Molecular Dynamic Simulation of mGluR5 Amino Terminal Domain: Essential Dynamics Analysis Captures the Agonist or Antagonist Behaviour of Ligands

Alessandro Casoni, Francesca Clerici and Alessandro Contini*

Dipartimento di Scienze Farmaceutiche - sezione di Chimica Generale e Organica “Alessandro Marchesini”, Università degli Studi di Milano, via Venezian, 21 20133 Milano, Italy. Fax: +39 02 50314476; Tel: +39 02 50314480; E-mail: alessandro.contini@unimi.it

Abstract

We describe the application of molecular dynamics followed by principal component analysis to study the inter-domain movements of the ligand binding domain (LBD) of mGluR5 in response to the binding of selected agonists or antagonists. Our results suggest that the method is an attractive alternative to current approaches to predict the agonist-induced or antagonist-blocked LBD responses. The ratio between the eigenvalues of the first and second eigenvectors (R_{1,2}) is also proposed as a numerical descriptor for discriminating the ligand behavior as a mGluR5 agonist or antagonist.

Keywords

Molecular dynamic simulations; glutamate metabotropic receptor; principal component analysis; essential dynamics

1. Introduction

Eight members of the G-protein-coupled metabotropic glutamate receptor (mGluR) family have been identified (mGluR1-8) and divided into three groups on the basis of sequence homology, second messenger coupling and pharmacology.[1, 2] mGluR family members all possess a large bi-lobed
extracellular N-terminal domain (NTD) which has been demonstrated to contain the glutamate binding site by both site-directed mutagenesis[3] and X-ray crystallography.[4] The N-terminal domain is linked via an extracellular cysteine-rich region to a typical G protein-coupled receptor (GPCR) trans membrane heptahelical domain which mediates G-protein activation. The Ligand Binding Domain (LBD) of mGluRs shows a significant homology with the Leucine/Isoleucine/Valine Binding Protein (LIVBP) member of Periplasmic Binding Protein family,[5] and, on the basis of this homology, a mechanism, known as Venus’ flytrap model, has been proposed for the activation of LBD of mGluRs.[6] Once the agonist has bound the receptor, the two amino-terminal domain (ATD) lobes (LB1 and LB2) collapse, thus trapping the ligand into the cleft. When occupied by a competitive antagonist, the bi-lobed extracellular domain exists in an open-inactive conformation (Figure 1).

**Figure 1.** Representation of the open-close equilibrium of mGluRs NTD. Agonists stabilize the closed-active conformation, while antagonists stabilize the open-inactive conformation.

mGluRs crystallized so far are indeed constituted of two promoters, either in a closed or in an open conformation.[4, 7, 8] Both open and closed conformations have been observed when L-glutamate was bound, whereas only open forms were detected with an antagonist bound. Thus, the crystallized complexes clearly confirmed the hypothesis that the LBD functioning is mediated through a Venus’
flytrap mechanism operated by the equilibrium between an open and a closed conformation.
Furthermore, the crystallization of agonist- and antagonist-bound LBDs supported the notion that agonists act by stabilizing the closed-active conformation while antagonists favor the open-inactive form.[4]
The Group I mGluR pharmacology appears to offer a considerable therapeutic potential.[9] Particularly mGluR5 antagonists exhibit anxiolytic activity in a broad range of preclinical models at doses which do not affect spontaneous locomotor activity.[10, 11] mGlu5 receptors have also been implicated in the development of treatments for the dependence to psychostimulants and other drugs of abuse.[12]
Although the benefits of available X-ray structures for the LBD complexed with ligands are evident in terms of novel drug design opportunities, the crystal structures only represent frozen snapshots of a dynamic and complex process. Molecular dynamics (MDs) are an important tool to study the functional dynamics of proteins and macromolecular complexes.[13-16] Some MD studies have already been carried out on the Venus’ flytrap architecture,[17-19] but a classical MD simulation may underestimate the entire displacement for proteins that require a relaxation time of the order of milliseconds or longer. Indeed, one of the major limitations of MDs is the length of achievable simulation times, typically of the order of tens to hundreds nanoseconds.[4] These times are much shorter than the time scale of many important biological processes, that take place on the millisecond scale.[20] The question is “how long should be a MD run to achieve meaningful results?” or “how long is long enough?” The answer should be “longer than the relaxation time of the property under investigation”, but several attempts have been made to extract essential functional features from short trajectories, with the aim to describe such features as trends.[21-26] However, the large number of variables typical of MDs of complex systems can induce a maddening noise which prevents a correct interpretation of results. For this reason, a method to exalt feature/noise ratio should be adopted to improve the analysis. One widely used
statistical approach to achieve a dimensionality reduction is the Essential Dynamics (ED) procedure.[23] ED is a very powerful analysis technique which exploits a Principal Component Analysis (PCA)[24] to identify the nature and relative importance of the essential motions of a macromolecule from a MD sampling. The result of an ED analysis is a set of eigenvectors, which describe the nature of motions in the Cartesian space, and eigenvalues, which represent the amount of variance described by each eigenvector. The ED method is able to separate the few large, mainly anharmonic motions (the essential subspace) in a MD trajectory from the small, Gaussian fluctuations (the near-constraints subspace).[27] The motions in the essential space are linked to the biological function of the protein. In the framework of a project aiming to develop novel mGluR5 ligands, we decided to apply the ED method to study the LBD motions in presence of either agonists or antagonists.

