JOURNAL OF CHEMICAL INFORMATION AND MODELING

Letter

pubs.acs.org/jcim

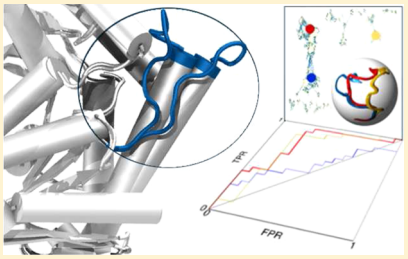
Collecting and Assessing Human Lactate Dehydrogenase-A Conformations for Structure-Based Virtual Screening

Rosa Buonfiglio,[†] Mariarosaria Ferraro,[‡] Federico Falchi,[†] Andrea Cavalli,[†,‡] Matteo Masetti,*[,†] and Maurizio Recanatini[†]

[†]Department of Pharmacy and Biotechnology, Alma Mater Studiorum—Università di Bologna, via Belmeloro 6, 40126 Bologna, Italy

[‡]Department of Drug Discovery and Development, Istituto Italiano di Tecnologia, via Morego 30, 16163 Genova, Italy

Supporting Information



ABSTRACT: Human lactate dehydrogenase-A (LDHA) is emerging as a promising anticancer target. Up to now, structure-based investigations for identifying inhibitors of this enzyme have not explicitly accounted for active site flexibility. In the present study, by combining replica exchange molecular dynamics with network and cluster analyses, we identified reliable LDHA conformations for structure-based ligand design. The selected conformations were challenged and validated by retrospective virtual screening simulations.

The energetic metabolism of normal differentiated cells depends primarily on oxidative phosphorylation. Conversely, most cancer cells show an upregulated anaerobic glycolytic pathway even in the presence of normoxic conditions. This metabolic shift, termed Warburg effect or aerobic glycolysis, is exploited by cancer cells to produce ATP and metabolites required to sustain a high biomass production.[1] The increased glucose uptake and (aerobic) glycolytic activity observed in cancer cells are due to the upregulation of key enzymes operated by several oncogenes such as *myc*, *ras*, and *akt*, as well as the hypoxia-inducible transcription factor HIF-1α. These enzymes protect cells from hypoxic conditions, often found in solid tumor lesions, and promote proliferation and survival in adverse environments. As a consequence, an enhancement of vascularization and the optimization of sugar metabolism via the Warburg effect result in a more malignant and aggressive tumor phenotype. Among these enzymes, lactate dehydrogenase (LDH, EC 1.1.1.27) plays a pivotal role in the metabolism of most tumors, representing an important means of NAD^+ regeneration essential to continue glycolysis.[2]

LDH is an oligomeric enzyme that catalyzes the NADH-dependent conversion of pyruvate to lactate.[3] Being assembled as a tetramer, five isoforms are conceivable basing on two different subunits, M and H (also known as A and B, respectively). LDHA (or M_4) is mainly found in skeletal muscle, whereas LDHB (or H_4), is located in heart, kidney, and brain. By ensuring energy production irrespectively of oxygen supply, LDHA acts as a tumor promoter, and for this reason, in recent years, it has emerged as a promising anticancer target.[4]

From a structural standpoint, the LDH tetramer shows a D_2 symmetry and is constituted by 330 amino acids per subunit (see Figure S1a, in the Supporting Information, SI). In each monomer, the Rossman-fold domain defines the NADH-cofactor binding site, while, in proximity of the nicotinamide pocket, a mixed α/β structure forms the substrate binding site. The latter is shielded from the solvent by a mobile loop (residues 98–110, hereafter referred to as "active site loop"), including the highly conserved Arg105 directly involved in catalysis.[3] The catalytic mechanism consists of an ordered sequence of events occurring on a millisecond time scale, and the loop motion has been hypothesized to be the rate limiting step of this mechanism.[5]

Despite the emerging role of LDHA in anticancer treatment, very few inhibitors have been reported so far in the literature. Recently, compounds showing an inhibitory activity in the micro-[6,7] and nanomolar[8,9] ranges have been identified by using structure-based drug design approaches. However, most of these computational studies have considered the "closed" conformation of the active site loop and, therefore, have not fully exploited the expected conformational flexibility of the LDHA binding site. One can actually argue that neglecting the conformational flexibility of the active site loop can hamper a proper exploitation of both structural information and computational tools in LDHA inhibitor research.

