



Article

# Structural Prediction of the Dimeric Form of the Mammalian Translocator Membrane Protein TSPO: A Key Target for Brain Diagnostics

Juan Zeng <sup>1,2,†</sup> , Riccardo Guareschi <sup>1,†</sup>, Mangesh Damre <sup>3,4,†</sup> , Ruyin Cao <sup>1</sup>, Achim Kless <sup>5</sup> , Bernd Neumaier <sup>6</sup> , Andreas Bauer <sup>7</sup>, Alejandro Giorgetti <sup>1,3,\*</sup>, Paolo Carloni <sup>1,8</sup> and Giulia Rossetti <sup>1,9,10,\*</sup>

- <sup>1</sup> Institute for Advanced Simulations (IAS)-5/Institute for Neuroscience and Medicine (INM)-9, Forschungszentrum Jülich, 52428 Jülich, Germany; j.zeng@fz-juelich.de (J.Z.); r.guareschi@fz-juelich.de (R.G.); caobb0214@gmail.com (R.C.); p.carloni@fz-juelich.de (P.C.)
- <sup>2</sup> Laboratory of Computational Chemistry and Drug Design, Laboratory of Chemical Genomics, Peking University Shenzhen Graduate School, 518055 Shenzhen, China
- <sup>3</sup> Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy; mangeshdamre007@gmail.com
- <sup>4</sup> Neurobiology, International School for Advanced Studies (SISSA), 34136 Trieste, Italy
- <sup>5</sup> Grünenthal Innovation, Translational Science & Intelligence, Grünenthal GmbH, 52078 Aachen, Germany; Achim.Kless@grunenthal.com
- <sup>6</sup> Institute for Neuroscience and Medicine (INM)-5, Forschungszentrum Jülich, 52428 Jülich, Germany; b.neumaier@fz-juelich.de
- <sup>7</sup> Institute for Neuroscience and Medicine (INM)-2, Forschungszentrum Jülich, 52428 Jülich, Germany; an.bauer@fz-juelich.de
- <sup>8</sup> RWTH Aachen University, Department of Physics, 52078 Aachen, Germany
- <sup>9</sup> Jülich Supercomputing Center (JSC), Forschungszentrum Jülich, 52428 Jülich, Germany
- <sup>10</sup> University Hospital Aachen, RWTH Aachen University, 52078 Aachen, Germany
- \* Correspondence: alejandro.giorgetti@univr.it (A.G.); g.rossetti@fz-juelich.de (G.R.); Tel.: +39-045-802-7905 (A.G.); +49-246-161-8933 (G.R.); Fax: +39-045-802-7929 (A.G.); +49-246-161-4823 (G.R.)
- † These authors contributed equally to this work.

Received: 25 July 2018; Accepted: 28 August 2018; Published: 31 August 2018



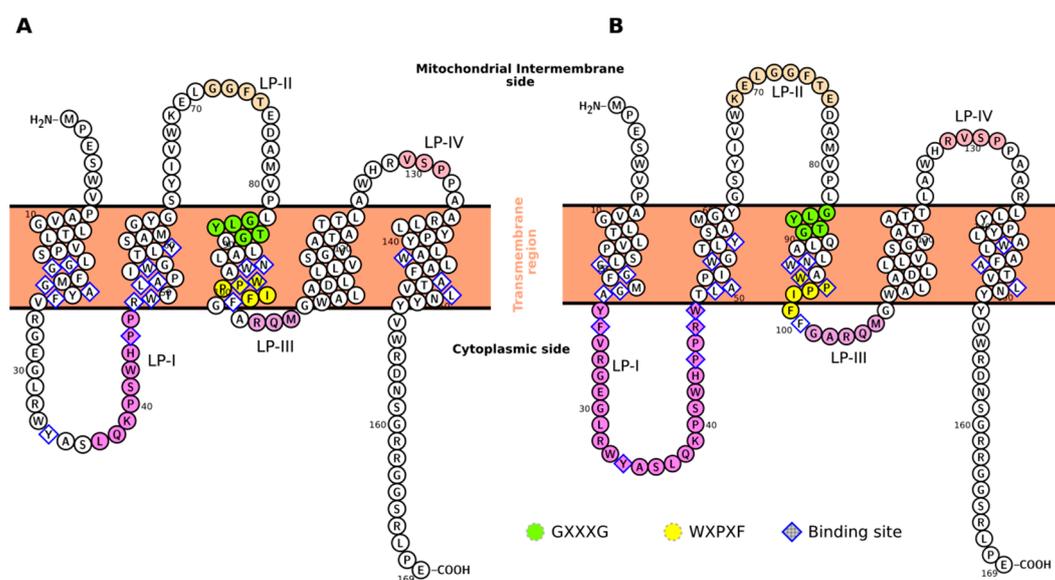
**Abstract:** Positron emission tomography (PET) radioligands targeting the human translocator membrane protein (TSPO) are broadly used for the investigations of neuroinflammatory conditions associated with neurological disorders. Structural information on the mammalian protein homodimers—the suggested functional state of the protein—is limited to a solid-state nuclear magnetic resonance (NMR) study and to a model based on the previously-deposited solution NMR structure of the monomeric mouse protein. Computational studies performed here suggest that the NMR-solved structure in the presence of detergents is not prone to dimer formation and is furthermore unstable in its native membrane environment. We, therefore, propose a new model of the functionally-relevant dimeric form of the mouse protein, based on a prokaryotic homologue. The model, fully consistent with solid-state NMR data, is very different from the previous predictions. Hence, it provides, for the first time, structural insights into this pharmaceutically-important target which are fully consistent with experimental data.

**Keywords:** TSPO; radioligands; PET; oligomerization; brain inflammation; homology modeling; molecular dynamics; docking

## 1. Introduction

The human translocator membrane protein (TSPO) [1–3] is a key biomarker for the diagnostics of inflammatory conditions in the brain [4]. Expression of human TSPO in the outer membrane of the mitochondria is indeed strongly up-regulated in areas of brain injury and in neuroinflammatory conditions, including those associated with Alzheimer’s and Parkinson’s disease [5]. Increased expression levels of human TSPO can be monitored by positron emission tomography (PET). PET uses radiolabeled human TSPO ligands to sensitively recognize lesions and active disease processes of the brain [6–8]. In addition, selective human TSPO ligands are expected to be therapeutic agents with a wide spectrum of action against psychiatric disorders and limited side effects [9,10]. Several specific diagnostic ligands show potent activity ( $K_i$ : 0.18–11 nM) for both rat and human TSPOs [11]. More recently, other studies have investigated the use of low-affinity TSPO binders to treat traumatic brain injury and brain inflammation [12–15]. These works have pointed out that also low-affinity binders of TSPO can be effectively employed to control apoptosis, to reduce neuronal degeneration and promote neurosteroidogenesis.

However, the mode of action of these ligands is still unclear due to the controversial structural information on TSPO. The protein topology consists of five transmembrane helices (TM-I to TM-V) connected by two loops (LP-I and LP-III) at the cytosol interface and two loops (LP-II and LP-IV) placed in the intermembrane space (Figure 1). Mammalian TSPOs can assemble in multimeric complexes (up to six units) *in vivo* [16,17], but the dimeric form is likely to be already functional [2,18]. Unfortunately, no oligomeric mammalian experimental structure is available.



