JOURNAL OF CHEMICAL INFORMATION AND MODELING

Letter

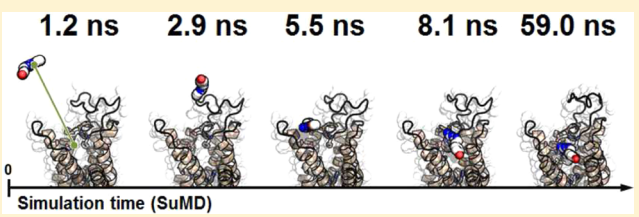
pubs.acs.org/jcim

Supervised Molecular Dynamics (SuMD) as a Helpful Tool To Depict GPCR–Ligand Recognition Pathway in a Nanosecond Time Scale

Davide Sabbadin and Stefano Moro*

Molecular Modeling Section (MMS), Dipartimento di Scienze del Farmaco, Università di Padova, via Marzolo 5, 35131 Padova, Italy

Supporting Information



ABSTRACT: Supervised MD (SuMD) is a computational method that allows the exploration of ligand–receptor recognition pathway investigations in a nanosecond (ns) time scale. It consists of the incorporation of a tabu-like supervision algorithm on the ligand–receptor approaching distance into a classic molecular dynamics (MD) simulation technique. In addition to speeding up the acquisition of the ligand–receptor trajectory, this implementation facilitates the characterization of multiple binding events (such as meta-binding, allosteric, and orthosteric sites) by taking advantage of the all-atom MD simulations accuracy of a GPCR–ligand complex embedded into explicit lipid–water environment.

G protein-coupled receptors (GPCRs) are membrane proteins that serve as crucial signal transduction machineries, linking various extracellular inputs with diverse cellular responses. Indeed, a large number of clinically relevant drugs elicit their therapeutic effect(s) through GPCRs.[1,2] During the past few years, crystallography of GPCRs has experienced an unpredictable growth, resulting in the determination of the structures of 20 distinct receptors that, including closely related subtype homology models, amounts to approximately 12% coverage of the human GPCR superfamily.[3] This high-resolution structural information is helping redefine our knowledge of how GPCRs recognize such a diverse classes of ligands and how they transmit signals across the cell membrane Moreover, they have provided an enormous opportunity for computational methodologies to make major contributions in this field. In particular, molecular dynamics simulations have become a driving factor in many areas of GPCR biophysics and molecular pharmacology, improving our understanding of ligand–receptor interaction, activation mechanisms, receptor hydration, and ligand-subtype selectivity.[4−8] Given that computers will continue to get faster and more structures will be solved, the importance of computational methods will only continue to grow, particularly as simulation research is more closely coupled to experimentation.

In fact, one of the most challenging issues for the future of drug discovery is the capability to understand the GPCR–ligand recognition pathway with the aim to facilitate the development of drug candidates with more favorable phamacodynamic profiles. Unfortunately, the recognition process between a ligand and its receptor is a very rare event to describe at the molecular level, and even with the recent GPU-based computing resources, it is necessary to carry out classical molecular dynamics (MD) experiments in a long microsecond time scale.[5,7] In order to overcome this limiting factor, we have implemented an alternative MD approach, named supervised molecular dynamics (SuMD), that enables us to follow GPCR–ligand approaching process within a time scale reduced up to 3 orders of magnitude compared to classical MD. SuMD enables the investigation of ligand–receptor binding events independently from the starting position, chemical structure of the ligand, and also from its receptor binding affinity.