It has been reported that taking into accout protein flexibility can improve virtual screening (VS) predictivity in the search for novel and structurally diverse inhibitors.[10] In this context, we here aimed at developing a computational strategy suited to sample, classify, and validate LDHA conformations to take advantage of the enzyme's intrinsic flexibility in structure-based inhibitor design.

The unit cell of the crystal structure of human LDHA (PDB ID: 1I10) consists of two tetramers (A–D and E–H) including NADH and oxamate, a substrate mimicking inhibitor.[3] Interestingly, while monomers A, B, and C show a closed active site loop conformation, chain D displays an unusual open conformation. Not surprisingly, the active site loop in the open

Published: October 19, 2013

   dx.doi.org/10.1021/ci400543y | *J. Chem. Inf. Model.* 2013, 53, 2792–2797

state is highly disordered, yielding the direct inclusion of such conformation in a VS campaign a tempting but unwise approach. It is indeed well-known that the VS performances are in general highly sensitive to the quality of the starting structure.[11] To take advantage of the open loop conformation, and possibly to collect further potentially useful structures, a robust procedure of sampling and analysis must be employed.

Several approaches have been proposed to address the issue of target protein flexibility.[12] For example, selected degrees of freedom of the protein can be included in the search algorithm of docking methods. According to another strategy, an ensemble of conformations generated either by molecular dynamics (MD) or Monte Carlo (MC) techniques can be used as discrete representatives of the protein in a sequential docking exercise (ensemble docking). In the context of a VS framework, the latter approach is often referred to as an ensemble-based VS (EBVS).[13] In this case, the possibility to sample relevant conformations strongly depends on the relaxation time of the conformational transition. Whenever slow motions are involved, standard sampling techniques become unproductive, and the conformational transition is considered to be a rare event compared to the time-scales accessible by simulations. Several workarounds have been conceived to overcome this limitation, including the use of high performing hardware or specialized sampling techniques, also known as enhanced sampling methods.[14] Among these, replica exchange MD (REMD)[15] provides a general and remarkably versatile framework to sample rare events. Within the REMD formalism, noninteracting copies of the system under investigation (replicas) are simulated in parallel at a different nominal temperature. At fixed time intervals, exchanges between replicas are attempted according to a Metropolis MC criterion. In this way, high energy barriers are efficiently crossed and the correct Boltzmann distribution of states at each temperature preserved. Here, the large-scale rearrangement of the active site loop involved in substrate recognition was sampled by using a hybrid solvent variant of the REMD method (hREMD).[16]

For sake of computational efficiency, instead of simulating the whole LDHA tetramer, a minimal model preserving the essential dynamical features of the complex was identified and validated by means of preliminary explicit solvent MD simulations. Indeed, experimental evidence supports the idea that the tetramer showing enzymatic activity contains at least two native subunits, and that a dimer can actually represent the minimal functional unit of the enzyme.[17] The dimeric protein consisting of chains A and B was therefore used as a reference to define a smaller system with a comparable stability. The final working model consisted of chain A plus α-C helix of monomer B (see Figure S1b and SI for further details). The stability of the minimal LDHA model compared to both the AB dimer and chain A alone is shown in SI Figure S2.

The protein conformational ensemble was therefore generated carrying out hREMD simulations with AMBER11,[18] using the parm99SB-ILDN force field.[19] Twenty-four replicas spanning a geometrically distributed temperature range of 200 K (from 300 to 500 K) were employed, and a solvation shell of 700 water molecules around the solute was preserved during MC exchanges. To improve the sampling, two hREMD simulations were performed starting from both a closed and an open active site loop conformation for a total aggregate time of almost 700 ns (further details are reported in the SI). The trajectories coming from the physical replicas were then carefully analyzed to characterize the most relevant loop conformations.

