulation using an NPT ensemble (T = 298 K, P = 1 bar). Next, all atoms of each system were again relaxed via 1500 steps of minimization, and then heated from 0 to 300 K via a 500-ps simulation using the NVT ensemble (T = 298 K). Seventy-five and a half nanoseconds of productive MD simulation were subsequently performed for each MMP using the NVT ensemble at constant temperature (T = 298 K).

3. Results and discussion

To demonstrate the utility of POVME, we used the new algorithm to study the pocket dynamics of three members of the matrix-metalloproteinase (MMP) family of proteins: MMP-2, MMP-3, and

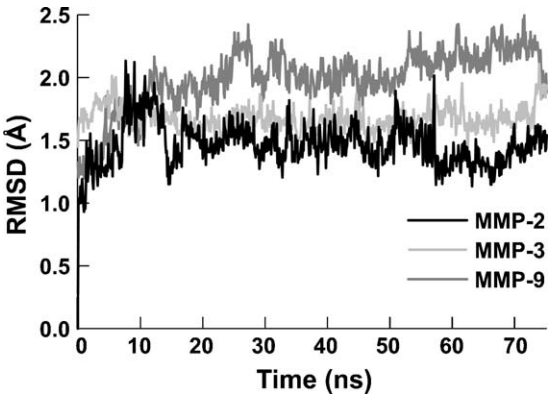

Fig. 2. The alpha-carbon RMSD between a fixed structure and the structures of the molecular dynamics simulations of MMP-2, MMP-3, and MMP-9, as a function of time. Relatively stable RMSD values suggest that these three systems are equilibrated and provide little evidence of differing dynamics.

MMP-9. MMPs are known to degrade components of the extracellular matrix and have been implicated in arthritis, multiple sclerosis, vascular disease [6], Alzheimer's disease [7], asthma [8–10], and cancer [11–13].

MMP-2, MMP-3, and MMP-9 all have similar structures; when crystallographic structures of the homologous regions of these three proteins are aligned and compared, the alpha-carbon RMS distance (RMSD) between the three is less than 1 Å in all cases (PDB IDs: 1QIB, 1B3D, and 1L6J) [3,5,14]. Despite these similarities, these MMPs have distinct substrate specificities [15,16]. As they are structurally similar, some have hypothesized that differences in protein dynamics are responsible for this specificity [17].

To study the pocket-volume dynamics of these three matrix metalloproteinases, we performed a molecular dynamics simulation of each. Following simulation, the homologous regions of all protein structures were aligned to a common structure, and the alpha-carbon RMSD between that common structure and each of the remaining frames was calculated (Fig. 2). As expected, each simulation was well equilibrated. Surprisingly, the data offered little indication that these three proteins have significantly different dynamics.

While additional metrics like RMSF and PCA might have identified differences in protein flexibility, we decided to see if differences could be detected by analyzing the binding-pocket volumes of these three MMPs over time. Frames were extracted from the molecular dynamics simulations at regular intervals of 10 ps, producing representative sets of 7550 protein structures for each simulation. After aligning all frames, a single volume could be defined that encompassed all binding pockets (Fig. 1a). The POVME algorithm was then used to calculate the binding-pocket volume of each structure.

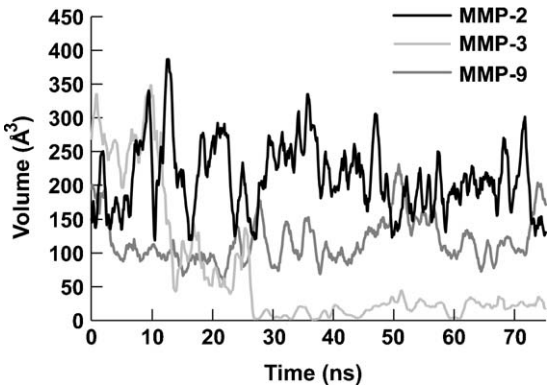

Fig. 3. The volumes of the binding pockets of MMP-2, MMP-3, and MMP-9, as a function of time. Though similar in structure, these three proteins have significantly different binding-pocket dynamics.

Fig. 3 shows how the pocket volume changed over the course of each molecular dynamics simulation. Significant differences in the pocket size are immediately evident. Though both MMP-2 and MMP-9 are gelatinases, MMP-2 tends to have the larger pocket. Additionally, the pocket of MMP-3 undergoes an impressive collapse at roughly 27 ns (Video 1). This collapse is particularly surprising given that the RMSD measurements over the course of the simulation gave no indication of such structural instability (Fig. 2). While the possibility that the system had not fully equilibrated cannot be ruled out, given the extensive minimization and equilibration steps that preceded the productive simulation, the fact that the overall RMSD suggested stability, and the published evidence suggesting that these pockets are highly dynamic [17], we suspect this pocket-closing event may be physiologically relevant.