SuMD is a standard MD simulation in which the ligand–receptor docking pathway is supervised by a tabu-like algorithm (Figure 1). During the production of the MD trajectory, the distance between the center of masses of the ligand atoms and the residues composing the orthosteric binding site of the

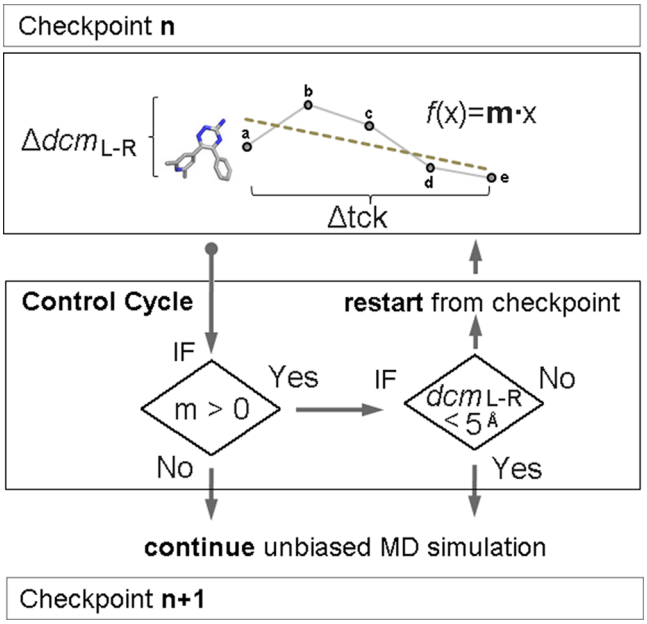


Figure 1. Scheme of the ligand–receptor distance vector (dcm_{L-R}) supervision algorithm implemented in the supervised molecular dynamics (SuMD) technique.

Published: January 23, 2014

dx.doi.org/10.1021/ci400766b | J. Chem. Inf. Model. 2014, 54, 372–376

GPCR ($dcm_{L\text{-}R}$) is monitored over a fixed time window (Δt_{ck}, e.g., 200 ps). An arbitrary number of distance points (n: a, b, c, d, e) per each checkpoint trajectory is collected in real time and a linear function $f(x) = m \times x$ is fitted on the distance points at the end of the checkpoint time. A supervision tabu-like algorithm is applied to increase the probability to produce ligand−receptor binding events without introducing bias to the MD simulation. More precisely, if $m < 0$, the ligand−receptor distance is likely to be shortened over the checkpoint time, and classic MD simulation is restarted from the last produced set of coordinates. Otherwise, the simulation is restored from the original set of coordinates and random velocities of each atom in the system, reassigned coherently to the NVT ensemble. The tabu-like supervision algorithm is perpetuated in time until the ligand−receptor distance ($dcm_{L\text{-}R}$) is less than 5 Å. To validate the methodology, we selected as a key study the human A_{2A} adenosine receptor ($hA_{2A}AR$) that has been recently crystallized with several ligands, both agonists and antagonists, characterized by different receptor binding affinities. In particular, we selected four crystal structures of the $hA_{2A}AR$ in complex with three strong binders such as 4-(2-(7-amino-2-(2-furyl)(1,2,4)triazolo(2,3-*a*)(1,3,5)triazin-5-yl-amino)ethyl)-phenol, ZM 241385 ($pK_D = 9.18 \pm 0.0$,[9] PDB code: 3EML[10]); 6-(2,6-dimethylpyridin-4-yl)-5-phenyl-1,2,4-triazin-3-amine, T4G ($pK_D = 8.9 \pm$ n.d.,[9] PDB code: 3UZA[11]); 4-(3-amino-5-phenyl-1,2,4-triazin-6-yl)-2-chlorophenol, T4E ($pK_D = 9.6 \pm$ n.d.,[9] PDB code: 3UZC[11]); and a weaker binder such as caffeine ($pK_D = 5.31 \pm 0.44$,[9] PDB code: 3RFM[9]).

In all analyzed cases, we were able to reproduce the complete binding process in a nanosecond time scale reproducing with high accuracy the crystallographic pose of each ligand. All SuMD trajectories were run in triplicate (see Supporting Information for more details). Moreover, using SuMD simulations, it is possible to easily determine and characterize all possible ligand binding sites that chronologically anticipate the orthosteric one. These sites are better known as meta-binding sites,[12] and in some cases, they may coincide with possible allosteric sites.[7] The SuMD approach has the potential to facilitate a better understanding of all GPCR−ligand recognition pathways thus increasing the potentiality of in silico screening to expedite drug development taking account of full protein flexibility, water-mediated ligand−receptor interactions, and the presence of the membrane environment as well.