Cluster analysis has been extensively used to classify configurations sampled by MD simulations. However, it is well-known that conventional cluster algorithms suffer from stability issues.[20,21] In addition, cluster analysis does not provide per se any useful interpretation of the phenomenon under investigation, which is indeed of vital importance when processing long MD simulations or trajectories resulting from enhanced sampling methods. Alternatively, one can try to reduce the dimensionality of data, using approaches including covariance analysis[22] or multidimensional scaling.[21] Further recent strategies in this context have been based on network analysis (or discontinuity graph approach), rather than relying on continuous projections.[23,24] The network concept is currently emerging as a useful tool to analyze protein dynamics, providing a visual and therefore intuitive classification of conformations otherwise difficult to be obtained by classical cluster analysis or conventional dimensionality reduction tools. Indeed, the network is perceived as a set of interacting elements (nodes), which are bound together by links (edges).

Here, network analysis was performed using the Cytoscape software.[25] The nodes of the network represented more than 6000 sampled configurations, and their similarity was described in terms of the RMSD metric by considering backbone atoms of the active site loop along with Glu101:CD, Glu103:CD, and Arg105:CZ (SI Figure S3). The inclusion of these side chain atoms in the similarity measure was required to properly trace the motion of these charged residues that is supposed to affect the loop geometry and consequently to influence EBVS performances. The network edges were built by connecting nodes having a pairwise RMSD below the threshold of 1.00 Å. Further details can be found in the SI.

The result of the application of the network analysis to the hREMD configurations is reported in Figure 1a. The network appears to be formed by a dense central core of nodes with sparsely branched islands. Interestingly, most of the sampled configurations were corresponding neither to the closed nor to

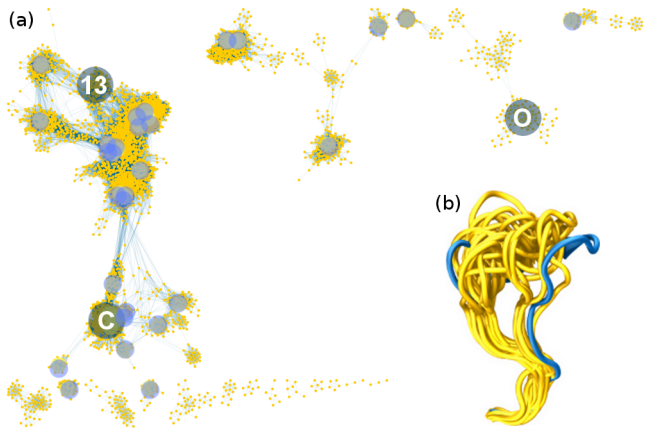


Figure 1. (a) Network representation of the sampled configurations. Nodes are represented as yellow dots, while links are shown as blue lines. The transparent blue circles correspond to location of the 30 collected seeds in the low dimensionality space obtained. Conformations C, O, and seed 13 are highlighted. (b) Active site loop conformations adopted by the 30 collected seeds (yellow tubes) together with conformation C and O (blue tubes on the left and on the right, respectively).

dx.doi.org/10.1021/ci400543y | J. Chem. Inf. Model. 2013, 53, 2792−2797

the open loop conformation, highlighting the central role of intermediate conformations in the winding mechanism of opening/closure of the active site loop. Nevertheless, the most interconnected regions of the network included configurations that were structurally related to the closed state. Conversely, the open loop conformation was found in an isolated island, and was linked to a few more nodes, reflecting its lower structural stability. The latter region also included some configurations with an even wider loop opening compared to the crystallographic open state conformation. Other sparse regions consisted of nodes without connections because of an unfavorable loop rearrangement. Some representative conformations are reported in SI Figure S4.