**Figure 1.** Topology of mTSPO for the models here considered. Specific regions and motifs discussed in the text are highlighted for mTSPO\_NMR\_monomer (A) and mTSPO\_Rs\_monomer (B). The topology is generated using the Protter web application [19].

Current information on the structural determinants and the subunit-subunit assembly of mammalian TSPOs is derived from solid-state NMR investigations on the mouse protein (mTSPO) [20]. This is expected to be similar to the human TSPO, because of the high mutual sequence identity (81%, as obtained from the analysis of the sequence identity performed with BLAST [21,22]). mTSPO’s monomer-monomer interface involves residues from the TM-III helix [20] (Figure 1 and Table 1). These include: (i) the G83XXXG87 (X = any residue) motif present across different protein classes’ interfaces, [23–27]; and (ii) the W95XPXF98 motif [28], occurring across eukaryotic and prokaryotic TSPOs. In the same study, the authors also presented a prediction of the functionally relevant dimer structure, based on a previously

published solution NMR structure of monomeric mTSPO [29] in complex with its high affinity ligand 1-(2-chlorophenyl)-*N*-methyl-*N*-(1-methylpropyl)-3-isoquinolinecarboxamide (PK11195) ( $K_d = 2.1\text{--}28$  nM, see Figure S1) [7,30,31] (mTSPO\_NMR\_monomer in this study, PDBiD: 2MGY). Unfortunately, this model [20] shows some inconsistencies. These are likely to be caused, at least in part, by intrinsic limitations of the solution NMR structural template [29]. Indeed, the protein's conformation may be affected by the use of ionic detergents (decylphosphocholine micelles) during the purification and the NMR measurements [28,32]. This concern is corroborated by binding assays of PK11195 in ionic detergents [30,32]: the affinity is abolished by purifying the protein in a sodium dodecyl sulfate detergent, and it is restored up to about 1.5 nM after the removal of SDS and reconstruction into liposomes [30]. Hence, the proposed solution NMR monomer structure may be in a conformational state different from the one in the mitochondrial membrane [28].

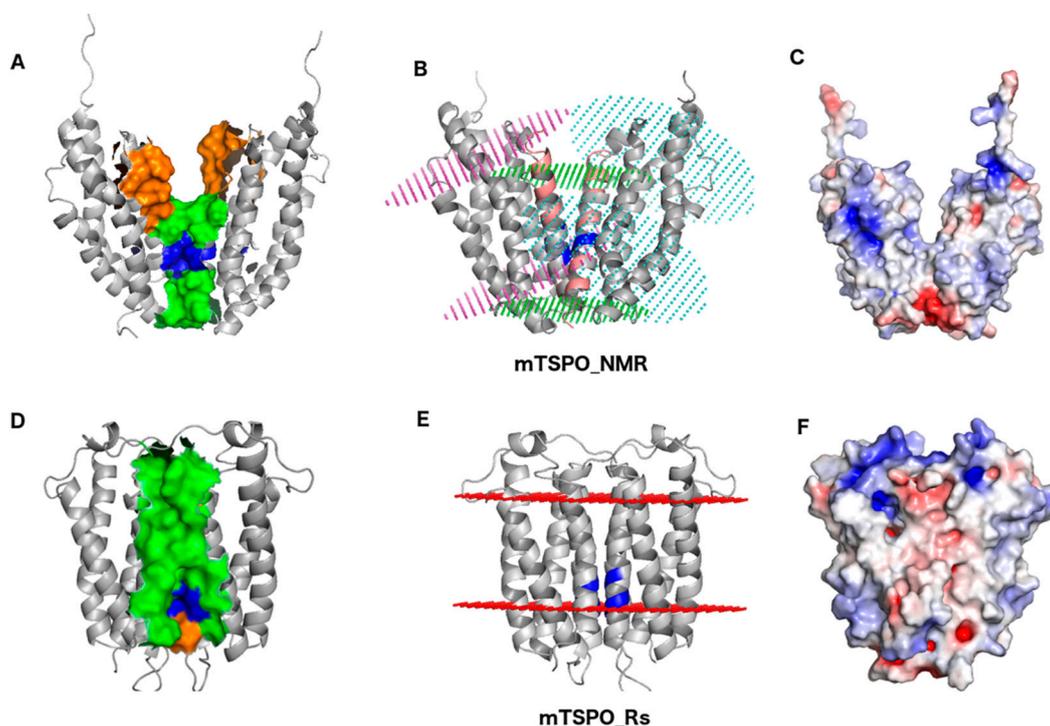
**Table 1.** Residues of the mTSPO/PK11195 complex located either in the protein binding cavities or at the dimer interface, according to the experimental data [20] and to the proposed dimer models. These models are based either on the NMR structure (mTSPO\_NMR, [20] and this work) or on the protein from *Rhodobacter Sphaeroides* (RsTSPO, this work). Residues at the interface accordingly to both the NMR data [20] and mTSPO\_Rs are in bold. mTSPO\_Rs' binding pockets do not include exactly the same residues because of differences of the ligands' binding poses, as obtained with docking simulations described in the text. On the contrary, mTSPO\_NMR binding pockets include by construction exactly the same residues. The residues within 5 Å from the ligand are assigned to the binding cavity and reported in this table. The corresponding list of the residues within 4 Å and 6 Å from the ligand is shown in Table S6 and Figure S5. We report also a list of the residues with a conservation rate higher than 85% in multiple sequence alignment across 148 homologous sequences to the mTSPO, as determined with the ConSurf server [33], along with the topological area of the protein where these sequences are found. A detailed discussion of the evolutionary coupling of the residues is presented in the Results section and in Figure S6.

| Region         | Inferred by Experiment  | mTSPO_NMR  |  | mTSPO_Rs  |  |  |  |
|----------------|---|--|--|---|--|--|--|
|                |   | Subunit A  | Subunit B  | Subunit A   | Subunit B  |  |  |
| Binding Cavity | G19, A23, V26, R27, H43, R46, L49, A50, I52, W53, W95, W107, A110, D111, L114, W143, A147, L150, N151 | G18, G19, F20, G22, A23, V26, R27, G30, L31, K39, P40, S41, H43, P44, P45, R46, L49, A50, I52, W53, W93, W95, W107, A108, A110, D111, L114, W143, F146, A147, T148, L150, N151 | G18, G19, F20, G22, A23, V26, R27, G30, L31, K39, P40, S41, H43, P44, P45, R46, L49, A50, I52, W53, W93, W95, W107, A108, A110, D111, L114, W143, F146, A147, T148, L150, N151 | P15, G18, G19, M21, G22, A23, F25, V26, R27, G28, E29, Y34, K39, H43, P44, R46, L49, A50, W53, W53, G54, Y57, N92, W93, W95, P96, F99, F100, L112, W114, F146, A147, T148, L150, N151, V154   | G18, M21, G22, A23, F25, Y34, H43, P44, R46, L49, A50, W53, L56, Y57, N92, W93, A9, W95, P96, P97, F99, F100, L112, V110, Y140, L141, W143, A147, L150 |  |  |
|                |   |  |  | V6, P7, G10, L11, L13, V14, L17, G18, F20, M21, Y24 V26, R27(A), M79, <b>V80</b> , L82, <b>G83</b> , L84, Y85, T86, G87, A90, L91, <b>W93</b> , A94, P97, <b>I98</b> , <b>A102</b> , Q104, W107, A108, A110, <b>D111</b> , L114, <b>V118</b> , A121, A125 |  |  |  |
|                |   | Dimer Interface  | <b>V80</b> , <b>G83</b> , Q88, N92, <b>W93</b> , W95, <b>I98</b> , F100, G101, <b>A102</b> , <b>D111</b> , <b>V118</b>   | F74, T75, E76, D77, M79, <b>V80</b> , P81, <b>G83</b> , L84, T86, G87, Q88, A90, L91  |  |  |  |
|                |   | Conserved Residues   | <b>LP-I</b> : L37, P40, P44, P45, <b>TM-II</b> : W53, L56, G61, <b>TM-III</b> : N92, W95, F99, F100, <b>TM-V</b> : L136, P139, Y140, W143, A147, L150, N151                    |   |  |  |  |
|                |   | Evolutionary Coupling  | P40, P45 coevolve with L150; P44, P45 with W95; W53 with L56, A147, L150; W95 with A147 and N151   |   |  |  |  |