ZM241385−HUMAN A_{2A} ADENOSINE RECEPTOR RECOGNITION MECHANISM

The ligand recognition pathway described by supervised molecular dynamics highlights two major interaction sites that anticipate the crystallographic binding conformation (b,c, Figure 2, panel A). In particular, extracellular loop 2 (EL2) and extracellular loop 3 (EL3) of the human A_{2A} adenosine receptor are involved in the ligand recognition process. The highlighted meta-binding sites are engaged in tuning ZM241385 orientation and conformation to appropriately reach (d, Figure 2, panel A) and fit (e, Figure 2, panel A) into the orthosteric binding site (Videos S1 and S2, Supporting Information). The antagonist, starting from a randomized set of coordinates at least 40 Å away from protein atoms (a, Figure 2, panel A) reaches the orthosteric binding site accurately reproducing the crystallographic pose in less than 60 ns.

The most energetically stable ligand−receptor complex structures (i, ii, iii, Figure 2, panel B and C) were extracted after an energy inspection of the conformational ensemble generated from the SuMD simulation after the orthosteric

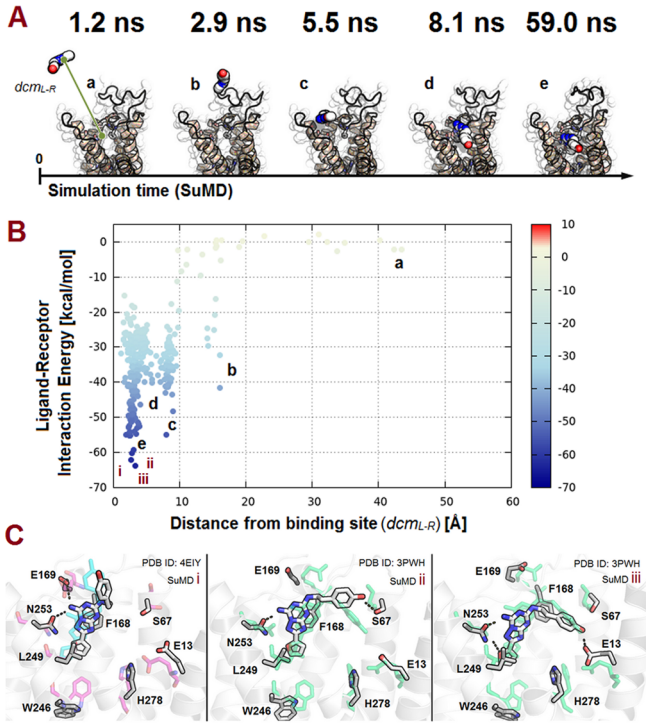


Figure 2. (A) Overview of the adenosine receptor antagonist ZM241385−human A_{2A} adenosine receptor recognition mechanism using supervised molecular dynamics (SuMD). Simulation time, when the depicted event occurs, is reported above the ligand−receptor representation. Ligand−receptor distance vector ($dcm_{L\text{-}R}$) is shown. van der Waals spheres represent ZM241385 atoms, and receptor ribbon representation is viewed from the membrane side facing transmembrane domain 6 (TM6) and transmembrane domain 7 (TM7). Hydrogen atoms are not displayed. (B) Ligand−receptor interaction energy landscape for the nonbiased ZM241385−human A_{2A} adenosine receptor recognition process. Some of the most important characterized binding sites (b, c, d) that anticipate the crystallographic information (e) are highlighted. Interaction energy values are expressed in kcal/mol. (C) Overview of the three most energetically stable binding conformation of ZM241385 inside the hA2A AR binding pocket generated from SuMD simulation (white sticks) in comparison with two representative XRAY structures, PDB ID: 3PWH (green sticks) and PDB ID: 4EIY (cyan sticks). The complexes are viewed from the membrane side facing TM6 and TM7 with the view of TM7 is partially omitted. Hydrogen atoms are not displayed, whereas hydrogen bond interactions are highlighted as yellow dashed lines.