The reduced dimensionality of the conformational space was further rationalized by means of cluster analysis. This allows one to identify a manageable number of conformations for subsequent EBVS studies. The cluster analysis was here performed directly on the network analysis outcome, using the MCODE algorithm[26] (see SI for further details). Thirty clusters were obtained that spanned the majority of the network, especially in proximity of the core (Figure 1a), and all of them turned out to show high clustering score. Visual inspection of these clusters demonstrated that this method was effective enough to discriminate different loop conformations, and also adequately accurate in describing specific side chain orientations. In Figure 1b, the leaders of the first 30 clusters (also referred to as "seeds" according to the MCODE terminology), along with the crystallographic closed and open state conformations of the active site loop, are shown.

Finally, the 30 collected seeds were used in a retrospective VS study to validate the selected conformations and confirm that the use of multiple protein conformations can increase the enrichment of known binders. Besides these seeds, the closed conformation coming from the crystal structure (hereafter referred to as "C conformation") was included in the screening because of its unquestionable experimental significance. To stay on the safe side, a representative conformation of the open state was also considered. However, due to the aforementioned issues related to the crystal structure reliability, we decided to include the most similar representative to the open state conformation found among the entire ensemble of configurations (hereafter referred to as "O conformation"). The whole set of 32 conformations was then challenged in order to assess their ability to discriminate between known active and inactive compounds coming from the BindingDB[27] (active/inactive ratio: 21/17), and known active compounds and decoys generated by the DUD-E database[28] (active/decoy ratio: 21/1950). The 21 active compounds show an affinity for LDHA less than 30 μM in terms of either IC_{50} or K_i value (see SI Table S1). The choice of a high threshold value and the heterogeneity (i.e., IC_{50} or K_i values) of the data were due to the low number of available LDHA inhibitors and the fact that most of them show low potency against the enzyme. Both data sets were finally docked into the LDHA binding pocket using all the previously selected 32 conformations, which are available upon request by the authors. All docking calculations and analysis were performed with the Schrödinger suite[29] (see the SI for further details).

In SI Table S2, the figures of merit computed for both BindingDB and DUD-E data sets against each protein conformation are reported. Analysis of DUD-E screening results showed that all the LDHA conformations had comparable recognition capability. BindingDB docking results turned out to be more informative and deserved a more in depth investigation. As shown in Figure 2, a high recognition

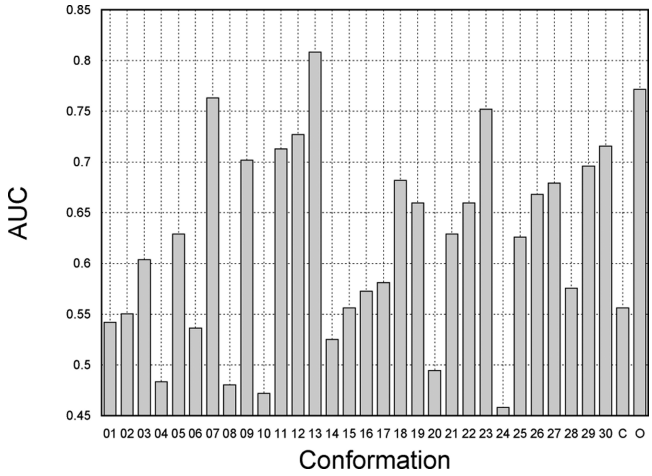


Figure 2. Area under the ROC curve calculated for the BindingDB data set using all the collected seeds plus the C and O conformations.

performance with an area under the ROC curve (here reported to as AUC) above 0.8 was achieved only by seed number 13, while C and O conformations showed lower AUC values (0.77 and 0.56, respectively). This result could be related to the different binding site volume displayed by these three conformations (see Table S3 and the SI for details). Seed 13, indeed, showed an intermediate loop conformation located between the open and the closed state (see Figure 3a), likely resulting in a more permissive structure, able to accommodate ligands of different size. The ROC curves for seed 13, C, and O are shown in Figure 3b, while the curves for the remaining conformations are reported in SI Figure S5. In SI Figure S6, a structural comparison between the best and the worst performing structure (seed 13 vs seed 24) is also reported. It is worth underlying that an interesting correlation in the recognition performance of BindingDB and DUD-E was observed. However, while O and seed 13 showed a high AUC value in both data sets, a dissimilar performance was obtained for C. Indeed, the latter early recognized large molecules along with small actives, compared to decoys. Nevertheless, some BindingDB inactives resulted well ranked probably because of overestimated favorable interactions promoted by a well-defined enclosed cavity.