Prompted by the pharmacological importance of the protein, here we have constructed a new model of dimeric mTSPO. We use as a template the prokaryotic protein X-ray structure, which features the same fold as the eukaryotic one [34]. Our prediction turns out to be fully consistent with the available experimental topological information, thus delving for the first time into the structural determinants of this exceptionally important target for pharmacological interventions and diagnostics.

## 2. Results

The only model reported so far for the functional, dimeric form of eukaryotic TSPOs (mTSPO\_NMR hereafter) is based on mTSPO\_NMR\_monomer structure [29]. As mTSPO\_NMR's 3D coordinates are currently not available, we have repeated the procedure reported in [20] to construct and analyze it. The mTSPO\_NMR model (Figure 2A) reveals some drawbacks: (i) Residues N92, W93, W95, I98, F100, G101, and A102, experimentally identified to be located at the monomer-monomer interface, are separated by a distance of 9 Å or more (Table 1 and Figure 2A); (ii) Conversely, there are six residues, from F74 to M79, located at the dimer interface, that were found elsewhere experimentally (Table 1 and Figure 2A); (iii) The predicted embedding of the protein into the membrane (as obtained with the 'Positioning of Proteins in Membranes', PPM, server [35]) shows that the dimer is significantly tilted with respect to the membrane plane. As a result, the thickness of the hydrophobic region of this model is far smaller (14 Å) than that required in a biological hydrophobic membrane region (Figure 2B and Table S2). Therefore, it is unlikely that mTSPO\_NMR can be correctly embedded in a biological membrane. We underline that the embedding in the mitochondria membrane is compatible with the generalized membrane used by the PPM server (<http://opm.phar.umich.edu/protein.php?search=tspo>). Indeed, the PPM database contains the monomer of mouse TSPO embedded in mitochondria membrane and either the tilt and membrane thickness is fully compatible with our modeled structure (Figure 2E); (iv) Several charged residues are exposed towards the membrane (see Figure 2C); (v) PK11195 forms highly stabilizing interactions in the *Bacillus cereus* TSPO/PK11195 X-ray structure (PDBiD: 4RYI) (these are  $\pi$ -stacking interactions with F90 and hydrogen bonds (H-bonds) with W51 (Figure S2B,C)) [36], not present in this model. The ligand forms here mostly hydrophobic interactions with the protein (Figure S2A). In particular, the two residues corresponding in the sequence alignment to F90 and W51 of the prokaryotic protein (W95 and W53) are not oriented in an optimal way inside the binding pocket (Figure S2), thus impeding a similar interaction pattern. This finding is not consistent with the experimental evidence that PK11195 binds stronger to mammalian TSPOs [7,30,31] than to prokaryotic ones [32].



**Figure 2.** The mTSPO\_NMR (A–C) and mTSPO\_Rs dimers (D–F). (A) The G83XXXG87 motif (blue surface) of mTSPO\_NMR is located at the dimer interface, consistently with the experimental findings [20] (Table 1). Other residues participating to the dimer interface in the model and according to experiment [20] are represented by green and orange surfaces, respectively. (B) Embedding of mTSPO\_NMR in a bilayer membrane according to the PPM server [35]. The membrane boundaries for the individual monomers are shown by pink and cyan spheres. Glycophorin A (PDBiD: 5EH4, shown as pink ribbons) has been used as template to build the dimer, following the procedure of [20]. The membrane boundaries of this protein derived from the Orientation of Proteins in Membranes (OPM) database [35] are shown as green spheres. The membrane boundaries of mTSPO\_NMR are not parallel to those of the template guiding the dimerization. (C) mTSPO\_NMR electrostatic surface potential. The surface potentials are calculated using APBS [37]. Red and blue surfaces represent negative and positive electrostatic potentials ( $-5$  kT/e,  $+5$  kT/e), respectively. The exposed positively charged (blue) surface in are generated by R27, R46, R156, and R161. (D) The blue surface shows the G83XXXG87 motif of mTSPO\_Rs. The green surface shows the other residues participating to the dimer interface in each monomer. The orange surface shows the residues participating to the dimer interface according to the experimental assignments data [20] not already included in the previous selections. (E) Embedding of mTSPO\_Rs in a bilayer membrane according to the PPM server calculations. The membrane boundaries are shown in red. (F) mTSPO\_Rs electrostatic surface potential. These residues that are exposed towards the lipid membrane in C, have now the sidechains oriented towards the cytoplasm.

We conclude that these inconsistencies with experimental data may limit the predictive power of this model. Part of these drawbacks might be caused by limitations of the template structure used.

Therefore, we decided to build a new model (mTSPO\_Rs hereafter), based on homology modeling using the corresponding prokaryotic dimer from RsTSPO, for which an X-ray structure is available (PDBiD: 4UC1) and whose folding is comparable with the ones of the eukaryotic protein [34]. This template appears rather suitable for this prediction.

Most importantly, our prediction turns out to be consistent with the results from solid-state NMR (Figure 2). The two subunits of mTSPO\_Rs arrange very differently than those of mTSPO\_NMR: they are oriented almost perpendicularly to the membrane bilayer. As a result, the thickness of the hydrophobic region ( $\sim 22$  Å, see Table S2 and Figure 2D,E) is compatible with that of other membrane

proteins [35]. Therefore, mTSPO\_Rs can be embedded into a membrane bilayer. The side chains of the charged residues point toward the cytosolic side, as expected for a membrane protein (Figure 2F). This observation is in line with [34], which compares the electrostatic surface of RsTSPO with that of mTSPO\_NMR\_monomer. The dimer interface is mainly hydrophobic (Table 1). It extends along 33 Å. It involves (i) V80, G83, W93, I98, and A102 from TM-III and V118 from TM-IV, consistently with experimental data [20] (Table 1); (ii) the G83XXXG87 motif, also in agreement with experimental data [20]; and (iii) the W95XPXF99 motif. Here, the side chains of P97 and I98 are directed to the interface, while the side chains of W95, P96, and F99 point towards the inner binding cavity of the protein. This is consistent with the experimental findings that assigns to the dimer interface residues W95 and I98 from this motif [20]. The other residues occurring at the dimer interface in the model, are placed in the proximity of other residues identified experimentally. Thus, we cannot exclude that small structural adjustments in the area of the interface could optimize the interactions, yielding monomer-monomer contacts fully consistent with the experimental assignments.

