

UC Riverside

UC Riverside Previously Published Works

Title

T-Analyst: a program for efficient analysis of protein conformational changes by torsion angles

Permalink

<https://escholarship.org/uc/item/9976q544>

Journal

Journal of Computer-Aided Molecular Design: Incorporating Perspectives in Drug Discovery and Design, 24(10)

ISSN

1573-4951

Authors

Ai, Rizi
Qaiser Fatmi, M.
Chang, Chia-en A.

Publication Date

2010-10-01

DOI

10.1007/s10822-010-9376-y

Peer reviewed

T-Analyst: a program for efficient analysis of protein conformational changes by torsion angles

Rizi Ai · M. Qaiser Fatmi · Chia-en A. Chang

Received: 10 March 2010 / Accepted: 26 July 2010 / Published online: 6 August 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract T-Analyst is a user-friendly computer program for analyzing trajectories from molecular modeling. Instead of using Cartesian coordinates for protein conformational analysis, T-Analyst is based on internal bond-angle-torsion coordinates in which internal torsion angle movements, such as side-chain rotations, can be easily detected. The program computes entropy and automatically detects and corrects angle periodicity to produce accurate rotameric states of dihedrals. It also clusters multiple conformations and detects dihedral rotations that contribute hinge-like motions. Correlated motions between selected dihedrals can also be observed from the correlation map. T-Analyst focuses on showing changes in protein flexibility between different states and selecting representative protein conformations for molecular docking studies. The program is provided with instructions and full source code in Perl.

Keywords Molecular dynamics · Entropy · Allosteric · Simulation · Drug screening · Docking · HIV protease · AMBER

Introduction

The conformational dynamics of proteins play important roles in their functions and regulating ligand binding. A fundamental appreciation of how proteins work requires study of conformations and dynamics, as well as changes between states of protein motions, such as folded/unfolded and ligand-bound/-free states. For example, protein allosteric

effects may be related by either or both conformational and dynamical changes [1–3]. Molecular dynamics (MD) simulations provide powerful tools for the exploration of the conformational energy landscape accessible to protein molecules because multiple conformations are difficult to probe experimentally [4–6]. Moreover, recent computer-aided drug discovery studies have focused on protein flexibility in molecular docking processes [7–9]. Since most docking software prefers proteins to be rigid to avoid intensive computational effort, a promising strategy is to dock ligands into protein conformation ensembles obtained from MD simulations [10, 11]. Several programs provide general tools or special modules for analysis of MD results and clustering conformations, but most are based on Cartesian coordinates [12–17]. T-Analyst uses internal bond-angle-torsion (BAT) coordinates, which are efficient in capturing side-chain rotamers and most low-frequency motions [18–20]. Our program provides useful tool to analyze MD trajectories. For example, users can easily view proper rotameric states of dihedrals from the output files instead of plotting and correcting them manually. This program allows for efficient analysis of MD simulations to study protein flexibility and extract structural information for virtual screening.

Method

Data manipulation and output

T-Analyst reads NAMD, Amber or CHARMM trajectory files and Amber topology file. The CHARMM-type topology files can be converted to Amber topology files easily with the freely available CHAMBER program [21]. The program implements amber2accent to transform Cartesian coordinates to BAT coordinates [22]. To eliminate repeats,

R. Ai · M. Qaiser Fatmi · C. A. Chang (✉)
Department of Chemistry, University of California Riverside,
Riverside, CA 92521, USA
e-mail: chiaenc@ucr.edu

only heavy-atom side-chain torsion angles and ω , φ and ψ angles in backbones are considered. Users can choose angles and residues for analysis, and a dihedral distribution and its rotations during a simulation are output as .agr-format files, which can be viewed with Xmgr/Grace. Moreover, results generated by T-Analyst can be visualized by freely available packages such as VMD, Xmgr/Grace or R [12, 23, 24]. Users can output a series of files, such as distributions of all or selected torsion angles in a protein, or sorted and/or unsorted standard deviations and entropy. T-Analyst also groups different conformations based on rotameric states of residues of interest and outputs coordinates of grouped conformations into different trajectory files. The program also computes pairwise cross-correlation coefficients for all pairwise dihedral angles that users selected. By examining the output correlation map, dihedrals that correlate with each other can be observed.

Torsion angle correction

To correctly reflect rotameric states and capture the angle periodicity, T-analyst detects the angle population in margins, such as $\pm 180^\circ$ and $360^\circ/0^\circ$, depending on how users output the angles. If discontinuity in a dihedral

distribution is detected (see Fig. 1a, c), the output angle range will be corrected; thus, energy wells can be illustrated and defined properly (Fig. 1b, d). For example, Fig. 1c illustrates a rotameric state of side-chain Ile 153 of ligand-bound tryptophan synthase (TRPS), whereby one energy well is split into two wells near -180° and $+180^\circ$. T-Analyst automatically detects the peak discontinuity at each edge of -180° and $+180^\circ$ and determines the range of each discontinued peak. Later, the whole population will be divided into left set (-180° to 0°) and right set (0° to $+180^\circ$) and be counted in each set. T-Analyst angle correction is then applied by moving the discontinued peak on the set with smaller population to the set with larger population. For example, in Fig. 1c, the set from -180° to 0° has the smaller population, so the peak near the -180° margin was shifted by $+360^\circ$ and merged with the other margin at $+180^\circ$.