In order to evaluate the enrichment in VS by considering a multiple protein conformation approach, ensemble-averaged, and Boltzmann-weighted averaged scores were calculated for all the ensembles consisting of 3 structures: C, O, and each of the 30 collected seeds. In Figure 3b, we report the ROC curve for the ensemble consisting of C, O, and seed 13, as one of the best combinations to increase enrichment. The inclusion of multiple protein conformations improved the overall performance of the VS, in agreement with previous studies in the field.[10] Other combinations provided similar results and are reported in the SI (Figure S7).

To further validate the usefulness of the selected conformations in view of prospect VS campaigns, three recently reported data set, not yet included in the BindingDB, were used.[9,30,31] These collections, hereafter referred to as Kohlmann, Dragovich, and Fauber data sets, consisted in an

dx.doi.org/10.1021/ci400543y | J. Chem. Inf. Model. 2013, 53, 2792−2797

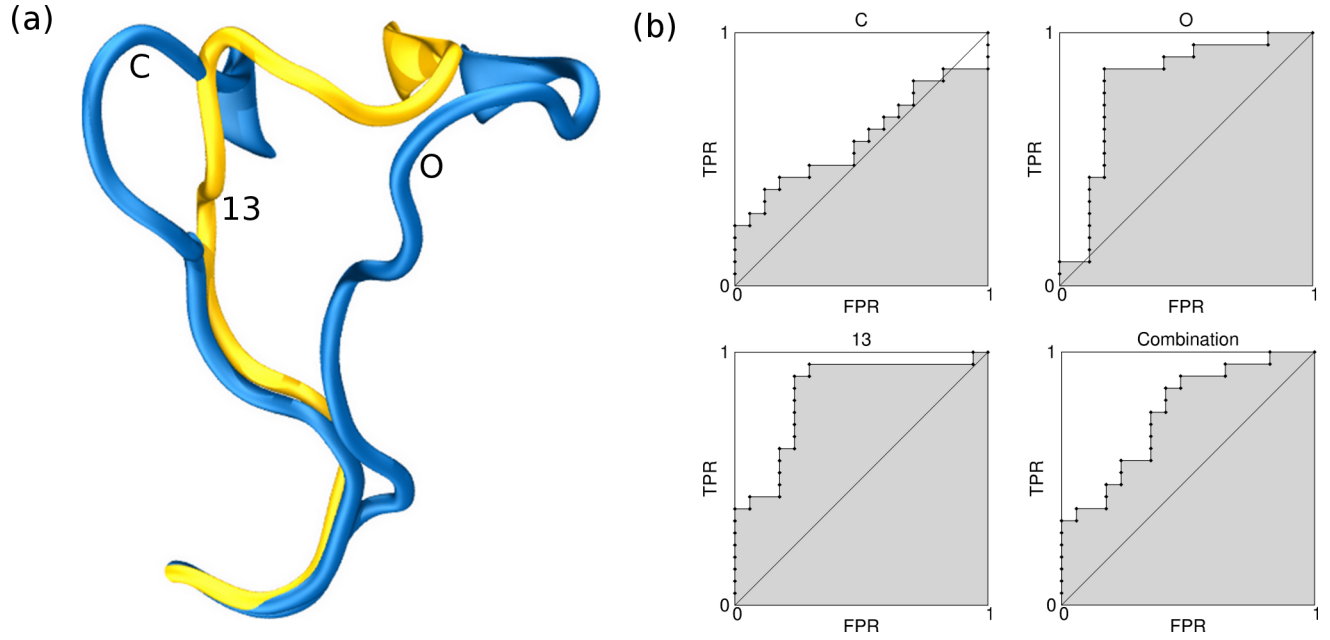


Figure 3. (a) Loop conformations for C, O, and seed 13. (b) ROC curves calculated for the BindingDB data set using conformations C, O, seed 13, and their combination. The ROC space is defined as true positive rate (TPR, *y* axis) versus false positive rate (FPR, *x* axis).

active/inactive ratio of 6/7, 5/11, and 18/15, respectively (racemic mixtures were not considered, see SI Table S1).