binding site recognition and compared to the XRAY structural information available (Figure 2, panel C). Upon recognition, ZM241385 exibits low fluctuation into the binding site (rmsf of the triazolotriazine core <2 Å over 5 ns) and the phenolethyl chain, attached to the triazolotriazine ring, explore the same diverse conformational landscape anticipated by XRAY crystallography. In fact, in the latest stage of formation of the high affinity antagonist−human A_{2A} adenosine receptor, the structural information extracted from SuMD simulation is undistinguishable (r.m.s.d. below crystallographic resolution limits) from the XRAY crystallographic structure available even in the case of different receptor constructs (Figure 2, panel C).

T4G−HUMAN A_{2A} ADENOSINE RECEPTOR RECOGNITION MECHANISM

As for ZM241385, the T4G recognition pathway highlights multiple ligand−receptor binding events that anticipate the

dx.doi.org/10.1021/ci400766b | J. Chem. Inf. Model. 2014, 54, 372−376

orthosteric binding site recognition. The 1,2,4-triazine derivative can be trapped in a transient pocket, named meta-binding site 2 (Videos S3 and S4,Supporting Information), located in the second extracellular loop 2 (ECL2). After this binding event (b, Figure 3, panel A), the aromatic substituents at the 1,2,4-triazine aromatic core are directed toward the third extracellular loop (EL3) that represents the common meta-binding site (c, Figure 3, panel A) explored by ZM241385 as well. In this case the antagonist, starting from a randomized set of coordinates at least 40 Å away from protein atoms (a, Figure 3, panel A) reaches the orthosteric binding site (d, Figure 3, panel A) accurately reproducing the crystallographic pose (e, Figure 3, panel A) in less than 65 ns (see Supporting Information for more details).

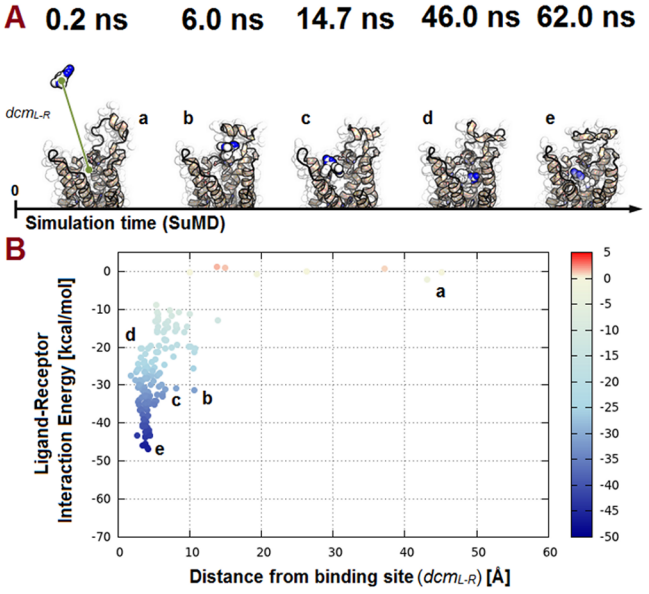


Figure 3. (A) T4G–human A_{2A} adenosine receptor recognition mechanism using supervised molecular dynamics (SuMD). Simulation time, when the depicted event occurs, is reported above the ligand–receptor representation. Ligand–receptor distance vector ($dcm_{L\text{-}R}$) is shown. van der Waals spheres represent T4G atoms, and receptor ribbon representation is viewed from the membrane side facing transmembrane domain 6 (TM6) and transmembrane domain 7 (TM7). Hydrogen atoms are not displayed. (B) Ligand–receptor interaction energy landscape for the nonbiased T4G–human A_{2A} adenosine receptor recognition process. Some of the most important characterized binding sites (b, c) that anticipate the crystallographic information (e) are highlighted. Interaction energy values are in kcal/mol.