Configurational entropy calculation and dihedral correlation

T-Analyst calculates configurational entropy, S_{conf} , for each torsional degree of freedom by the Gibbs entropy formula: $TS_{\text{conf}(i)} = -RT \sum P_i \ln(P_i)$, where P_i is the

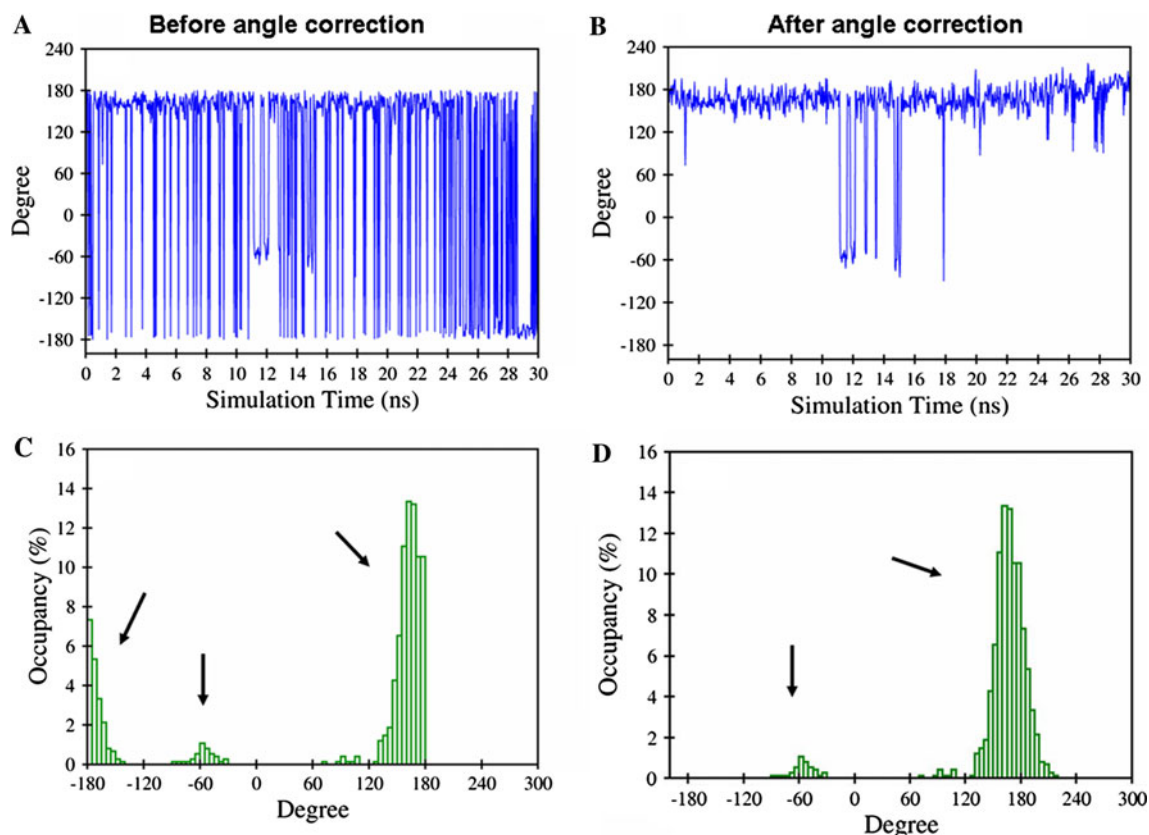


Fig. 1 Torsion angle distribution of side-chain χ^2 of Ile153 of TRPS. Torsion angle: **a** before angle correction and **b** after angle correction. Histogram of torsion angle distribution output by T-Analyst: **c** before angle correction and **d** after angle correction

probability distribution of angle i , R is the gas constant and T is the absolute temperature. T-Analyst calls Xmgr/Grace to generate histogram for each degree of freedom. The bin size for each P_i is 0.5° for ω angle, 1° for φ and ψ angles and 5° for side-chain dihedrals. The value of $TS_{\text{conf}(i)}$ has unit of kcal/mol, which allows for direct comparison with energy calculations. Summing $TS_{\text{conf}(i)}$ provides a quick approximation of the entropic contribution of a system, although coupling between torsions is ignored [25].

T-Analyst also computes the quasi-harmonic approximation (QH) from BAT coordinates. The covariance matrix C can be computed, with the probability distribution functions approximated by a multidimensional Gaussian distribution function [26]. The configurational entropy from QH is computed by $TS_{\text{QH}} = 1/2 nRT + 1/2 RT \ln[(2\pi)^n \det(C)]$, where n is the number of torsions. Although QH assumes that the probability distribution function is Gaussian, which is accurate for torsions that have only one rotameric state, TS_{QH} provides an upper bound limit for the configurational entropy [25, 27]. Moreover, the off-diagonal elements of the covariance matrix indicate the degree of significance of the coupling between the given torsions. Entropy computed from only the diagonal elements of the covariance matrix, $TS_{\text{QH_diag}}$, is also computed. If TS_{QH} equals $TS_{\text{QH_diag}}$, then there is no coupling among these torsions.

The extent to which pairs of dihedrals are correlated with one another can be assessed by examining the magnitude of their cross-correlation coefficients. T-Analyst computes a correlation matrix of dihedrals and calls levelplot function in R to plot a correlation map. Users can select dihedral angles, e.g. backbone φ and ψ angles of selected residues, to plot a correlation matrix. Typical characteristics of a correlation map include a line of strong cross-correlation along the diagonal (where matrix element $i = j$), and off-diagonal cross-correlations. The high diagonal values are set to 1.00. Off-diagonal correlations can be either positive or negative, and non-zero values may indicate potentially interesting correlations between two close proximity or non-contiguous regions of a protein system.