Compounds from the Kohlmann data set were obtained by fragment-based design, and their binding mode was characterized by X-ray crystallography.[9] Although the protein exhibited similar fold to the human LDHA/NADH/oxamate ternary complex, the active site loop was reported as disordered in the original paper. By challenging our collected LDHA conformations against this data set, AUC values spanning between 0.64 and 1.00, were recorded, with the highest recognition performance shown by O and C, while seed 13 displayed a lower value (see SI Table S4). The unsatisfying performance of seed 13 was related to the incorrect binding mode of two active compounds that were wrongly docked into the binding site because of the specific loop geometry and side chains orientation. Although this finding was not in line with BindingDB results, binding mode analysis and comparison with crystallized ligand−protein complexes highlighted that seed 13 reproduced well accommodated small molecules, such as the cognate ligand from PDB ID 4I8X. Moreover, bulky molecules (for example, PDB ID 4I9H) were properly located within the binding pocket, for C, O and seed 13, with the exception of the two molecules above-mentioned for this latter conformation.

As regards the Dragovich data set, novel LDHA inhibitors were identified and then optimized by high-throughput screening, and a crystal structure was obtained.[30] As described by the authors, the mobile loop resulted in a relatively open state, with Arg105 well displaced from the binding site. This data set confirmed O as one of the best structures in terms of AUC metric, followed by seed 13 among the best ranked conformations. Conversely, the poor recognition performance showed by C (see SI Table S4) could be related to the proximity of the Arg105 side chain to the substrate binding site, as found in this conformation. The agreement between docking results and crystallographic data confirmed the importance of a large pocket in accommodating this class of inhibitors.

Finally, the Fauber data set reported a new class of LDHA inhibitors identified via biochemical screening and validated using biochemical and biophysical assays.[31] A crystal structure of an optimized compound bound to the protein was also obtained. The AUC metric evaluated for the collected conformations against this data set showed seed 13 as the second best structure, followed by O and C (see SI Table S4). In this case, seed 13 confirmed to be one of the best conformations in recognizing active compounds, in agreement with BindingDB results. In terms of binding mode, most of these molecules interacted with conserved residues involved in catalysis, in line with the cognate ligand of the solved crystal.

In conclusion, by combining a sampling approach based on a REMD framework with network and cluster analyses, we collected several conformations of the LDHA binding pocket that were assessed for their effectiveness to discriminate between known active inhibitors, and known inactive compounds from BindingDB and *bona fide* inactive decoys. By doing so, not only we were able to characterize the most relevant conformations of the LDHA binding pocket at a fully atomistic description, highlighting salient features of the active site loop winding motion, but we also sorted out a protein conformation (seed number 13) endowed with a striking potential for successful prospective VS campaigns. Further validation of the collected conformations with more recent ligands confirmed C, O, or seed 13 among the best performing structures. While individually none of these conformations performed optimally in all the available data sets, we demonstrated that the use of multiple LDHA conformations is crucial to discriminate between active and inactive compounds belonging to diverse chemical classes. Even though our findings are mainly relevant in the context of binding sites flanked by large mobile loop moieties, we believe that the present strategy, besides boosting structure-based studies in the field of LDHA discovery, might also provide a general approach to include flexibility in EBVS studies.