2. Computational methods


Homology models of the open-form of mGluR5, were built using Modeller 9v8 [28] and the best model selected on the basis of DOPE score.[29] To select the most appropriate templates for our query sequence (UniProtKB P41594) over the nine similar chains of mGluR1 in the PDB database,[4] we assessed the structural and sequence similarity between the possible templates by performing an iterative least-squares superposition of the nine 3D structures, using the multiple sequence alignment as its starting point (Figure S1, Supporting Information). Next we compared the structures according to the constructed alignment. From the obtained pairwise sequence distance matrix (Table S1, Supporting Information) it can be observed that the B chain of structures 1EWK, 1EWV and 1EWT are the most suitable templates.[4] We initially considered a single template strategy by only considering 1EWKB due to its higher resolution, but the plotted DOPE score profiles (Figure S2, supporting Information) of
the best model (DOPE score = −53528.8) with respect to the template show regions of relatively high energy for the long active site loop between residues 94 and 118. This problem was partially overcame by using a multiple template strategy considering all the three above mentioned structures, as evidenced by the better DOPE score obtained for the new model (−55193.0) and by the plot reported in Figure S2, Supporting Information. A final loop-optimization was then attempted by building 50 different and independently optimized loop conformations for the 94-118 loop region, leading to the final model showing a DOPE score of −55510.7 (see Figure S3, Supporting Information). Finally the structure was minimized using a steepest descent/conjugate gradient method in vacuo using GROMACS software and the GROMOS96 forcefield.[30, 31] The agonists (L-glutamate, L-quisqualate,[32] and 2-chloro-5-hydroxyphenylglycine (CHPG)[33]) and antagonists (S-4-carboxyphenylglicine (S4-CPG),[34] LY344545,[35] and α-methyl-4-carboxyphenylglycine (MCPG)[36]) studied in this work (Figure 2) were docked into the first lobe (LB1, the binding region) of mGluR5 model with the Autodock v3.0.5 package (see Supporting Information for further details)[37]

Figure 2. Evaluated ligands of mGluR5. Agonists: (1) L-Glutamic acid, (2) L-Quisqualic acid, (3) CHPG; antagonists: (4) S4-CPG, (5) LY344545, (6) MCPG.
2.3. MD Simulations

The PRODRG server was used to derived specific force field parameters for the herein investigated ligands.[38] For all ligands, carboxylic and amino functional groups were considered as conjugated bases and acids, respectively. The MD calculations were made by using the GROMACS v4.5 software package with the GROMOS96 53a6 force field.[31, 39] In each simulation, the initial structure was solvated by a periodic water box using the SPC water model[40] and ions were added to neutralize the total charge of the systems. A non-bond pair list cutoff of 10 Å was used and the pair list was updated every 10 time steps. The long-range electrostatic interactions were treated with the Particle-Mesh Ewald method and a time step of 2 fs was used for numerical integration. Then, energy minimizations were carried out with the Conjugate Gradient algorithm, followed by 200 ps of equilibration (100 ps with NVT and 100 ps with NPT ensemble) at 310 K with position restraints on protein and ligands (fc=1000) and by 100 ps of NPT equilibration at 310 K with position restraints on protein backbone. Production runs were then performed in NPT ensemble at 310 K for 20 ns. Further computational details may be found in the Supporting Information. Only the Cα atoms (482) were included in the ED analyses, as it has been shown that this simplification better detects the large-scale concerted motions in proteins.[23, 41]