The topology of the monomers in the proposed model presents few, still significant differences with respect the ones in mTSPO\_NMR (Figure 1B) in the dispositions of the transmembrane and loops regions. In turn, the position of several residues located in the binding site differ markedly from those of mTSPO\_NMR (Figure S3): W95, F99, and F100 orient towards the binding cavity, creating an aromatic pocket suitable to accommodate the PK11195 ligand and offering multiple potential anchor points for  $\pi$ -stacking interactions. Within the well-known limitations associated with the docking procedures to a homology model [38], the induced-fit docking [39,40] of PK11195 using Glide [41–43] and MOE [44,45] scoring functions, confirms that this is the case. Indeed, the ligand establishes at least one  $\pi$ -stacking and one H-bond in each of the binding sites. The residues involved in these interactions are W53, W93, and P96 (Figure S2D,E). As expected, in both subunits the ligand shares the same binding pose, and only small differences can be observed in the two binding pockets. The interactions patterns are not too different from those found in BcTSPO/PK11195 X-ray structure [36]. With the same approach, the binding modes of other known mammalian TSPO tracers in mTSPO\_Rs have been also predicted (see Table S7 of the Supporting Information). The protein-ligand interactions turn out to be overall similar to those discussed for PK11195.

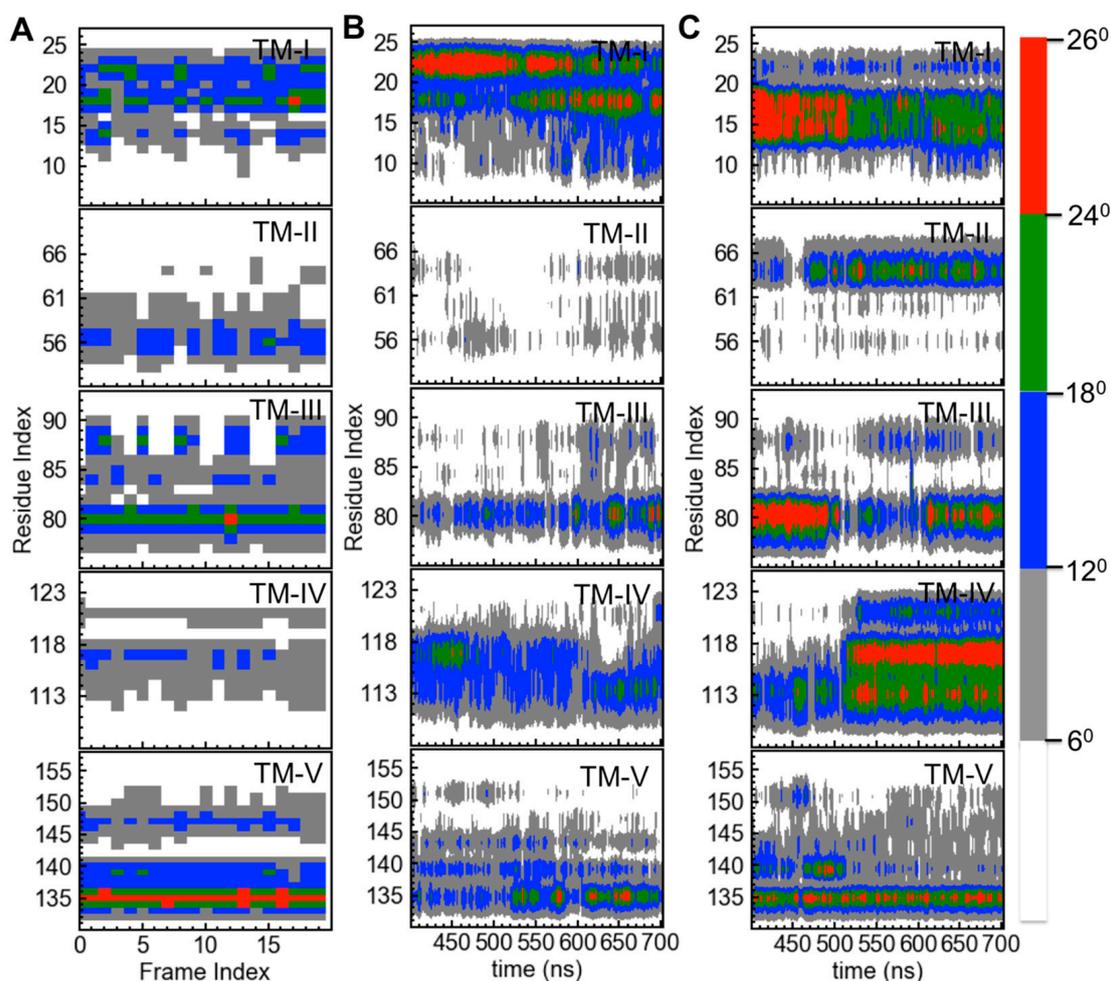
We analyzed the coevolution among relevant regions of the protein (Table 1 and Figure S6). W95 from the W95XP97XF99 motif, and N151, belonging to the cholesterol recognition amino acid consensus (CRAC) motif (residues 149–156) [20,46,47] have coevolved. Interestingly, the W95XP97XF99 motif has also coevolved with binding site residues in LP-I and TM-II residues. Moreover, CRAC domain residue L150 has coevolved with residues of the binding region. These include P40, P44, and T48, placed in LP-I, as well as W53 placed in TM-II. Furthermore, residues Y140 and L144 belonging to the mirror cholesterol recognition amino acid consensus motif (CARC) [47] have coevolved with binding-site residues of TM-II (W53, Y57). Our model shows that the W95XP97XF99 motif is involved in ligand binding and receptor dimerization. Our coevolution analysis accordingly suggest that W95XP97XF99 is evolutionary linked to other binding sites residues, as well as with the cholesterol-binding motif. This is in line with the suggested role of cholesterol in regulating TSPO dimerization [20]. Notably, the W95XP97XF99 motif is located at the dimer interface only in our model.

We conclude that our prediction is overall in good agreement with experimental data and appears a reasonable model of the functional form of mTSPO. In the following section, we discuss the feasibility of the template structures used for the structural predictions discussed above.