T4E–HUMAN A_{2A} ADENOSINE RECEPTOR RECOGNITION MECHANISM

The extracellular loop 3 (EL3) of the human A_{2A} adenosine receptor plays a crucial role in the recognition of T4E. In fact, starting from a randomized set of coordinates at least 40 Å away from protein atoms (a, Figure 4, panel A), the antagonist makes contact with EL3 (b, c, Figure 4, panel A) and eventually reaches the orthosteric binding site (d, Figure 4, panel A) and makes contact that accurately reproduces the crystallographic structure (e, Figure 4, panel A) in less than 110 ns. The recognition mechanism, investigated using SuMD, is reported in Videos S5 and S6 of the Supporting Information.

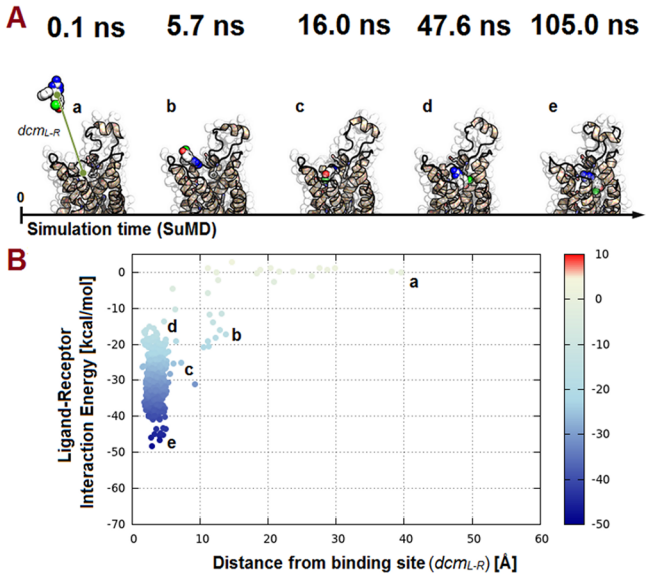


Figure 4. (A) T4E–human A_{2A} adenosine Receptor recognition mechanism using supervised molecular dynamics (SuMD). Simulation time, when the depicted event occurs, is reported above the ligand–receptor representation. Ligand–receptor distance vector ($dcm_{L\text{-}R}$) is shown. van der Waals spheres represent T4E atoms, and receptor ribbon representation is viewed from the membrane side facing transmembrane domain 6 (TM6) and transmembrane domain 7 (TM7). Hydrogen atoms are not displayed. (B) Ligand–receptor interaction energy landscape for the nonbiased T4E-human A_{2A} adenosine receptor recognition process. Some of the most important characterized binding sites (b, c) that anticipate the crystallographic information (e) are highlighted. Interaction energy values are in kcal/mol.

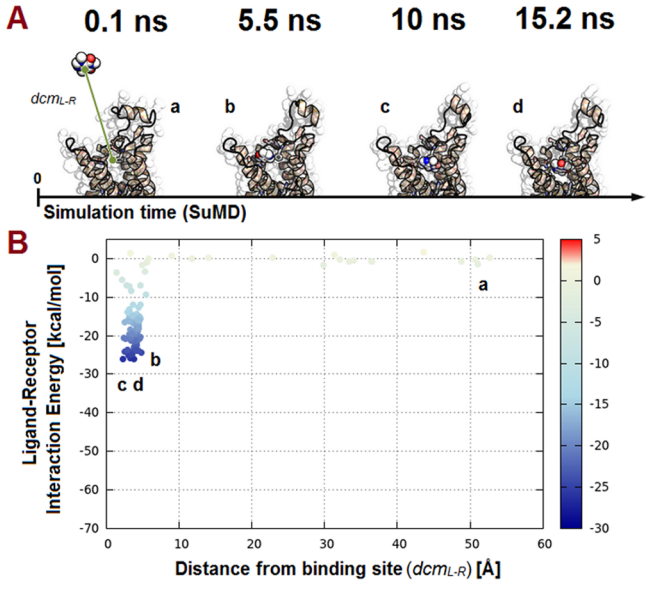


Figure 5. (A) Caffeine–human A_{2A} adenosine receptor recognition mechanism using supervised molecular dynamics (SuMD). Ligand–receptor distance vector ($dcm_{L\text{-}R}$) is shown. van der Waals spheres represent caffeine atoms, and receptor ribbon representation is viewed from the membrane side facing transmembrane domain 6 (TM6) and transmembrane domain 7 (TM7). Hydrogen atoms are not displayed. (B) Ligand–receptor interaction energy landscape for the nonbiased caffeine–human A_{2A} adenosine receptor recognition process. Some of the most important characterized binding sites (b, c) that anticipate the crystallographic information (e) are highlighted. Interaction energy values are in kcal/mol.