3. Result and discussion

It is known that group I mGluR ligands bind to the LB1 lobe and both agonists and antagonists share similar binding modes.[4, 7, 42, 43] However, while antagonists stabilize the open-inactive conformation, glutamate and other agonists stabilize the closed-active conformation. For this reason,
additional H-bonds can be observed in the closed-active conformation between glutamate and LB2 residues Asp305 and Arg310, this latter mediated by a water bridge.\[42\] Since our aim was to study, through MD simulations, the receptor motion following ligand binding, all of our simulation originated from the receptor in the open conformation, where the different ligand were docked. The reliability of our starting geometries was then assessed by comparing the binding conformations obtained by docking with those experimentally observed for MCPG \[7\] and LY341495 \[43\] antagonists, crystalized in the open-inactive conformation of mGluR1 receptor (having the 85\% of sequence identity with mGluR5 and a fully conserved binding site), and with that of glutamate crystallized in the closed-active conformation of mGluR5.\[42\]

Concerning agonists, all our starting complex showed H-bonds with key amino acids reported by literature, namely Tyr 64, Ser 152, Ser 173 and Thr 175.\[4, 7, 42\] In particular, as shown in Figure 3, glutamate 1 interacts by its amino acidic carboxy group with the NH of both Ser 152 and Thr 175 (donor-acceptor distances = 2.95 and 3.19 Å, respectively), while the amino group interacts with the Ser 173 carbonyl (donor-acceptor distance = 3.03 Å); interactions can also be observed between the glutamate side chain carboxy group and both the Tyr 64 OH and Arg 68 guanidine groups (donor-acceptor distances = 3.46 and 2.33 Å, respectively).
Figure 3. Bounded complexes used as starting geometries for MD. Carbons of compounds 1-6 are colored in grey, green, orange, ochre, purple and cyan, respectively.

Similar interactions were found for quisqualate 2, where its carboxy group makes an H-bond with the Ser 152 NH (donor-acceptor distance = 3.07 Å) while the amino group interacts with the Ser 173 carbonyl (donor-acceptor distance = 3.02 Å) and the Thr 175 OH group (donor-acceptor distance = 3.26 Å). Also in this case, a weaker interactions was observed between the Tyr 64 OH group and the 3-carbonyl of the oxadiazolidinyl moiety of 2 (donor-acceptor distance = 3.75 Å).

Finally, CHPG 3 also showed H-bonds between the carboxy group and the NH of both Ser 152 and Thr 175 (donor-acceptor distances = 2.99 and 3.20 Å, respectively) and between the amino group and the OH of Thr 175 (donor-acceptor distances = 2.68 Å); a weak interaction was found between the 5-OH group of 3 and the Tyr 64 OH (donor-acceptor distance = 3.65 Å), while no interactions were observed with Ser 173 (a distance of 4.05 Å was measured between the CHPG nitrogen and the Ser 173 carbonyl oxygen).

Concerning antagonists, both the structurally related compounds S4-CPG 4 and MCPG 6 show a binding mode comparable to the one observed in the crystal structure of 6 complexed with the mGluR1 receptor.[7] In particular, H-bonds were observed between the aminoacidic carboxy group of 4 and 6 with the Ser 152 NH (donor-acceptor distances = 3.03 and 3.37 Å, respectively), with the former compound making an additional interaction with the Thr 154 NH (donor-acceptor distance = 3.21 Å). The amino group of both compounds makes an H-bond with the Thr 154 OH (donor-acceptor distances = 2.84 and 2.83 Å, respectively), while for both 4 and 6 the aromatic carboxy group interacts with the Tyr 43 OH (donor-acceptor distances = 2.81 and 3.19 Å, respectively).
In the framework of a project aiming to develop novel mGluR5 ligands, we decided to study the LBD motions in presence of agonists and antagonists through MD simulations, followed by ED analysis.

Figure 4. Snapshot of the defined feature: $d_1$ distance between Ca of Tyr 64 (LB1) and Arg 310 (LB2).