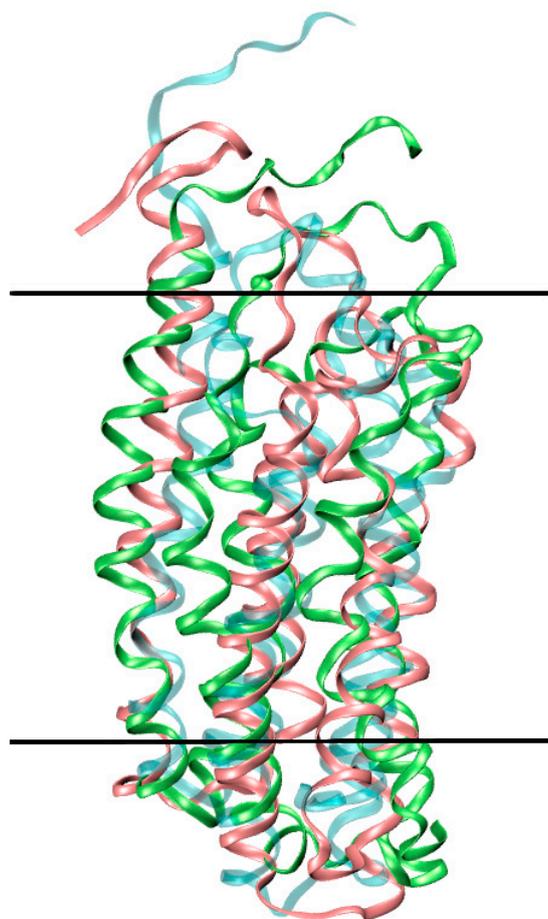
#### *Analysis of the Templates for mTSPO\_NMR and for mTSPO\_Rs*

The ionic detergents used to determine the mTSPO\_NMR\_monomer structure [29] may affect significantly the structure of the receptor. Indeed, these modify the binding affinity of TSPO for PK11195 [30,32], possibly because of structural changes of the receptor. 700-ns MD simulations of mTSPO\_NMR\_monomer with and without ligand, in explicit solvent and embedded in a membrane environment (see Methods for details) suggest that this is indeed the case. The obtained conformations

appeared to be equilibrated after 400 ns and they lead to a different conformation compared with the NMR structure (See SI). While the ligand mostly retains its binding pose (see SI), the bending of the helices (particularly of TM-II and TM-IV, Figure 3) increases, especially in the presence of the ligand. This leads to a distortion of the helix bundle (Figure 4) and to a shrinking of the protein along the direction perpendicular to the membrane, especially in the holo protein.



**Figure 3.** The helix bending in mTSPO\_NMR\_monomer. The values are either reported for the deposited 20 NMR structures of mTSPO\_NMR\_monomer (PDBiD: 2MGY) (A), for mTSPO\_NMR\_monomer without (B) or with PK11195 (C) as a function of the simulation time. Only the last 300 ns are shown. The Figure was made using the “bendix” plugin of the VMD program [48].



**Figure 4.** Distortion of mTSPO\_NMR\_monomer in the lipid bilayer. The initial NMR structure (cyan) is superimposed to mTSPO\_NMR\_monomer in complex with PK11195 (green) and mTSPO\_NMR\_monomer without ligand (pink) after MD equilibration. The boundaries of the membrane are approximately shown as black lines.

Moreover, we have analyzed the membrane embedding of the monomers that build our new mTSPO\_Rs dimer and compared it to the embedding of the mTSPO\_NMR\_monomer structure (PDBiD: 2MGY). Using the QMEANBrane program [49], we observe that the monomer of our model (mTSPO\_Rs\_monomer) shows higher scores than the NMR monomeric structure, indicating a better membrane embedding already at the monomeric level (Figure S4).

All the multimeric TSPO structures are either dimers or trimers of RsTSPO from *Rhodobacter sphaeroides* and BcTSPO from *Bacillus cereus* (Table S1). They have a sequence identity with mTSPO is of 32% and 27%, respectively [28]. These proteins (i) have been crystallized in lipidic cubic phase conditions [34,50] which have proven to be a reliable method to determine X-ray structures of membrane proteins; (ii) they present a correctly folded membrane protein according to the QMEANBrane server, which takes into account general statistical characteristics of membrane proteins, but not membrane composition [49]; and (iii) they feature a characteristic motif located at the subunit/subunit interface (A75XXXA79 in TM-III of RsTSPO and G44XXXG48 in TM-II of BcTSPO). Both motifs are involved in the stabilization of helix-helix contacts across different transmembrane proteins [51]. The alignment of RsTSPO and mTSPO sequences (Table S1) shows that the A75XXXA79 motif matches the G83XXXG87 motif of mTSPO—as a crucial element in the dimerization of mTSPO [20]—and in the resulting model the latest is correctly located at the dimer interface. This is not the case for the G44XXXG48 motif, which does not align with the G83XXXG87 motif of mTSPO and in the resulting model it would not be located at the dimer interface. Thus,

BcTSPO, by lacking this crucial feature, appears not to be a suitable template for a structural prediction of the mTSPO functional dimer.

Although it shares a modest sequence identity with the mouse protein [28], all the results reported above leads us to suggest that RsTSPO represents a fairly good template for our structural predictions.

### 3. Discussion

In this paper we propose an alternative dimer model of mammalian TSPO based on the deposited X-ray structure of bacterial RsTSPO (mTSPO\_Rs) with respect the one based on the NMR structure of mTSPO (mTSPO\_NMR). The major differences can be appreciated from Figure 2 and they regard (i) the dimer interface, (ii) the embedding in a membrane bilayer, and (iii) the binding pocket and the effect on ligand binding. For all of these aspects, the mTSPO\_Rs dimer model appears more suitable than mTSPO\_NMR to describe the physiologically-relevant dimeric structure. The residues at the dimer interface in mTSPO\_Rs match those identified experimentally in solid state NMR [20], the simulated embedding of mTSPO\_Rs in a membrane is reasonable, and the predicted binding of the prototypical ligand PK11195 is favored in mTSPO\_Rs with respect to mTSPO\_NMR. Remarkably, our proposed model can rationalize the higher affinity of PK11195 for mammalian TSPO with respect the bacterial protein. This is not the case for mTSPO\_NMR, where the type and number of interactions do not seem consistent with the measured nanomolar affinity [7,30,31].

Our model also shows that the W95XP97XF99 motif is involved in both ligand binding and dimerization. Accordingly, we found that W95XP97XF99 residues have coevolved with residues responsible for ligand binding and with the ones in the cholesterol binding motifs. This is in agreement with the suggested role of cholesterol binding on the monomer-dimer equilibrium [20]. Furthermore, the combined molecular modeling and co-evolution analyses suggest a possible interplay between the W95XP97XF99 dimerization motif and cholesterol and ligand binding during evolution. Notably this motif is not part of the dimer interface in mTSPO\_NMR, while it is in mTSPO\_Rs.

In conclusion, our model is in agreement with the available experimental data and can explain several aspects of TSPO features related to dimerization and cholesterol binding.

### 4. Materials and Methods

#### 4.1. Bioinformatics Analyses

We used the ConSurf server (<http://consurf.tau.ac.il/2016/>) [33] to perform a conservation analysis across TSPO homologous sequences of TSPO. We set a cut-off value of 95% of sequence identity and minimal cut off value of 35% homologue identity to retrieve the homologous sequences to the mTSPO sequence. Further obtained multiple sequence alignment visualized online (<http://molsim.sci.univr.it/TSPO/>).

The mTSPO sequence was submitted to the HHPRED server [52,53] to identify the most suitable template. The Modeller program [54] then generated multiple homology models of the monomer based on the template. We used a very slow mode of MD annealing technique for precise refinements. We selected the best homology models according to DOPE [55] and GA341 scores [56], as well as the model quality according to the QMEANBrane module from Swiss-model [49] (Figure S4).