dx.doi.org/10.1021/ci400766b | J. Chem. Inf. Model. 2014, 54, 372–376

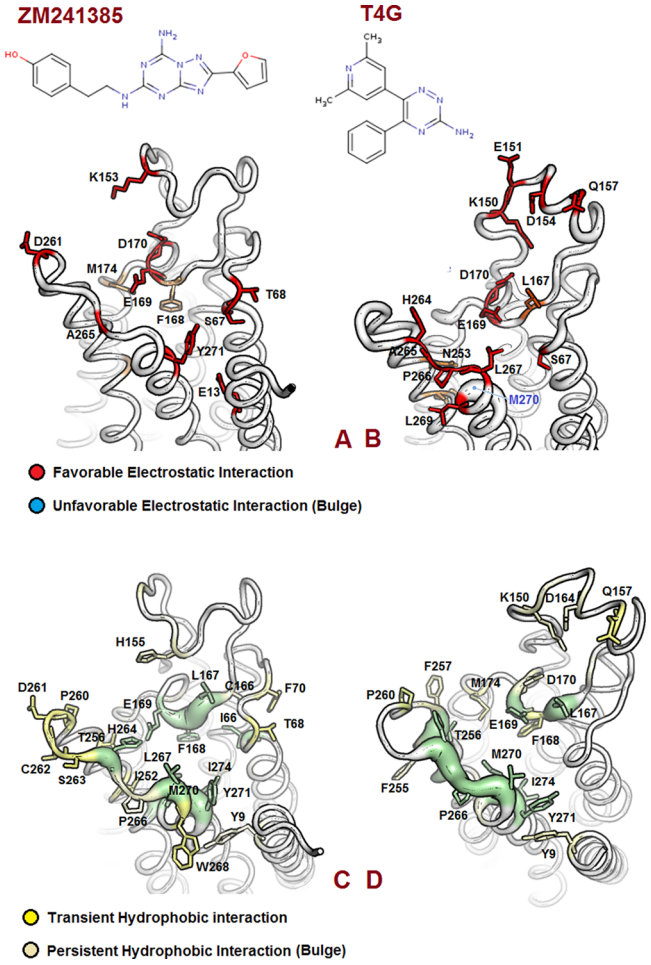


Figure 6. Electrostatic (A, B) and hydrophobic (C, D) contributions to the interaction energy of each receptor residue involved in the binding with the high affinity hA2A AR antagonists ZM241385 and T4G during the metabinding sites recognition process. Contributions to ligand binding were calculated during the first 15 ns of SuMD simulations. Ribbon representation is viewed from the extracellular side, and hydrogen atoms are not displayed.