The distance between the Ca of Tyr 64 and Arg 310 (hereafter referred as $d_1$), positioned on the surface of the first and the second lobe, respectively (Figure 4), was defined as the geometrical feature followed to obtain information on the direction and the magnitude of the inter-domain movement. This distance has been chosen in analogy to structural analyses conducted on the mGluR1, where the two residues correspond to Tyr 74 and Arg 323, respectively.[4] Indeed, those residue belongs to different lobes and are both involved in glutamate recognition in the activated form,[4] while only Tyr 74 is involved in the recognition of antagonists blocking the receptor in its open-inactive form.[43] For these reasons, the distance between such residues provide a clean indication of the LB1 and LB2 collapsing during the mGluR activation process.

Trajectory analyses show that the RMSD (calculated using the first frame as a reference) after about 10 ns reaches a plateau which is generally higher for agonists respect to antagonists (see Figure 5, top, and Figure S4, Supporting Information), meaning that, for the former set of compounds, the equilibrium geometry is somehow farther from the starting open-inactive conformation.
Figure 5. Top: time-evolution of RMSD (top) for mGluR5 complexed with 1 (black) or 5 (red); bottom: time-evolution of distance $d_1$ computed for agonist 1 (black) and antagonist 5 (red).
This different behavior is not surprising, considering how agonists and antagonists exert their action: the former lead to receptor activation, thus driving the switching between the open-inactive and the closed-active receptor conformation, while the latter tend to stabilize the open-inactive conformation.

The time-evolution of $d1$ confirm this observation, being fairly constant for antagonists (although some fluctuations are observed due to the oscillating inter-lobe motion), while sharply decreasing, especially in the first half of the simulation trajectory, for agonists (see Figure 5, bottom, and Figure S5, Supporting Information). Taken together, the time-evolution of both RMSD and $d1$ suggests that the switching between open-inactive to closed-active conformations induced by agonists occurs, in the adopted simulation conditions, in the first 10 ns of MD trajectory.

A further indication of the ligand character (agonist or antagonist) can be obtained through an H-bond analysis of the MD trajectory. Figure 7 shows the time evolution of the number of H-bonds that are formed between the two lobes, LB1 and LB2 for mGluR5 complexes with agonist 1 and antagonist 5 over 20 ns of MD simulation.

The comparison of the two data-set showed that, soon after the beginning of the production run, agonist 1 started to induce the collapse of the two lobes and, consequently, the formation of new inter-lobe H-bonds, while antagonist 5 stabilized the open-inactive conformation, thus limiting the number of H-bonds that can be formed between the two lobes.
Figure 6. Time-evolution of H-bonds between lobes LB1 and LB2 (LB1 = Ser 22 – Val 191 and Asp 333 – Gly 460; LB2 = Ala 196 – Leu 329 and Tyr 466 – Lys 503) for mGluR5 complexes with agonist 1 (Glu) and antagonist 5 (LY344545).

Similar conclusions were also be drawn by comparing the time-evolution of inter-lobe H-bonds obtained for agonist 3 with those obtained for antagonists 4-6 (see Figure S6, Supporting Information), but less clean results were obtained from the H-bond analysis of agonist 2 complex, thus preventing a reliable discrimination of the ligand character. We thus decided to apply ED analyses in order to obtain a better and cleaner description of the inter-lobe motion during the activation process.

The output trajectories were thus processed by PCA where the protein Cα were considered for the construction of the covariance matrix. In all cases, the first 10 eigenvectors, over 1446, represented at least the 70% of all movements (see Figure S7, Supporting Information).

Particularly interesting were the projections of trajectories on the first and second eigenvectors, which made possible a visual representation of motions represented by each eigenvector (see video_c1_ev1
and video_c1_ev2, Electronic Supporting Information, to visualize agonist 1 trajectory projections on eigenvectors 1 and 2; see video_c4_ev1 and video_c4_ev2, Electronic Supporting Information, for projections on eigenvectors 1 and 2 of antagonist 4). By making a parallelism with IR vibrational modes, eigenvectors 1 and 2 represent the LB1 and LB2 scissoring and twisting, respectively (Figure 7). The trajectory projection over eigenvector 1, which has the largest eigenvalue and describes the collapse of LB1 and LB2, provides a clear indication of the ATD motion upon receptor activation due to the binding of agonists.