We cross checked our alignment and model by performing an extensive conservation analysis to identify the evolutionary relevance of the different amino acid positions in the protein, based on the phylogenetic relation between homologous sequences to the mTSPO. A total of 1940 sequences were used to build the Multiple Sequence Alignment, 1882 of which are unique and 148 are homologous sequences to the mTSPO (visualized online <http://molsim.sci.univr.it/TSPO/>). As discussed in the main text, we selected RsTSPO (PDBiD: 4UC1) as a template to build dimer for mTSPO. The homology model for the dimer has been generated using Modeller [54] and validated similarly as monomer homology model.

CoeViz, a web-based tool [57], is used to identify the evolutionary coupling among the residues and the functionally related residues in mTSPO. Evolutionary coupling residue pairs were selected based on their pairwise  $\chi^2$  scores [58] (cutoff value > 0.7, See Table 1 and Figure S6).

The Positioning of Proteins in Membranes (PPM) server [35] assigns spatial positions of molecules inside membranes by optimizing their transfer energy from water to the lipid bilayer [59]. It was used here to assess the embedding of each studied system in the membrane bilayer. Molecular graphics were drawn using Pymol [60] and VMD [61].

The accession numbers (PDBiD) of the proteins relevant for this work are: 2MGY, 4UC1, 4RYI, and 5EH4.

The 3D structures of the dimer models of mTSPO discussed in this work are available upon request.

#### 4.2. MD Simulations of mTSPO\_NMR\_monomer

The first frame of the NMR ensemble (PDBiD 2MGY) is embedded in a membrane composed by phosphatidylcholine (POPC, 31%), phosphatidylethanolamine (POPE, 41%) and cholesterol (CHL, 28%) and enclosed in a water box with dimensions (nm)  $10.1 \times 10.1 \times 12.8$ . The total charge of this system is kept neutral and the ionic strength is set to 0.15 M by adding potassium ( $K^+$ ) and chloride ( $Cl^-$ ) ions. The whole system includes mTSPO, PK11195, POPC (84 molecules), POPE (116 molecules), CHL (98 molecules), water (27126 molecules),  $K^+$  (118 ions), and  $Cl^-$  (124 ions). In the simulation of the apo protein, only the ligand PK11195 is removed.

The AMBER99SB-ILDN force fields [62], the Slipids [63,64], and TIP3P [65] were used for the protein and ions, the lipids, and water, respectively. The General Amber force field (GAFF) parameters [66] were used for PK11195, along with the RESP atomic charge [67,68] fitted with the electrostatic potential (ESP) from Gaussian 09 [69] calculation with the HF-6-31G\* basis set [70,71]. The topology of PK11195 was then converted to GROMACS format using the ACPYPE tool [72].

The Particle Mesh Ewald method [73] was used to treat the long-range electrostatic interaction with a real space cutoff of 1.2 nm. A 1.2 nm cutoff was also used for the short-range non-bonded interaction. A time step of 2 fs was set. The LINCS algorithm [74] was applied to constrain all bonds involving hydrogen atoms. Constant temperature and pressure conditions were achieved via independently coupling protein/PK11195, lipids, solvent, and ions to a Nosè-Hoover thermostat [75] at 315 K and Parrinello-Rahman barostat [76] at 1 atm. For each simulation, the systems underwent a two-steps minimization without restraints. The minimization with conjugated gradient method [77] followed the steepest descent method. Then 1-ns simulated annealing and 10-ns equilibration with positional restraint using a force constant of  $1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$  on the heavy atoms of the protein were carried out. The last 700-ns MD at 315 K and 1 atm was performed with frames collected every 20 ps.

The root mean square deviation of backbone atoms (N,  $C_\alpha$ , C atoms, bb-RMSD) and the root mean square fluctuation (RMSF) were calculated using the *g\_rms* and *g\_rmsf* codes from the Gromacs package [78]. The 13th structure of the mTSPO NMR ensemble (PDBiD: 2MGY) was selected as the reference. This choice was based on the results of pairwise RMSD analysis, hydrogen bonds count, and secondary structure analysis performed on the twenty structures of NMR ensemble. The secondary structure was defined and calculated with *do\_dssp* [79] embedded in Gromacs. The distortion of helix was evaluated with the Bendix [48] plug-in in VMD software [61], using default parameters. The *g\_cluster* module with *gromos* method in Gromacs [80] was used to perform cluster analysis on the converged part of the MD simulation (after 400 ns). The cutoff based on bb-RMSD for clustering is set to 1.2 Å. From the cluster analysis, the central frame of the most populated cluster is selected as representative structure of mTSPO\_NMR\_monomer after MD equilibration for the holo and apo proteins simulated. The *trj\_cavity* package [81] within GROMACS was selected to calculate the volume of the binding cavity, using a grid spacing of 1.4 Å. The number of H-bonds was calculated with *g\_hbond* in GROMACS. The PAD flexibility index [82], the distribution of geometry center, dihedral

angles, rotation angle of PK11195 and contact percentage between PK11195 and residues are all calculated with an in-house code. The contact percentage is defined as the minimum distance between any atoms of side chains and ligand, both excluding hydrogen atoms, less than 4.0 Å. Water molecule with distance less than 5 Å from the geometric center is accounted as molecules within the binding pocket. MD simulations were performed using Gromacs package [78] on supercomputer.

## 5. Conclusions

Here, we have presented a computational study aiming to shed light on the structural determinants of the functional form of mammalian TSPO. The reliability of the dimer construct based on the NMR monomer of the mouse protein ([20] and this work) appears weakened by some inconsistencies, possibly caused by the fact that the NMR experiments were carried out under conditions far from the physiological ones. Therefore, we have proposed an alternative model, quite different from the previous one, which is in good agreement with the available experimental findings and supports dimer formation. This model may foster pharmacological studies aimed at a rational design of novel TSPO targeting ligands.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/19/9/2588/s1>.

**Author Contributions:** J.Z. performed and analyzed the M.D. simulations and contributed to the writing of the manuscript; R.G. performed and analyzed the docking simulations and contributed to the writing of the manuscript; M.D. performed the bioinformatics calculations and contributed to the writing of the manuscript; R.C. set up the M.D. simulations; A.G., A.B., B.N., P.C., A.K., and G.R. conceived the project and contributed to the writing of the manuscript.

