Journal of Molecular Graphics and Modelling 52 (2014) 1–10

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Constrained geometric simulation of the nicotinic acetylcholine receptor

William J. Belfield[a], Daniel J. Cole[a,b], Ian L. Martin[c], Mike C. Payne[a], P.-L. Chau[d,*]

[a] Theory of Condensed Matter Group, Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, United Kingdom
[b] Department of Chemistry, Yale University, New Haven, CT 06520-8107, United States
[c] Department of Pharmacology, University of Alberta, Edmonton, Alberta, Canada
[d] Bioinformatique Structurale, CNRS UMR 3528, Institut Pasteur, 75724 Paris, France

ARTICLE INFO

Article history:
Accepted 2 May 2014
Available online 6 June 2014

Keywords:
Nicotinic acetylcholine receptor
FRODA
FIRST
Membrane protein
Cys-loop receptors
Ligand-gated ion channels

ABSTRACT

Constrained geometric simulations have been performed for the recently published closed-channel state of the nicotinic acetylcholine receptor. These simulations support the theory that correlated motion in the flexible β-sheet structure of the extracellular domain helps to communicate a "conformational wave", spreading from the acetylcholine binding pocket. Furthermore, we have identified key residues that act at the interface between subunits and between domains that could potentially facilitate rapid communication between the binding site and the transmembrane gate.

1. Introduction

The nicotinic acetylcholine receptor (nAChR) is the archetypal member of the cys-loop family of ligand gated ion channels. The structure of the receptor, from *Torpedo marmorata*, has been determined by cryoelectron microscopy to a resolution of 4 Å (PDB structure code: 2BG9). It is composed of five subunits, two α-subunits together with single copies of β, γ and δ, arranged pseudosymmetrically around the integral ion channel, which spans the cell membrane and is selective for cations [1]. Recent studies, subsequent to time limited exposure of this protein to the natural agonist, have furnished images of the open channel structure [2], providing a structural image at both the ground state and the activated state of this important oligomeric protein. The physiological relevance of this receptor has led to detailed study of the biophysical consequences of its activation [3]. Of particular value here is rate-equilibration free energy relationship (REFER) analysis that has supplied a time-sensitive trace exploring the movement of individual structural blocks within the protein during the activation process [4]: thus providing an opportunity to explore the detail of the structural change in this protein that occurs between cryoelectron microscopic images of the ground and activated states.

Despite these experimental studies, little is known at an atomistic level concerning the gating mechanism, namely the movements that occur following ACh binding that result in a widening of the transmembrane pore. In this respect, molecular dynamics (MD) is a valuable computational tool for investigating interactions and correlations in a structure at the level of individual atoms and residues. So far, the only MD simulation performed on the 2BG9 coordinates is that of Liu et al. [5]. In this work, the authors show that inward movement of the C-loop, at the ACh-binding pocket, causes rotational movement of the transmembrane helices around the pore axis. In addition, Gao et al. [6] performed simulations on the acetylcholine-binding protein (AChBP), Law et al. [7] used the AChBP crystal coordinates and the nAChR cryo-electron microscopy coordinates as templates to build a combined model of the $(\alpha 7)_5$-subtype of the nAChR, and Cheng et al. [8] created a different model of the $(\alpha 7)_5$-subtype of the nAChR, while subsequent work used principal component analysis to identify the movements responsible for gating [9]. Despite providing valuable information, all this work was carried out before the open-channel structure of the nAChR had been determined by experiment and so does not benefit from the additional information this structure contains.

Recently, Unwin and Fujiyoshi [2] proposed a gating mechanism for moving from the closed-channel form to the open-channel form based on their experimental structures. They suggest that binding of ACh causes an orthogonal displacement of β-sheets of the α_γ-subunit. This structural shift then induces a displacement of the

* Corresponding author. Tel.: +33 1 45688546; fax: +33 1 45688719.
E-mail address: pc104@pasteur.fr (P.-L. Chau).

http://dx.doi.org/10.1016/j.jmgm.2014.05.001
1093-3263/

β-subunit away from the pore axis, which, in turn, destabilises the pentameric arrangement of helices making up the central hydrophobic gate. This destabilisation permits the M2 helices of both the α_γ- and δ-subunits to adopt straighter conformations, thus widening the pore radius. In their model, the opening of the channel is asymmetric, starting with the α_γ-subunit, moving out to the β-subunit and then the other three subunits.

In order to investigate the gating mechanism of the nAChR in more atomistic detail, one could employ MD simulations coupled with artificial forces to bias the trajectory [10,11] between the open and closed states. However, these methods are costly in computer time and require the application of arbitrary biasing forces which make it difficult to interpret the results quantitatively. In this work, we investigate the use of constrained geometric simulations, as implemented in the FRODA software [12], followed by cross-correlation analysis, principal component analysis and site-directed mutagenesis, to study the motion of the nAChR. In particular, we investigate how networks of constraints, identified by the FIRST software, lead to correlated motions in the protein that are consistent with, and hence support, the gating mechanism proposed by Unwin and Fujiyoshi [2].