Figure 7. Schematic representation of movements represented by the first and second eigenvectors.

To better evidence the differences between agonists and antagonists, principal components where analyzed on the first 10 ns, the time lag where the agonists-induced switching between open-inactive and closed-active conformation was predicted to occur, as well as on the full trajectory, for comparison. Indeed, as represented in Figures 8 for selected agonist 1 and antagonist 5 (see Figure S7, Supporting Information, for compounds 1-6) only a minimal variation is observed for the latter upon eigenvalues computed on the full trajectory or on the first half, while a dramatic change is observed for agonist 1, where the first eigenvalue raises from 86.4 to 109.9 amu nm$^2$ and the second eigenvalue lowers from 44.5 to 28.1 amu nm$^2$. 

13
Figure 8. Comparison of the eigenvalues of the covariance matrix obtained for the agonist 1 (left) and antagonist 5 (right) by analyzing the whole 20 ns trajectory (red, dashed) or the first 10 ns (black, plain).

This is not surprising, as MD simulations were conducted starting from the open-inactive conformation; for this reason, the binding with an agonist provokes the lobes collapse and eigenvector 1 decidedly prevails, until an equilibrium geometry (the closed-active conformation) is reached. As soon as this equilibrium is obtained, the motion represented by eigenvector 1 becomes less relevant, while other motions, such as the lobes twisting described by eigenvector 2, relatively raise in importance. Conversely, antagonists stabilize the starting open-inactive conformation and, for this reason, the amount of motion represented by eigenvectors 1 and 2 is relatively constant in time, as no large conformational changes are expected to occur.
Those findings suggested the possibility to derive from ED analyses a simple numerical descriptor able to discriminate different compounds, in terms of agonist or antagonist behavior. In fact, thinking upon the dynamic activation process of the ATD domain, two major assumptions can be made:

a) The activation process presents a prevalent movement, expressed by the collapse of the two lobes.

b) When this approach disappears, other movements become prevailing.

Upon the above assumptions, we identified the ratio between the first and the second eigenvalues (hereafter referred as R$_{1,2}$) as a potential numerical descriptor of mGluRs ligand behavior. Figure 9 shows the R$_{1,2}$ values obtained from the ED analyses of compounds 1-6, performed on the first 10 ns of MD trajectory.

![Bar chart showing R$_{1,2}$ values for different compounds](image)

**Figure 9.** Ratios between first and second eigenvector (R$_{1,2}$) obtained from ED analyses performed on the first 10 ns or on the whole 20 ns production run.
In these conditions, the difference between agonists and antagonists is rather evident with the former being characterized by $R_{1,2}$ between 3.1 (obtained for compound 3) and 4.8 (obtained for compound 2), with the natural mGluR agonist glutamate showing an $R_{1,2} = 3.9$. Conversely, antagonists all showed $R_{1,2}$ values below 2.5, ranging from 1.8 obtained for compound 6 and 2.4 obtained for 4. Interestingly, the highest $R_{1,2}$ was obtained for compound 2, where standard MD trajectory analyses leaded to an uncertain discrimination of the ligand character.

4. Conclusion

In conclusion we have simulated the activation process of ATD of mGluR5 and we have demonstrated how the ED statistical approach can be helpful for solving the time-scale problem in the activation of “fly-trap” model receptors. MD simulations and the subsequent ED analyses of the ensembles revealed that the large-scale concerted motions are dominated by ATD rearrangement.

Particularly this analysis is able to discriminate the binding mode of agonists, able to induce the switching between the open-inactive and open-active receptor conformation, and antagonists, which conversely prevent the lobes collapse. Moreover, an easily obtainable dynamic descriptor of the ligand character, herein defined as $R_{1,2}$, has been proposed. Indeed, we are confident that a reliable assessment of the ligand character might be achieved by complementing traditional analyses of other MD simulations, such as the time-evolution of RMSD, inter-lobe H-bonds and distances, with the $R_{1,2}$ value obtained from ED analyses, a dynamical descriptor which might also be helpful to derive new original QSAR parameters.

Acknowledgement
We would like to thank CT Laboratories (Sanremo – IM – Italy) for financial support. We acknowledge the CINECA “progetto di migrazione” for the availability of high performance computing resources and support.

References