**Funding:** Funding from the Dual<sup>2</sup>Pet grant by the European Union and North Rhine-Westphalia (reference number EFRE-0800988) is acknowledged. We thank the National Natural Science Foundation of China for financial support (grant no. 21603006 to J.Z). International Postdoctoral Exchange Fellowship Program 2016 under grant no. 2016002 also supported this work.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

|                   |  |
|-------------------|--|
| TSPO              | Translocator membrane protein  |
| PET               | Positron emission tomography   |
| RsTSPO            | <i>Rhodobacter Sphaeroides</i>   |
| BcTSPO            | <i>Bacillus Cereus</i>   |
| mTSPO_NMR_monomer | Monomer of mTSPO solved by solution NMR experiment, PDBiD: 2MGY              |
| mTSPO_NMR         | Dimer model of mTSPO. The prediction is based on mTSPO_NMR_monomer structure |
| mTSPO_Rs          | Dimer model of mTSPO. The prediction is based on the RsTSPO structure        |
| MD                | Molecular dynamics   |
| bb-RMSD           | Root-mean square deviation of backbone atoms (N, C <sub>α</sub> , C atoms)   |
| RMSF              | Root-mean square fluctuation   |

## References

1. Papadopoulos, V.; Baraldi, M.; Guilarte, T.R.; Knudsen, T.B.; Lacapère, J.-J.; Lindemann, P.; Norenberg, M.D.; Nutt, D.; Weizman, A.; Zhang, M.-R.; et al. Translocator protein (18kDa): New nomenclature for the peripheral-type benzodiazepine receptor based on its structure and molecular function. *Trends Pharmacol. Sci.* **2006**, *27*, 402–409. [[CrossRef](#)] [[PubMed](#)]
2. Beatriz, C.; Leo, V.; Moshe, G. Role of mitochondrial translocator protein (18 kDa) on mitochondrial-related cell death processes. *Recent Pat. Endocr. Metab. Immune Drug Discov.* **2013**, *7*, 86–101. [[CrossRef](#)]
3. Papadopoulos, V. On the role of the translocator protein (18-kDa) TSPO in steroid hormone biosynthesis. *Endocrinology* **2014**, *155*, 15–20. [[CrossRef](#)] [[PubMed](#)]

4. Costa, E.; Auta, J.; Guidotti, A.; Korneyev, A.; Romeo, E. The pharmacology of neurosteroidogenesis. *J. Steroid Biochem. Mol. Biol.* **1994**, *49*, 385–389. [[CrossRef](#)]
5. Rupperecht, R.; Papadopoulos, V.; Rammes, G.; Baghai, T.C.; Fan, J.; Akula, N.; Groyer, G.; Adams, D.; Schumacher, M. Translocator protein (18 kDa) (TSPO) as a therapeutic target for neurological and psychiatric disorders. *Nat. Rev. Drug Discov.* **2010**, *9*, 971–988. [[CrossRef](#)] [[PubMed](#)]
6. Gerhard, A. TSPO imaging in parkinsonian disorders. *Clin. Transl. Imaging* **2016**, *4*, 183–190. [[CrossRef](#)] [[PubMed](#)]
7. Trapani, A.; Palazzo, C.; de Candia, M.; Lasorsa, F.M.; Trapani, G. Targeting of the translocator protein 18 kDa (TSPO): A valuable approach for nuclear and optical imaging of activated microglia. *Bioconjugate Chem.* **2013**, *24*, 1415–1428. [[CrossRef](#)] [[PubMed](#)]
8. Dolle, F.; Luus, C.; Reynolds, A.; Kassiou, M. Radiolabelled molecules for imaging the translocator protein (18 kDa) using positron emission tomography. *Curr. Med. Chem.* **2009**, *16*, 2899–2923. [[CrossRef](#)] [[PubMed](#)]
9. Kim, T.H.; Pae, A.N. Translocator protein (TSPO) ligands for the diagnosis or treatment of neurodegenerative diseases: A patent review (2010–2015; part 1). *Expert Opin. Ther. Pat.* **2016**, *26*, 1325–1351. [[CrossRef](#)] [[PubMed](#)]
10. Kim, T.H.; Pae, A.N. Translocator protein (TSPO) ligands for the diagnosis or treatment of neurodegenerative diseases: A patent review (2010–2015; part 2). *Expert Opin. Ther. Pat.* **2016**, *26*, 1353–1366. [[CrossRef](#)] [[PubMed](#)]
11. Fukaya, T.; Kodo, T.; Ishiyama, T.; Kakuyama, H.; Nishikawa, H.; Baba, S.; Masumoto, S. Design, synthesis and structure-activity relationships of novel benzoxazolone derivatives as 18kDa translocator protein (TSPO) ligands. *Bioorg. Med. Chem.* **2012**, *20*, 5568–5582. [[CrossRef](#)] [[PubMed](#)]
12. Vainshtein, A.; Veenman, L.; Shterenberg, A.; Singh, S.; Masarwa, A.; Dutta, B.; Island, B.; Tsoglin, E.; Levin, E.; Leschiner, S.; et al. Quinazoline-based tricyclic compounds that regulate programmed cell death, induce neuronal differentiation, and are curative in animal models for excitotoxicity and hereditary brain disease. *Cell Death Discov.* **2015**, *1*, 15027. [[CrossRef](#)] [[PubMed](#)]
13. Simon-O'Brien, E.; Gauthier, D.; Riban, V.; Verleye, M. Etifoxine improves sensorimotor deficits and reduces glial activation, neuronal degeneration, and neuroinflammation in a rat model of traumatic brain injury. *J. Neuroinflammation* **2016**, *13*, 203. [[CrossRef](#)] [[PubMed](#)]
14. Costa, B.; Cavallini, C.; Da Pozzo, E.; Taliani, S.; Da Settimo, F.; Martini, C. The anxiolytic etifoxine binds to TSPO Ro5-4864 binding site with long residence time showing a high neurosteroidogenic activity. *ACS Chem. Neurosci.* **2017**, *8*, 1448–1454. [[CrossRef](#)] [[PubMed](#)]
15. Chen, Y.; Veenman, L.; Singh, S.; Ouyang, F.; Liang, J.; Huang, W.; Marek, I.; Zeng, J.; Gavish, M. 2-Cl-MGV-1 ameliorates apoptosis in the thalamus and hippocampus and cognitive deficits after cortical infarct in rats. *Stroke* **2017**, *48*, 3366–3374. [[CrossRef](#)] [[PubMed](#)]
16. Delavoie, F.; Li, H.; Hardwick, M.; Robert, J.-C.; Giatzakis, C.; Péranzi, G.; Yao, Z.-X.; Maccario, J.; Lacapère, J.-J.; Papadopoulos, V. In vivo and in vitro peripheral-type benzodiazepine receptor polymerization: Functional significance in drug ligand and cholesterol binding. *Biochemistry* **2003**, *42*, 4506–4519. [[CrossRef](#)] [[PubMed](#)]
17. Boujrad, N.; Vidic, B.; Papadopoulos, V. Acute action of choriogonadotropin on Leydig tumor cells: Changes in the topography of the mitochondrial peripheral-type benzodiazepine receptor. *Endocrinology* **1996**, *137*, 5727–5730. [[CrossRef](#)] [[PubMed](#)]
18. Papadopoulos, V.; Fan, J.; Zirkin, B. Translocator protein (18 kDa): An update on its function in steroidogenesis. *J. Neuroendocrinol.* **2017**, *30*, e12500. [[CrossRef](#)] [[PubMed](#)]
19. Omasits, U.; Ahrens, C.H.; Müller, S.; Wollscheid, B. Protter: Interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **2014**, *30*, 884–886. [[CrossRef](#)] [[PubMed](#)]
20. Jaipuria, G.; Leonov, A.; Giller, K.; Vasa, S.K.; Jaremko, L.; Jaremko, M.; Linser, R.; Becker, S.; Zweckstetter, M. Cholesterol-mediated allosteric regulation of the mitochondrial translocator protein structure. *Nat. Commun.* **2017**, *8*, 14893. [[CrossRef](#)] [[PubMed](#)]
21. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)] [[PubMed](#)]