Clustering protein conformations

Our program clusters protein conformations on the basis of user-selected rotameric states of residues. Although RMSD-based clustering methods are mostly applied to group conformations with significant differences, small fluctuations are challenging to detect with classical RMSD-based clustering methods. This module is particularly useful for choosing representative conformations based on side-chain rotations of key residues. Users can input specific torsions with rotameric states of interest and the range of each torsion to run T-Cluster, the second part of T-Analyst. T-Analyst

will provide all the combinations of groups for further analysis. The program is sensitive to dihedral rotations and can efficiently group user picked backbone or side-chain dihedrals into separate trajectory files. A report file is also generated to record information about each group.

Molecular dynamics simulations

Molecular dynamics simulations on ligand-free and ligand-bound TRPS were performed using the NAMD package [23]. Standard simulation procedures (e.g. ([28])) were followed using the Amber 10 package and ff03 Amber force field and general Amber force field [6, 29]. Initial coordinates of TRPS were taken from PDB code 2J9X and 3CEP [30, 31]. Briefly, after preparation of the system by sequential steps of energy minimization and equilibration, the 30 ns production runs were carried out at 298 K and 1 atm. The systems were solvated by a 12 Å TIP3P water box. Snapshots of the atomic coordinates were recorded every 1 ps. As T-Analyst does not require too many frames to run the analysis, snapshots were saved every 20 ps for T-Analyst and 1,500 frames were used. Molecular dynamics simulations on HIV-1 protease were initiated from crystallographic coordinates with a semi-open flap conformation (PDB code 1HHP) [32]. Amber ff99 force field was used for the protein. Aqueous solvation was modeled implicitly by using the Generalized Born approach [33] and temperature was maintained at 298 K by using Langevin dynamics. Standard simulation procedures (e.g. [34]) were followed with the Amber 9 package. Since the free protease predominantly populates the semi-open conformation, we took a 1.5 ns MD simulation and saved it as one 1,500 frame trajectory which had one flap open state for our analysis.

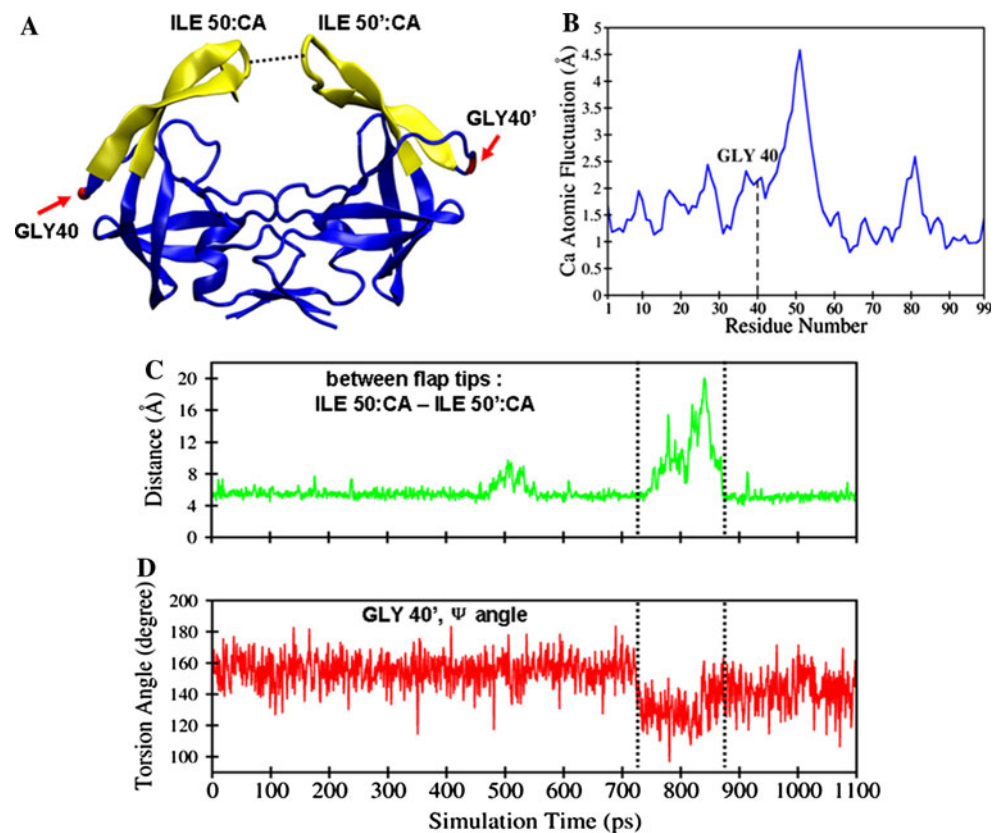
Analysis of MD simulations

Analysis of protein conformations

Analysis of small backbone fluctuations, as well as conformational changes, involves investigating loop or side-chain motions. T-Analyst adopts torsion angle analysis, which allows for accurate expression of bond rotations, which allows for accurate expression of bond rotations. For example, Fig. 1a shows a rotameric state of side-chain Ile 153 of ligand-bound TRPS. Standard deviations of this torsion before and after angle correction are 136.4° and 49.4° , respectively. Large differences in standard deviations usually indicate changes in rotameric states. Of note, proper angle correction is necessary for computing accurate rotamers and their standard deviations.