dx.doi.org/10.1021/ci400543y | *J. Chem. Inf. Model.* 2013, 53, 2792−2797

■ ASSOCIATED CONTENT

Supporting Information

Simulation setups and analysis procedures. Additional references. LDHA tetramer and minimal model, Figure S1. Structural stability comparison, Figure S2. Atoms employed in the similarity measure, Figure S3. Example of representative collected conformations, Figure S4. ROC curves for BindingDB data set, Figure S5. Structural comparison between the best and worst performing collected conformations, Figure S6. Area under the ROC curve for the multiple protein conformation ensembles, Figure S7. Structures and activities of the 21 known inhibitors for the BindingDB, Kohlmann, Dragovich, and Fauber data sets, Table S1. Figures of merit for Binding-DB and DUD-E data set, Table S2. Binding site volume displayed by C, O, and seed 13, Table S3. AUC for Kohlmann, Dragovich, and Fauber data sets displayed by C, O, and seed 13, Table S4. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: matteo.masetti4@unibo.it.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to thank Angelo Favia for the stimulating discussions during the early stages of the work and Giovanni Bottegoni for useful suggestions. We are also grateful to Giuseppina Di Stefano and Luigi Fiume for valuable insights on LDH. We acknowledge the financial support of Alma Mater Studiorum—University of Bologna.

■ REFERENCES

(1) Vander Heiden, M. G.; Cantley, L. C.; Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **2009**, *324*, 1029−1033.

(2) Walenta, S.; Mueller-Klieser, W. F. Lactate: mirror and motor of tumor malignancy. *Semin. Radiat. Oncol.* **2004**, *14*, 267−274.

(3) Read, J. A.; Winter, V. J.; Eszes, C. M.; Sessions, R. B.; Brady, R. L. Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins* **2001**, *43*, 175−185.

(4) Yeluri, S.; Madhok, B.; Prasad, K. R.; Quirke, P.; Jayne, D. G. Cancer's craving for sugar: an opportunity for clinical exploitation. *J. Cancer. Res. Clin. Oncol.* **2009**, *135*, 867−877.

(5) Qiu, L.; Gulotta, M.; Callender, R. Lactate dehydrogenase undergoes a substantial structural change to bind its substrate. *Biophys. J.* **2007**, *93*, 1677−1686.

(6) Manerba, M.; Vettraino, M.; Fiume, L.; Di Stefano, G.; Sartini, A.; Giacomini, E.; Buonfiglio, R.; Roberti, M.; Recanatini, M. Galloflavin (CAS 568-80-9): a novel inhibitor of lactate dehydrogenase. *ChemMedChem* **2012**, *7*, 311−317.

(7) Granchi, C.; Roy, S.; Giacomelli, C.; Macchia, M.; Tuccinardi, T.; Martinelli, A.; Lanza, M.; Betti, L.; Giannaccini, G.; Lucacchini, A.; Funel, N.; Leon, L. G.; Giovannetti, E.; Peters, G. J.; Palchaudhuri, R.; Calvaresi, E. C.; Hergenrother, P. J.; Minutolo, F. Discovery of N-hydroxyindole-based inhibitors of human lactate dehydrogenase isoform A (LDH-A) as starvation agents against cancer cells. *J. Med. Chem.* **2011**, *54*, 1599−1612.

(8) Ward, R. A.; Brassington, C.; Breeze, A. L.; Caputo, A.; Critchlow, S.; Davies, G.; Goodwin, L.; Hassall, G.; Greenwood, R.; Holdgate, G. A.; Mrosek, M.; Norman, R. A.; Pearson, S.; Tart, J.; Tucker, J. A.; Vogtherr, M.; Whittaker, D.; Wingfield, J.; Winter, J.; Hudson, K. Design and synthesis of novel lactate dehydrogenase A inhibitors by fragment-based lead generation. *J. Med. Chem.* **2012**, *55*, 3285−3306.

(9) Kohlmann, A.; Zech, S. G.; Li, F.; Zhou, T.; Squillace, R. M.; Commodore, L.; Greenfield, M. T.; Lu, X.; Miller, D. P.; Huang, W. S.; Qi, J.; Thomas, R. M.; Wang, Y.; Zhang, S.; Dodd, R.; Liu, S.; Xu, R.; Xu, Y.; Miret, J. J.; Rivera, V.; Clackson, T.; Shakespeare, W. C.; Zhu, X.; Dalgarno, D. C. Fragment Growing and Linking Lead to Novel Nanomolar Lactate Dehydrogenase Inhibitors. *J. Med. Chem.* **2013**, *56*, 1023−1040.