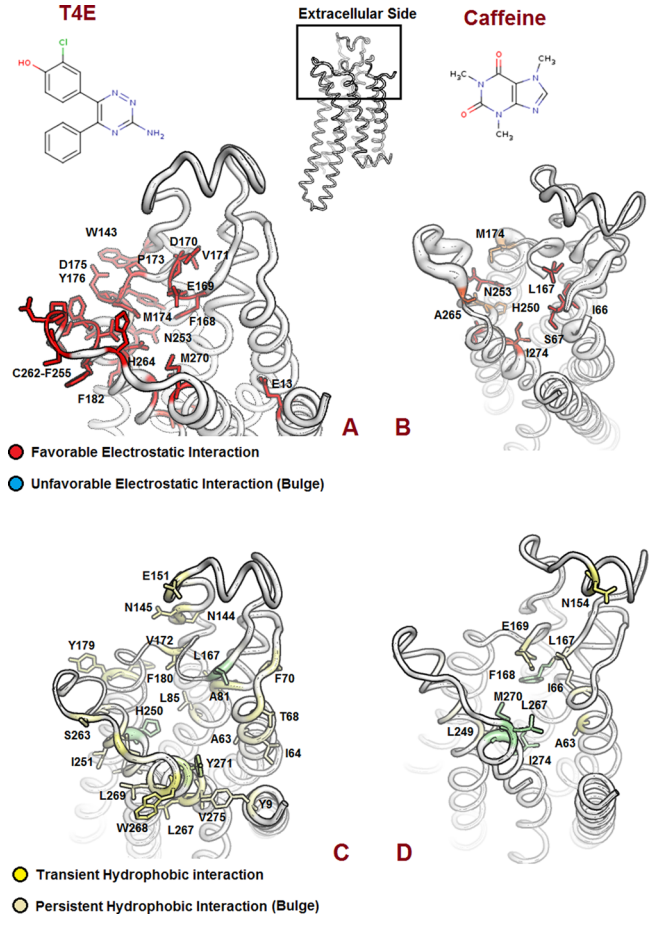


Figure 7. Electrostatic (A, B) and hydrophobic (C, D) contributions to the interaction energy of each receptor residue involved in the binding with the high affinity hA2A AR antagonist T4E and low affinity caffeine during the metabinding sites recognition process. Contributions to ligand binding were calculated during the first 15 ns of SuMD simulations (10 ns for caffeine). Ribbon representation is viewed from the extracellular side, and hydrogen atoms are not displayed.

CAFFEINE–HUMAN A_{2A} ADENOSINE RECEPTOR RECOGNITION MECHANISM

As reported for the high-affinity human A_{2A} adenosine receptor antagonists, the purine derivative caffeine recognition mechanism is mediated (b, Figure 5, panel A) by the extracellular loop 3 (EL3). Upon binding (c, d, Figure 5, panel A), the weak antagonist shows fragment-like features and lacks strong interactions with the binding site (rmsf > 4 Å). The complete binding event, described using SuMD, is reported in Videos S7 and S8 of the Supporting Information.

Moreover, supervised molecular dynamics simulations recognize the critical role of the hA2A AR extracellular loops in the ligand recognition process that have been postulated using site-directed mutagenesis in the past.[12−14] The complex evolving network of interactions has been depicted using a simplified ribbon representation of the receptor that comprises a quantitative estimate of the occurring ligand−protein mutual recognition process (Figures 6 and 7). In fact, SuMD could represent a powerful tool to assist the design a focused set of aminoacid mutation experiments in order to infer their role on the molecular recognition process. A critical analysis of the interaction maps reported in detail in Figures 6 and 7 highlights the involvement of the vast majority of the residues located in the extracellular loop 2 (EL2) and extracellular loop 3 (EL3) of hA2A AR in ligand recognition, thus confirming the crucial role of the acidic residues located in EL2[13] (E151, D170, E169).

CONCLUSIONS

In the present work, we have presented a helpful approach to simulate a receptor−ligand recognition pathway based on a GPU-driven membrane molecular dynamics simulation in an appreciable short time scale. Combining two very well-known and consolidated computational approaches such as molecular dynamics and a tabu-like algorithm to supervise the evolution of receptor−ligand trajectory, we are able to accurately completely explore the receptor−ligand event in a nanosecond time scale. This approach is also very useful to analyze both orthosteric and allosteric binding events broadening our perspectives in several scientific areas from molecular pharmacology to drug discovery. In particular supervised MD (SuMD) can be applied in a drug design campaign for lead optimization in a high-throughput level in order to design novel binders with preferable pharmacodynamic profiles. Moreover, SuMD represents a powerful tool to assist the design site-directed mutagenesis experiments in order to investigate the molecular recognition process.