With careful superposition, flexible regions in proteins can be revealed by computing root mean square deviation and fluctuation (RMSD and RMSF, respectively) in

Fig. 2 Analysis of the hinge regions of HIV-1 protease. **a** An open conformation of HIV-1 protease obtained from the MD simulation. Flaps: residues 43–58, yellow; elbows: residues 37–42; hinge: Gly 40, red. **b** RMSF of C α atoms. **c** Distance between two flap tips: residues Ile 50 and Ile 50'. **d** Ψ angle distribution of Gly 40'; values within the dashed lines, between 720 and 880 ps, correspond to open flap conformations shown in **c**



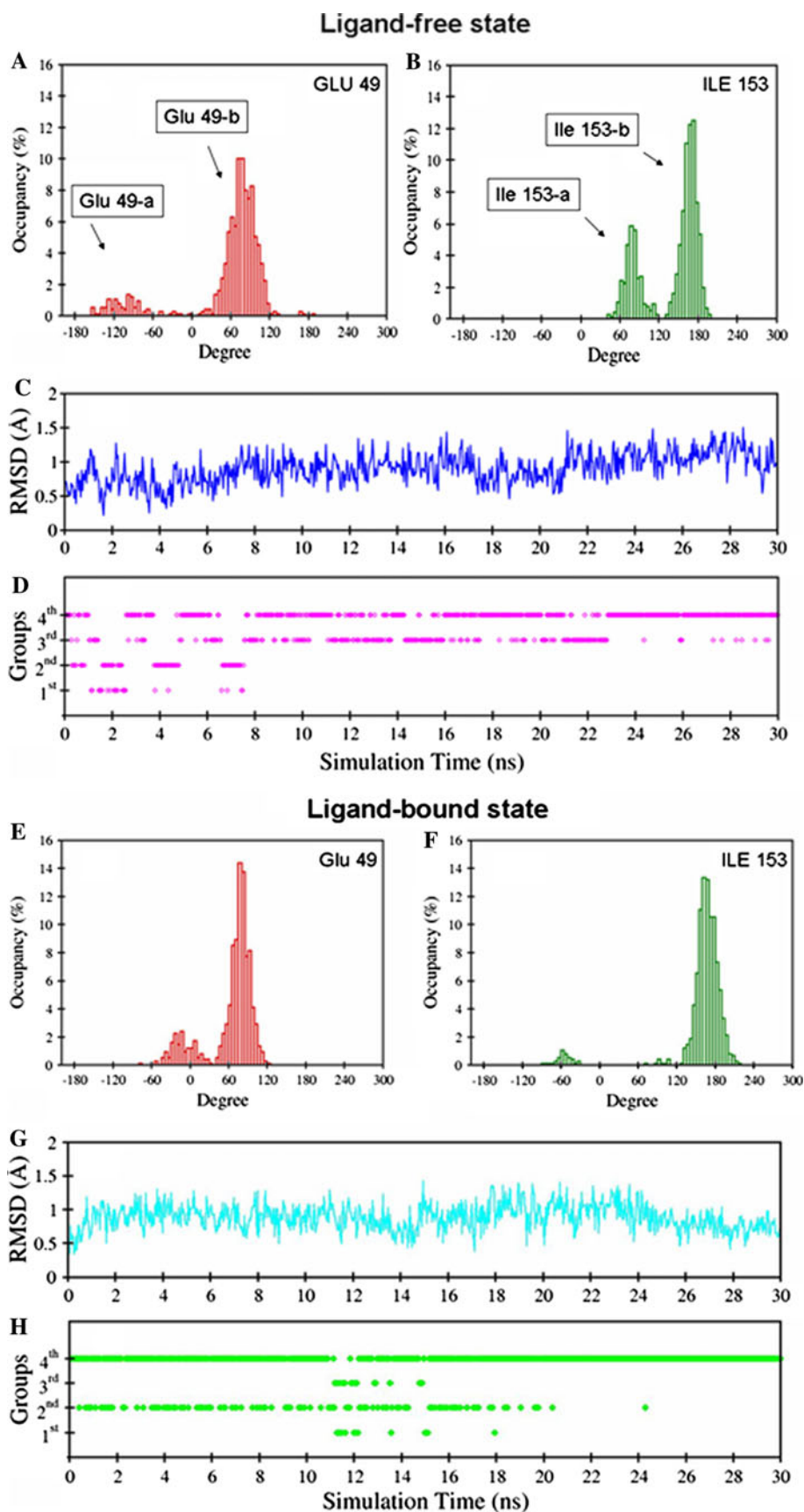
Cartesian coordinates [35]. Sometimes one or a few residues may serve as “hinges,” which mainly control movements of a flexible region. While the RMSD/RMSF method may not be sensitive enough to detect hinges, T-Analyst can illustrate torsion angles in hinges, which are associated with protein motions. User can compare multiple torsion distributions with protein motion to detect the hinge region. Figure 2a shows MD simulations of HIV-1 protease with open flaps, and the distance between flap tips is in Fig. 2c. Highly flexible regions such as flaps and elbows are shown in the RMSF plot (Fig. 2b) but determining whether any residues behave as hinges is not straightforward [34]. Data from T-Analyst imply that the flap β -hairpins move relatively rigidly, but some residues of the elbow, such as Gly 40, act as hinges (Fig. 2d).

One common method to express protein flexibility is by showing their rotameric states computed from corrected torsion angle distributions. When comparing the rotameric states for proteins between different states, e.g. ligand-bound/-free states, folded/unfolded states, users can have valuable information regarding the protein conformational or flexibility changes between different states. Different ligand mechanisms, such as induced fit or conformational selection (population shift) can be studied [1, 36].