2. Methods

FRODA is a protocol for performing geometrically constrained simulations of proteins. It is a computationally inexpensive method for sampling protein conformational space. Rather than the Newtonian dynamical approach of MD simulations, it employs a Lagrangian constraints-based approach, outputting an ensemble of structures that obey pre-defined constraints on bond lengths and angles, hydrogen bonds and hydrophobic interactions. Introducing such constraints allows efficient sampling of a relevant subspace of the total conformational search space, at the cost of being limited to a fixed constraints topology and generating an athermal ensemble. Explicit electrostatic interactions are not included in the computational model, and the effect of the solvent is included only indirectly by identifying and constraining hydrophobic interactions in the protein. FRODA has been widely used, for example, in studying the impact of intracellular flexibility upon the conductance of the 5-HT_3 ion channel [13], the effects of correlated motion on the transport of electronic excitations through photosynthetic bacteria complexes [14], protein–protein docking involving multiple conformational changes [15], elucidating cisplatin cross-linking in calmodulin [16], and monitoring flexibility of myosin during the ATPase cycle [17]. In comparative tests, fluctuations in an ensemble of FRODA-generated HIV-1 transactivation responsive region RNA structures agreed well with MD and NMR fluctuations [18] and FRODA outperformed MD in sampling transient pockets at protein–protein interfaces [19], indicating that dynamics of the native state emerge naturally from a simple network of native contacts.

FRODA simulations were performed with the FIRST/FRODA version 6.2 software downloaded from http://flexweb.asu.edu/. Full details of the method may be found elsewhere [12,20]. FRODA takes, as an input, a decomposition of the protein into rigid and flexible regions provided by the FIRST software.

The starting structures for the simulations were those of Unwin and Fujiyoshi [2], in the open- and closed-channel states (4AQ9 and 4AQ5, respectively). The structures were protonated using the AMBER11 package [21]. The default protonation state of titratable residues at pH 7 was used, and the resulting structures were checked by eye. The membrane was not added to the computational model, though, as we shall show later, the transmembrane domain comprises a series of large rigid clusters that are expected to be unaffected by this approximation. Missing residues in the β8-β9 loops of subunits β, δ and γ were not included, because they are expected to be flexible, and are distant from our main sites of study (the α_γ subunit and its interface with β). Suitable structures for input to FIRST were generated *via* 250 steps of steepest descent minimisation followed by 250 steps of conjugate gradient minimisation, both in the gas phase using the AMBER11 package [21]. To generate rigid cluster decompositions, the hydrogen bond energy cut off (E_{cut}) was set to −4.2 kJ/mol. The principal component subspaces spanned by FRODA simulations have been shown to be very robust with respect to the chosen value of E_{cut} [22].

For FRODA dynamics, a step length of 0.1 Å [22] was employed. Sixteen simulations, each sampling 275,000 configurations, were run starting from the open-channel structure, and sixteen were run starting from the closed-channel structure. For each state, the sixteen simulations started from the same structure, but used a different random seed for the atomic moves, thus ensuring independent trajectories. Our results for the open-channel structure are qualitatively similar to those of the closed-channel structure, and so, for simplicity, we focus exclusively on the closed-channel form in what follows. A total of 17,600 configurations were stored for analysis. The stereochemical quality of the snapshots was ascertained via PROCHECK [23] – for all snapshots, the maximum number of bad contacts was 23 and fewer than 2.5% of the amino acids were in disallowed regions of the Ramachandran plot. The total computational time for all 16 runs was 1650 h on a single Intel Sandy Bridge core, which represents a substantial saving over typical MD simulations.

Root-mean-square fluctuations (RMSF) and cross-correlation analysis was performed on the C_α atoms of the stored configurations using the PTRAJ module of AMBER11 [21]. The cross-correlation coefficient C_{ij} between atoms i and j is defined as follows:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\sqrt{\langle \Delta r_i \cdot \Delta r_i \rangle \langle \Delta r_j \cdot \Delta r_j \rangle}} \quad (1)$$

where Δr_i is the displacement vector for atom i and angular brackets denote an average over the stored configurations. The resulting coefficient is a measure of the correlated nature of the motion of pairs of residues and ranges from −1 (anti-correlated) to +1 (correlated). This simple measure allows identification of long-range correlations that are not intuitive from the structure alone, and will be used in this study as a starting point for the identification of rigid clusters that may aid communication between the ACh binding site and the pore-lining helices. We note that recent improvements to the measurement of correlation coefficients, based on the calculation of mutual information, allow the identification of non-linear and non-colinear correlations that may be missed by Eq. (1) [24]. However, Eq. (1) has been shown to be accurate for the strongest correlations, and indeed we have confirmed that the pattern of correlations in the α_γ subunit is robust with respect to the correlation measure.

We have also performed a principal components analysis (PCA). PCA is a transformation generated by diagonalising the covariance matrix of the atomic fluctuations from their average positions, as generated from the ensemble of stored conformations. The elements of the covariance matrix are defined by:

$$\Gamma_{ij} = \langle (x_i(t) - \langle x_i(t) \rangle) \cdot (x_j(t) - \langle x_j(t) \rangle) \rangle \quad (2)$$

where $x_i(t)$ and $x_j(t)$ are the Cartesian coordinates for atoms i and j at time t, with the angle brackets indicating averaging over the ensemble. The new coordinate basis defined by the eigenvectors (arranged in descending order of eigenvalue) has the property that each eigenvector accounts for progressively less of the observed variance in the data.