4. Conclusion

Herein we have presented a novel algorithm, called POVME, for measuring the volume of ligand-binding pockets. POVME has been implemented in Python and is easily editable, platform independent, fast, and open source. To demonstrate the utility of this new algorithm, we used POVME to show that MMP-2, MMP-3, and MMP-9, three matrix metalloproteinases that are structurally similar, in fact have dynamically different binding sites, potentially explaining the differences in their substrate specificities. POVME can be downloaded for free from http://www.nbcr.net/POVME/.

Acknowledgements

This work was supported in part by funding from NIH GM31749, NSF MCB-0506593, and MCA93S013 to JAM. Additional support from the Howard Hughes Medical Institute, the National Center for Supercomputing Applications, the San Diego Supercomputer Center, the W.M. Keck Foundation, the National Biomedical Computational Resource, the Center for Theoretical Biological Physics, and the NSF Supercomputer Centers is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.10.007.

References

[1] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph 14 (1996) 33–38.
[2] J.D. Durrant, C.A.F. de Oliveira, J.A. McCammon, Including receptor flexibility and induced fit effects into the design of MMP-2 inhibitors, J. Mol. Recognit. 23 (2010) 173–182.
[3] V. Dhanaraj, M.G. Williams, Q.-Z. Ye, F. Molina, L.L. Johnson, D.F. Ortwine, et al., X-ray structure of gelatinase A catalytic domain complexed with a hydroxamate inhibitor, Croat. Chem. Acta 72 (1999) 575–591.
[4] A.G. Pavlovsky, M.G. Williams, Q.Z. Ye, D.F. Ortwine, C.F. Purchase, A.D. 2nd, White, et al., X-ray structure of human stromelysin catalytic domain complexed with nonpeptide inhibitors: implications for inhibitor selectivity, Protein Sci. 8 (1999) 1455–1462.
[5] P.A. Elkins, Y.S. Ho, W.W. Smith, C.A. Janson, K.J. D'Alessio, M.S. McQueney, et al., Structure of the C-terminally truncated human ProMMP9, a gelatin-binding matrix metalloproteinase, Acta Crystallogr. D: Biol. Crystallogr. 58 (2002) 1182–1192.
[6] J. Hu, P.E. Van den Steen, Q.X. Sang, G. Opdenakker, Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases, Nat. Rev. Drug Discov. 6 (2007) 480–498.
[7] L.M. Coussens, Z. Werb, Matrix metalloproteinases and the development of cancer, Chem. Biol. 3 (1996) 895–904.
[8] D. Cataldo, C. Munaut, A. Noel, F. Frankenne, P. Bartsch, J.M. Foidart, et al., Matrix metalloproteinases and TIMP-1 production by peripheral blood granulocytes from COPD patients and asthmatics, Allergy 56 (2001) 145–151.
[9] G. Mautino, N. Oliver, P. Chanez, J. Bousquet, F. Capony, Increased release of matrix metalloproteinase-9 in bronchoalveolar lavage fluid and by alveo-

lar macrophages of asthmatics, Am. J. Respir. Cell Mol. Biol. 17 (1997) 583–591.
[10] K. Prikk, P. Maisi, E. Pirila, M.A. Reintam, T. Salo, T. Sorsa, et al., Airway obstruction correlates with collagenase-2 (MMP-8) expression and activation in bronchial asthma, Lab. Invest. 82 (2002) 1535–1545.
[11] P.D. Brown, Clinical studies with matrix metalloproteinase inhibitors, APMIS 107 (1999) 174–180.
[12] L.J. Denis, J. Verweij, Matrix metalloproteinase inhibitors: present achievements and future prospects, Invest. New Drugs 15 (1997) 175–185.
[13] S.M. Wojtowicz-Praga, R.B. Dickson, M.J. Hawkins, Matrix metalloproteinase inhibitors, Invest. New Drugs 15 (1997) 61–75.
[14] L. Chen, T.J. Rydel, F. Gu, C.M. Dunaway, S. Pikul, K.M. Dunham, et al., Crystal structure of the stromelysin catalytic domain at 2.0 A resolution: inhibitor-induced conformational changes, J. Mol. Biol. 293 (1999) 545–557.
[15] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30.
[16] J. Wixon, D. Kell, The Kyoto encyclopedia of genes and genomes—KEGG, Yeast 17 (2000) 48–55.
[17] C.A. de Oliveira, M. Zissen, J. Mongon, J.A. McCammon, Molecular dynamics simulations of metalloproteinases types 2 and 3 reveal differences in the dynamic behavior of the S1′ binding pocket, Curr. Pharm. Des. 13 (2007) 3471–3475.