Here we use TRPS, a protein with strong allosteric regulation of substrate binding, as an example. In both ligand-free and ligand-bound TRPS, the major peak of Glu 49

remains for both states to keep its favorite rotameric state (Fig. 3a, e), which is often assumed [37]. However, instead of dropping the minor peak, the peak changed while the substrate binds. We found that the change may be due to different hydrogen-bonds formations. In the ligand-bound state, Glu 49 form stable hydrogen-bonds with ligand 3-indole-D-glycerol-3'-phosphate (IGP), Tyr 173 and Val 23, and a residue near Glu 49, Tyr 175, also forms a hydrogen-bond with IGP. These interactions stabilize the local environment for the ligand which is necessary for catalysis. In the ligand-free state, since there is no ligand to interact with the protein in the binding site, more free space is allowed for side-chains to move around, though less stable hydrogen-bonds between side-chains can still formed. Without the ligand, the hydrogen-bond between Glu 49 and Tyr 173 is absent and Tyr 173 flips away from the binding site. Further detailed analysis shows that the distance between Tyr173:OH and Glu49:CD changes from 3.6 ± 0.16 Å in ligand-bound state (red dashed lines in Fig. 4a) to 17.5 ± 0.63 Å in ligand-free state (red dashed lines in Fig. 4b). As a result, instead of completely diminishing one rotameric state of Glu 49, the interaction induces a new rotameric state to alter the local environment to accommodate the ligand. Similarly, the major peak of residue Ile 153 in Fig. 3b, f remains in the same position in both states, but it is more populated in the ligand-bound state. Upon IGP binding, the methyl group of Ile 153 interacts with the

Fig. 3 Clustering conformations of TRPS. **a–d** Ligand-free state, **e–h** ligand-bound state. Torsion angle distribution of side-chain χ^3 of Glu 49 in **a** ligand-free and **e** ligand-bound state. Torsion angle distribution of side-chain χ^2 of Ile 153 in **b** ligand-free and **f** ligand-bound state. **c, g** RMSD of Glu 49 and Ile 153 after structure alignment. **d, h** conformational groups computed by T-Analyst over the course of MD simulations. Each *small diamond* represents one frame



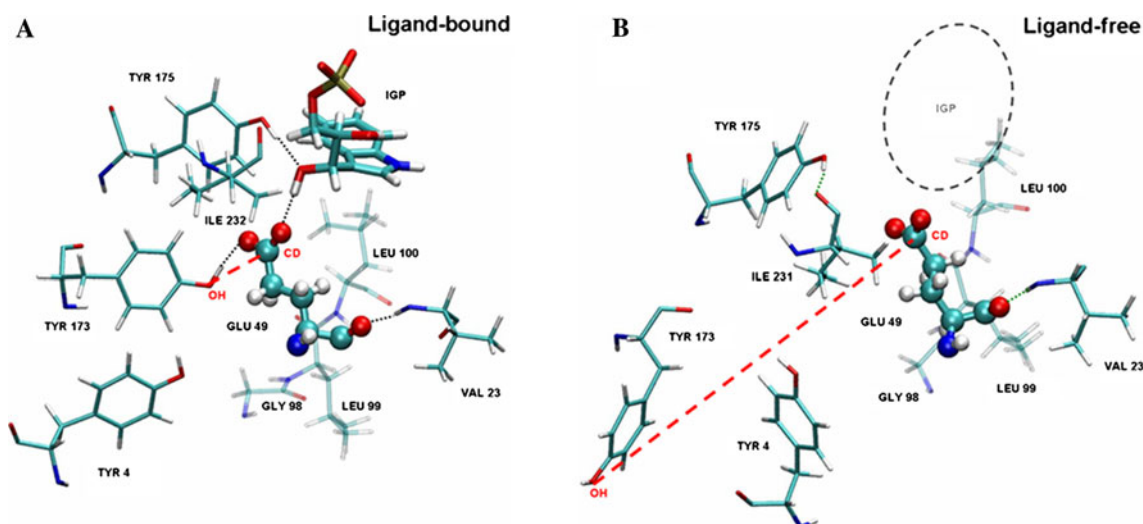


Fig. 4 Snapshots of surrounding residues of Glu 49 in **a** ligand-bound and **b** ligand-free state. Grey dotted lines indicate stable hydrogen-bonds in **a** and green dotted lines indicate weak hydrogen-

non-polar region of the ligand and stays in one peak. Therefore, the population of the rotameric states is shifted.

Clustering protein conformations

Docking potential ligands to target protein is one of the key steps in drug design and discovery process. During ligand–protein docking process, the position of side-chains in binding site can affect ligand–protein interaction significantly and directly, especially when using rigid protein conformations. Here we use the program to group conformations and then select representative conformations for performing docking studies to the α -subunit of TRPS. TRPS is a potential antibiotic target, and two torsion angles, Glu 49 and Ile 153, are known to directly involve in ligand–protein binding (see Fig. 3). Figure 3a, b show the distribution of side-chain torsion angles in both residues in ligand-free state which is output by T-Analyst after angle correction. Two populations are shown in each of the torsion angle

bonds in **b**. Red dashed lines indicate the distance between Tyr173:OH and Glu49:CD. The circle labeled with IGP in **b** indicates the position of the ligand binding pocket

distributions. T-Analyst can group the trajectories into four groups which are characterized by—group 1 (Glu 49-a, Ile 153-a), group 2 (Glu 49-a, Ile 153-b), group 3 (Glu 49-b, Ile 153-a), group 4 (Glu 49-b, Ile 153-b). Figure 3c, g show the RMSD distribution of Glu 49 and Ile 153 with simulation time. Figure 3d, h are the group distributions for the two torsion angles. Notably, although different protein conformations are usually clustered based on computed RMSD, the value is less sensitive to small scale conformational changes, such as side-chain rotations in a protein’s active site. As illustrated in Fig. 3c, g, there is no clue from RMSD that may be used for grouping. In contrast, T-Analyst also provides information regarding jumping between groups during a course of MD simulation, see Fig. 3d, h.