(10) Bottegoni, G.; Rocchia, W.; Rueda, M.; Abagyan, R.; Cavalli, A. Systematic Exploitation of Multiple Receptor Conformations for Virtual Ligand Screening. *PLoS ONE* **2011**, *6*, e18845.

(11) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* **2005**, *49*, 5912−5931.

(12) Lill, M. A. Efficient Incorporation of Protein Flexibility and Dynamics into Molecular Docking Simulations. *Biochemistry* **2011**, *50*, 6157−6169.

(13) Amaro, R. E.; Li, W. W. Emerging Methods for Ensemble-Based Virtual Screening. *Curr. Top. Med. Chem.* **2010**, *10*, 3−13.

(14) Rocchia, W.; Masetti, M.; Cavalli, A. *Enhanced Sampling Methods in Drug Design Physico-Chemical and Computational Approaches to Drug Discovery*; Luque, J., Barril, X., Eds.; RSC Publishing, 2012.

(15) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(16) Okur, A.; Wickstrom, L.; Layten, M.; Geney, R.; Song, K.; Hornak, V.; Simmerling, C. Improved Efficiency of Replica Exchange Simulations through Use of a Hybrid Explicit/Implicit Solvation Model. *J. Chem. Theor. Comput.* **2006**, *2*, 420−433.

(17) Wang, X. C.; Jiang, L.; Zhou, H. M. Minimal functional unit of lactate dehydrogenase. *J. Protein Chem.* **1997**, *16*, 227−231.

(18) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER11*; University of California, San Francisco. 2010.

(19) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950−1958.

(20) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theor. Comput.* **2007**, *3*, 2312−2334.

(21) Rajan, A.; Freddolino, P. L.; Schulten, K. Going beyond Clustering in MD Trajectory Analysis: An Application to Villin Headpiece Folding. *PLoS ONE* **2010**, *5*, e9890.

(22) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential dynamics of proteins. *Proteins: Struct. Funct. Bioinf.* **1993**, *17*, 412−425.

(23) Mishra, S.; Caflisch, A. Dynamics in the Active Site of β-Secretase: A Network Analysis of Atomistic Simulations. *Biochemistry* **2011**, *50*, 9328−9339.

(24) Campbell, Z. T.; Baldwin, T. O.; Miyashita, O. Analysis of the Bacterial Luciferase Mobile Loop by Replica-Exchange Molecular Dynamics. *Biophys. J.* **2010**, *99*, 4012−4019.

(25) Smoot, M. E.; Ono, K.; Ruscheinski, J.; Wang, P. L.; Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **2011**, *27*, 431−432.

(26) Bader, G.; Hogue, C. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinf.* **2003**, *4*, 2.

(27) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a web-accessible database of experimentally determined

protein-ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198−D201.
(28) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582−6594.
(29) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(30) Dragovich, P. S.; Fauber, B. P.; Corson, L. B.; Ding, C. Z.; Eigenbrot, C.; Ge, H.; Giannetti, A. M.; Hunsaker, T.; Labadie, S.; Liu, Y.; Malek, S.; Pan, B.; Peterson, D.; Pitts, K.; Purkey, H. E.; Sideris, S.; Ultsch, M.; VanderPorten, E.; Wei, B.; Xu, Q.; Yen, I.; Yue, Q.; Zhang, H.; Zhang, X. Identification of substituted 2-thio-6-oxo-1,6-dihydropyrimidines as inhibitors of human lactate dehydrogenase. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3186−3194.
(31) Fauber, B. P.; Dragovich, P. S.; Chen, J.; Corson, L. B.; Ding, C. Z.; Eigenbrot, C.; Giannetti, A. M.; Hunsaker, T.; Labadie, S.; Liu, Y.; Liu, Y.; Malek, S.; Peterson, D.; Pitts, K.; Sideris, S.; Ultsch, M.; VanderPorten, E.; Wang, J.; Wei, B.; Yen, I.; Yue, Q. Identification of 2-amino-5-aryl-pyrazines as inhibitors of human lactate dehydrogenase. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 5533−5539.