dx.doi.org/10.1021/ci400766b | J. Chem. Inf. Model. 2014, 54, 372−376

■ ASSOCIATED CONTENT

Supporting Information

Complete experimental section and additional result discussion for ZM241385, T4E, T4G, and caffeine−human A_{2A} adenosine receptor recognition mechanism. This work was supported by the University of Padova, Italy, and the Italian Ministry for University and Research (MIUR), Rome, Italy. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: +39 049 8275366. Tel: +39 049 8275704. E-mail: stefano.moro@unipd.it.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

ARs: adenosine receptors; EL2: second extracellular loop; EL3: third extracellular loop; GPCRs: G protein-coupled receptors; GPU: graphics processing unit; MD: molecular dynamics; SuMD: supervised molecular dynamics; n.d: not determined; POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; T4E: 4-(3-amino-5-phenyl-1,2,4-triazin-6-yl)-2-chlorophenol; T4G: 6-(2,6-dimethylpyridin-4-yl)-5-phenyl-1,2,4-triazin-3-amine; TM: transmembrane; ZM 241385: 4-(2-(7-amino-2-(2-furyl)(1,2,4)triazolo(2,3-a)(1,3,5)triazin-5-yl-amino)ethyl)-phenol

■ ACKNOWLEDGMENTS

S.M. participates in the European COST Action CM1207 (GLISTEN).

■ REFERENCES

(1) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727−730.

(2) Drews, J. Drug discovery: A historical perspective. *Science* **2000**, *287*, 1960−1964.

(3) Jacobson, K. A.; Costanzi, S. New insights for drug design from the X-ray crystallographic structures of G-protein-coupled receptors. *Mol. Pharmacol.* **2012**, *82*, 361−371.

(4) Dror, R. O.; Jensen, M. Ø.; Borhani, D. W.; Shaw, D. E. Exploring atomic resolution physiology on a femtosecond to millisecond timescale using molecular dynamics simulations. *J. Gen. Physiol.* **2010**, *135*, 555−562.

(5) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13118−13123.

(6) Buch, I.; Giorgino, T.; De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10184−10189.

(7) Dror, R. O.; Green, H. F.; Valant, C.; Borhani, D. W.; Valcourt, J. R.; Pan, A. C.; Arlow, D. H.; Canals, M.; Lane, J. R.; Rahmani, R.; Baell, J. B.; Sexton, P. M.; Christopoulos, A.; Shaw, D. E. Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* **2013**, *503*, 295−299.

(8) Selvam, B.; Wereszczynski, J.; Tikhonova, I. G. Comparison of dynamics of extracellular accesses to the β1 and β2 adrenoceptors binding sites uncovers the potential of kinetic basis of antagonist selectivity. *Chem. Biol. Drug Des.* **2012**, *80*, 215−226.

(9) Doré, A. S.; Robertson, N.; Errey, J. C.; Ng, I.; Hollenstein, K.; Tehan, B.; Hurrell, E.; Bennett, K.; Congreve, M.; Magnani, F.; et al. Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure* **2011**, *19*, 1283−1293.

(10) Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211−1217.

(11) Congreve, M.; Andrews, S. P.; Doré, A. S.; Hollenstein, K.; Hurrell, E.; Langmead, C. J.; Mason, J. S.; Ng, I. W.; Tehan, B.; Zhukov, A.; et al. Discovery of 1,2,4-triazine derivatives as adenosine A(2A) antagonists using structure based drug design. *J. Med. Chem.* **2012**, *55*, 1898−1903.

(12) Moro, S.; Hoffmann, C.; Jacobson, K. A. Role of the extracellular loops of G protein-coupled receptors in ligand recognition: A molecular modeling study of the human P2Y1 receptor. *Biochemistry* **1999**, *38*, 3498−3507.

(13) Kim, J.; Jiang, Q.; Glashofer, M.; Yehle, S.; Wess, J.; Jacobson, K. A. Glutamate residues in the second extracellular loop of the human A2A adenosine receptor are required for ligand recognition. *Mol. Pharmacol.* **1996**, *49*, 683−691.

(14) Kim, J.; Wess, J.; van Rhee, A. M.; Schöneberg, T.; Jacobson, K. A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human A2a adenosine receptor. *J. Biol. Chem.* **1995**, *270*, 13987−13997.