Configurational entropy

Protein allosteric effects or post-translational modifications such as phosphorylation do not always involve substantial

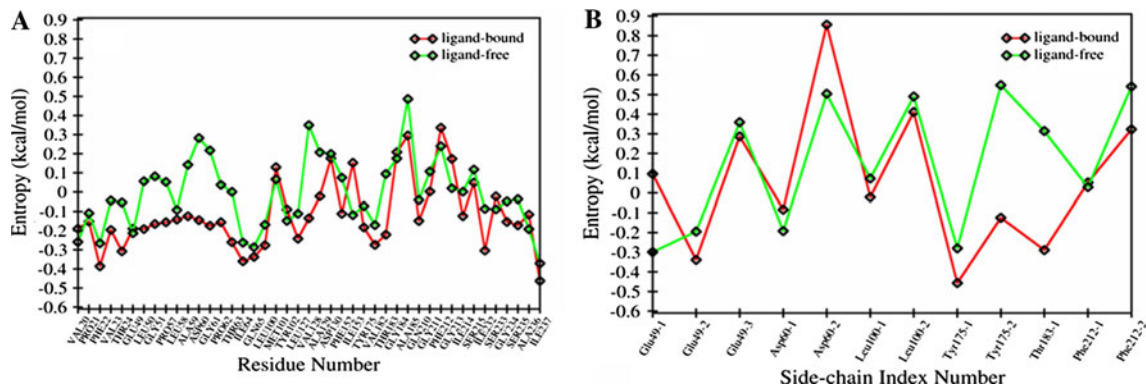


Fig. 5 Entropy distributions in the binding site of TRPS. **a** Entropy distribution of Ψ torsion angles of residues within 8\AA of ligand IGP and **b** side-chain torsion angles that directly interact with IGP

conformational changes. Recent experiments confirmed that in some cases, visual inspection of the active/inactive states may not reveal differences in the shape of the ligand binding site, but changes in protein dynamics [3]. Therefore, the magnitude of configuration entropy computed from dihedral degrees of freedom provides a direct way to examine protein flexibility.

The total entropy of a molecule in solution can be separated into two parts: solvent entropy associated with water motion and configurational entropy (S_{conf}) associated with molecular motion. The latter can be used to represent protein flexibility. Therefore, quantifying the configurational

entropy, especially changes, could help explain important biochemical processes such as protein folding and ligand–protein binding. Figure 5 shows S_{conf} calculated from selected backbone and side-chain dihedrals in the binding site of TRPS. T-Analyst suggested that the flexibility of a loop region changes significantly when the ligand binds into the pocket.

The calculations can provide quantitative information regarding entropy gain or loss. For example, in Fig. 5b, the side-chain entropies of Tyr 175 and Thr 183 drop notably, around 0.7 and 0.6 kcal/mol, respectively. In the ligand-bound state, Tyr 175 forms a stable hydrogen-bond with

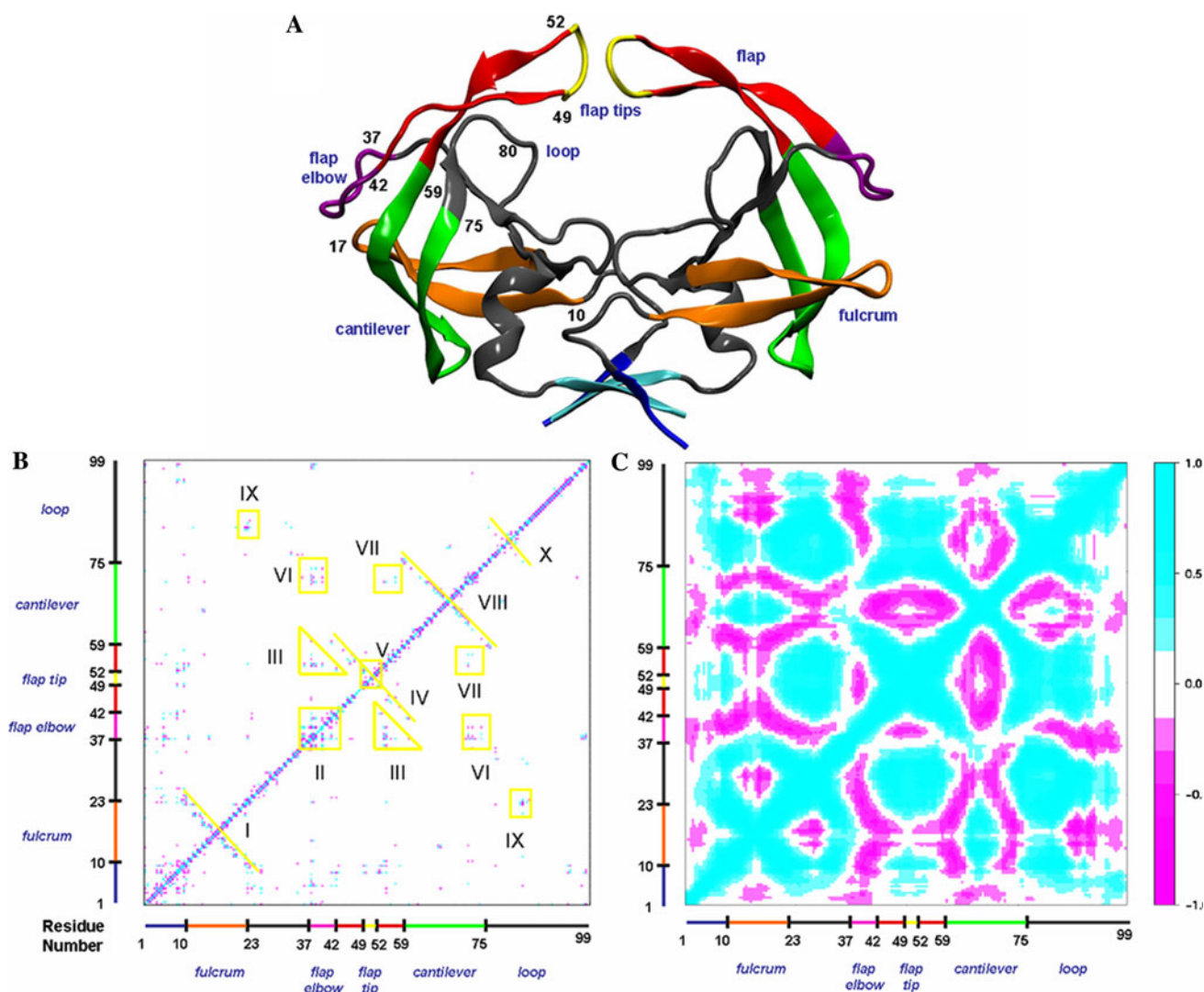


Fig. 6 Structure anatomy of HIV-1 protease and correlation maps of one subunit of HIV-1 protease in BAT coordinates and Cartesian coordinates. **a** Structure anatomy of HIV-1 protease. Color indicates distinct regions. Flaps: residues 43–58, red; flap tips: residues 49–52, yellow; flap elbows: residues 37–42, magenta; cantilevers: residues 59–75, green; fulcrums: residues 10–23, orange; loop area near flaps: residues 76–84, grey; and interleaved β -strand motif forming the dimer interface: residues 1–4 and 96–99, blue/cyan. **b** Φ and Ψ

torsion angle correlation map generated by T-Analyst using BAT coordinates. Yellow lines and boxes indicate correlations between I: fulcrum; II: flap elbow, III: flap with flap elbow, IV: flap, V: flap tip, VI: flap elbow with cantilever, VII: flap with cantilever, VIII: cantilever, IX: fulcrum with loop, and X: loop. **c** Correlation map generated by the use of the backbone $C\alpha$ and N atoms in the Cartesian coordinates with the Bio3D package

ligand IGP which stabilized the phenol ring, and Thr 183 also forms hydrogen-bonds with Asp 60 and Gly 61. In ligand-free state, both side-chains move more freely as there is no stable hydrogen-bond formed.

Although in most cases, one can observe entropy decrease upon ligand binding, there may be some exceptions which are of particular interest and worth further analysis. The entropy changes computed by T-Analyst provide guidance for users to pick up regions of a protein to do detail dynamic analysis. For example, T-Analyst showed an entropy gain in the second side-chain of Asp 60 in the presence of ligand, though not largely. Based on the information, we carried out further investigation near Asp 60, and found that two oxygen atoms in carboxyl group can form hydrogen-bonds with Tyr 102 and Thr 183 in both ligand-bound and ligand-free states. Interestingly, in the ligand-bound state, the two oxygen atoms of Asp 60 can flip very often but retain two stable hydrogen-bonds with Tyr 102 and Thr 183 alternatively. The presence of IGP provides a more hydrophilic environment around the carboxyl group of Asp 60; thus, the oxygen atoms can flip more freely. As a result, the local entropy increases without losing hydrogen-bonding and reducing electrostatic attraction. In contrast, when IGP is absent, the local environment around Asp 60 is mainly hydrophobic. Therefore, Asp 60 forms hydrogen-bond with Thr 183, the other hydrophilic residue, and the residues become less flexible.

Dynamical cross-correlation between dihedral angles

T-Analyst also performs cross-correlation analysis of a trajectory. The resulting cross-correlation map allows the identification of the correlated and anti-correlated motions involved in an entire protein or user selected dihedrals. Similar to calculating standard deviations, dihedral angles used to generate a correlation plot also need to run angle correction, or discontinuities in margins ($\pm 180^\circ$ or $360^\circ/0^\circ$) can cause errors when computing their cross-correlation coefficients.

Here we use HIV-1 protease as a test case, which continues to be one of the primary targets of anti-AIDS drug discovery. Understanding the dynamics of free HIV-1 protease has profound implications for designing new therapeutic agents, such as allosteric inhibitors. Figure 6a shows the structure anatomy of the protein. Yellow boxes in the correlation map, computed by the backbone φ and ψ angles with T-Analyst (see Fig. 6b), show the highly correlated regions within the HIV protease monomer. For example, boxes I, IV, VIII and X indicate flexible and concerted motions within fulcrum, flap, cantilever and loop regions. Other boxes highlighted in Fig. 6b indicate the correlations between flap elbow, flap tip, cantilever, fulcrum and loop regions. It has been suggested that motions

of the flap tip and elbow are highly correlated, which are illustrated in our plot as well [34]. Interestingly, correlations between the flap elbow and cantilever are identified (Box VI), and the motions of flap tip and cantilever are also correlated (Box VII), even though not quite strongly. We compared our plot to another correlation map generated by the use of the backbone $C\alpha$ and N atoms in the Cartesian coordinates with the Bio3D package (Fig. 6c) [38]. Both plots suggested similar correlations. However, because of the nature of Cartesian coordinates, patterns of correlations are harder to distinguish in Fig. 6c.

Conclusions

MD simulations provide invaluable conformational and dynamical landmarks useful for designing new experiments and for theoretical studies. The current analysis method we describe, T-Analyst, can help exam protein motion, identify structural and dynamic features, reveal changes of flexibility in different states, and group conformations based on dihedral rotamers. Analyzing the growing MD data may be the most time-consuming step in simulation studies, and our program facilitates this work. The program can be freely downloaded from <http://research.chem.ucr.edu/groups/chang/tools.htm>.

Acknowledgments We thank Dr. Thomas Steinbrecher for helping with amber2accent.pl. This work was supported by start-up funds and the Regents Faculty Fellowship from the University of California, Riverside, and the National Science Foundation (MCB-0919586).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Cui Q, Karplus M (2008) *Protein Sci* 17(8):1295
2. Kern D, Zuiderweg ERP (2003) *Curr Opin Struct Biol* 13(6):748
3. Tsai CJ, del Sol A, Nussinov R (2008) *J Mol Biol* 378(1):1
4. Karplus M, McCammon JA (2002) *Nat Struct Biol* 9(9):646
5. Gumbart J, Wang Y, Aksimentiev A, Tajkhorshid E, Schulten K (2005) *Curr Opin Struct Biol* 15(4):423
6. Deng YQ, Roux B (2009) *J Phys Chem B* 113(8):2234
7. May A, Sieker F, Zacharias M (2008) *Curr Comput Aided Drug Des* 4(2):143
8. Wong CF (2008) *Biochimica Et Biophysica Acta-Proteins Proteomics* 1784(1):244
9. Totrov M, Abagyan R (2008) *Curr Opin Struct Biol* 18(2):178
10. Carlson HA, McCammon JA (2000) *Mol Pharmacol* 57(2):213
11. Amaro RE, Baron R, McCammon JA (2008) *J Comput Aided Mol Des* 22(9):693
12. Humphrey W, Dalke A, Schulten K (1996) *J Mol Graph* 14(1):33
13. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caffisch A,

- Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) *J Comput Chem* 30(10):1545
14. Christen M, Hunenberger PH, Bakowies D, Baron R, Burgi R, Geerke DP, Heinz TN, Kastenhof MA, Krautler V, Oostenbrink C, Peter C, Trzesniak D, Van Gunsteren WF (2005) *J Comput Chem* 26(16):1719
 15. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) *J Comput Chem* 26(16):1701
 16. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) *J Comput Chem* 26(16):1668
 17. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. Volume 102
 18. Chang CE, Gilson MK (2003) *J Comput Chem* 24(16):1987
 19. McClendon CL, Friedland G, Mobley DL, Amirkhani H, Jacobson MP (2009) *J Chem Theory Comput* 5(9):2486
 20. Killian BJ, Kravitz JY, Somani S, Dasgupta P, Pang YP, Gilson MK (2009) *J Mol Biol* 389(2):315
 21. Crowley MF, Williamson MJ, Walker RC (2009) *Int J Quantum Chem* 109(15):3767
 22. Shao JY, Tanner SW, Thompson N, Cheatham TE (2007) *J Chem Theory Comput* 3:2312
 23. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. Volume 26
 24. R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>
 25. Chang C-EA, McLaughlin WA, Baron R, Wang W, McCammon JA (2008) *Proc Natl Acad Sci USA* 105(21):7456
 26. Levy RM, Karplus M, Kuschick J, Perahia D (1984) *Macromolecules* 17(7):1370
 27. Meirovitch H, Cheluvajara S, White RP (2009) *Curr Protein Pept Sci* 10(3):229
 28. Fatmi MQ, Ai R, Chang CEA (2009) *Biochemistry* 48(41):9921
 29. Wang JM, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. Volume 25
 30. Ngo H, Kimmich N, Harris R, Niks D, Blumenstein L, Kulik V, Barends TR, Schlichting I, Dunn MF (2007) Allosteric regulation of substrate channeling in tryptophan synthase: Modulation of the L-Serine reaction in stage I of the ss-reaction by alpha-site ligands. Volume 46
 31. Barends TRM, Domratheva T, Kulik V, Blumenstein L, Niks D, Dunn MF, Schlichting I (2008) *ChemBiochem* 9(7):1024
 32. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) *The Protein Data Bank*. Volume 58
 33. Onufriev A, Bashford D, Case DA (2000) *J Phys Chem B* 104(15):3712
 34. Hornak V, Okur A, Rizzo RC, Simmerling C (2006) *Proc Natl Acad Sci USA* 103(4):915
 35. Damm KL, Carlson HA (2006) *Biophys J* 90(12):4558
 36. Okazaki KI, Takada S (2008) Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms. Volume 105
 37. Kuhlman B, Baker D (2000) Native protein sequences are close to optimal for their structures. Volume 97
 38. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD (2006) Bio3d: an R package for the comparative analysis of protein structures. Volume 22