22. Altschul, S.F.; Wootton, J.C.; Gertz, E.M.; Agarwala, R.; Morgulis, A.; Schäffer, A.A.; Yu, Y.-K. Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* **2005**, *272*, 5101–5109. [[CrossRef](#)] [[PubMed](#)]
23. Brosig, B.; Langosch, D. The dimerization motif of the glycophorin a transmembrane segment in membranes: Importance of glycine residues. *Protein Sci.* **2008**, *7*, 1052–1056. [[CrossRef](#)] [[PubMed](#)]
24. Russ, W.P.; Engelman, D.M. The GxxxG motif: A framework for transmembrane helix-helix association. *J. Mol. Biol.* **2000**, *296*, 911–919. [[CrossRef](#)] [[PubMed](#)]
25. Senes, A.; Gerstein, M.; Engelman, D.M. Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with  $\beta$ -branched residues at neighboring positions. *J. Mol. Biol.* **2000**, *296*, 921–936. [[CrossRef](#)] [[PubMed](#)]
26. Senes, A.; Ubarretxena-Belandia, I.; Engelman, D.M. The  $C\alpha$ —H $\cdots$ O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9056. [[CrossRef](#)] [[PubMed](#)]
27. Doura, A.K.; Fleming, K.G. Complex interactions at the helix-helix interface stabilize the glycophorin a transmembrane dimer. *J. Mol. Biol.* **2004**, *343*, 1487–1497. [[CrossRef](#)] [[PubMed](#)]
28. Li, F.; Liu, J.; Liu, N.; Kuhn, L.A.; Garavito, R.M.; Ferguson-Miller, S. Translocator protein 18 kDa (TSPO): An old protein with new functions? *Biochemistry* **2016**, *55*, 2821–2831. [[CrossRef](#)] [[PubMed](#)]
29. Jaremko, L.; Jaremko, M.; Giller, K.; Becker, S.; Zweckstetter, M. Structure of the mitochondrial translocator protein in complex with a diagnostic ligand. *Science* **2014**, *343*, 1363–1366. [[CrossRef](#)] [[PubMed](#)]
30. Lacapère, J.-J.; Delavoie, F.; Li, H.; Péranzi, G.; Maccario, J.; Papadopoulos, V.; Vidic, B. Structural and functional study of reconstituted peripheral benzodiazepine receptor. *Biochem. Biophys. Res. Commun.* **2001**, *284*, 536–541. [[CrossRef](#)] [[PubMed](#)]
31. Scarf, A.M.; Kassiou, M. The translocator protein. *J. Nucl. Med.* **2011**, *52*, 677–680. [[CrossRef](#)] [[PubMed](#)]
32. Li, F.; Xia, Y.; Meiler, J.; Ferguson-Miller, S. Characterization and modeling of the oligomeric state and ligand binding behavior of purified translocator protein 18 kDa from *Rhodospirillum rubrum*. *Biochemistry* **2013**, *52*, 5884–5899. [[CrossRef](#)] [[PubMed](#)]
33. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **2016**, *44*, W344–W350. [[CrossRef](#)] [[PubMed](#)]
34. Li, F.; Liu, J.; Zheng, Y.; Garavito, R.M.; Ferguson-Miller, S. Crystal structures of translocator protein (TSPO) and mutant mimic of a human polymorphism. *Science* **2015**, *347*, 555–558. [[CrossRef](#)] [[PubMed](#)]
35. Lomize, M.A.; Pogozheva, I.D.; Joo, H.; Mosberg, H.I.; Lomize, A.L. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* **2012**, *40*, D370–D376. [[CrossRef](#)] [[PubMed](#)]
36. Guo, Y.; Kalathur, R.C.; Liu, Q.; Kloss, B.; Bruni, R.; Ginter, C.; Kloppmann, E.; Rost, B.; Hendrickson, W.A. Structure and activity of tryptophan-rich TSPO proteins. *Science* **2015**, *347*, 551–555. [[CrossRef](#)] [[PubMed](#)]
37. Lerner, M.G.; Carlson, H.A. *APBS Plugin for PyMOL*; University of Michigan: Ann Arbor, MI, USA, 2006.
38. Chen, Y.-C. Beware of docking! *Trends Pharmacol. Sci.* **2015**, *36*, 78–95. [[CrossRef](#)] [[PubMed](#)]
39. Totrov, M.; Abagyan, R. Flexible ligand docking to multiple receptor conformations: A practical alternative. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178–184. [[CrossRef](#)] [[PubMed](#)]
40. Sotriffer, C.A. Accounting for induced-fit effects in docking: What is possible and what is not? *Curr. Top. Med. Chem.* **2011**, *11*, 179–191. [[CrossRef](#)] [[PubMed](#)]
41. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [[CrossRef](#)] [[PubMed](#)]
42. Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196. [[CrossRef](#)] [[PubMed](#)]
43. Schrödinger, L.L.C. *Glide*; Schrödinger, LLC: New York, NY, USA, 2017.
44. Chemical Computing Group Inc. *Molecular Operating Environment (MOE)*; Chemical Computing Group Inc.: Montreal, QC, Canada, 2017.
45. Corbeil, C.R.; Williams, C.I.; Labute, P. Variability in docking success rates due to dataset preparation. *J. Comput. Aided Mol. Des.* **2012**, *26*, 775–786. [[CrossRef](#)] [[PubMed](#)]

46. Fantini, J.; Barrantes, F. How cholesterol interacts with membrane proteins: An exploration of cholesterol-binding sites including CRAC, CARC, and tilted domains. *Front. Physiol.* **2013**, *4*, 31. [[CrossRef](#)] [[PubMed](#)]
47. Fantini, J.; Di Scala, C.; Evans, L.S.; Williamson, P.T.F.; Barrantes, F.J. A mirror code for protein-cholesterol interactions in the two leaflets of biological membranes. *Sci. Rep.* **2016**, *6*, 21907. [[CrossRef](#)] [[PubMed](#)]
48. Dahl, A.C.; Chavent, M.; Sansom, M.S. Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics* **2012**, *28*, 2193–2194. [[CrossRef](#)] [[PubMed](#)]
49. Studer, G.; Biasini, M.; Schwede, T. Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics* **2014**, *30*, i505–i511. [[CrossRef](#)] [[PubMed](#)]
50. Landau, E.M.; Rosenbusch, J.P. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14532. [[CrossRef](#)] [[PubMed](#)]
51. Kleiger, G.; Grothe, R.; Mallick, P.; Eisenberg, D. GXXXG and AXXXA: Common  $\alpha$ -helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry* **2002**, *41*, 5990–5997. [[CrossRef](#)] [[PubMed](#)]
52. Söding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **2005**, *21*, 951–960. [[CrossRef](#)] [[PubMed](#)]
53. Zimmermann, L.; Stephens, A.; Nam, S.-Z.; Rau, D.; Kübler, J.; Lozajic, M.; Gabler, F.; Söding, J.; Lupas, A.N.; Alva, V. A completely reimplemented mpi bioinformatics toolkit with a new hhpred server at its core. *J. Mol. Biol.* **2017**, *430*, 2237–2243. [[CrossRef](#)] [[PubMed](#)]
54. Webb, B.; Sali, A. Protein Structure Modeling with MODELLER. In *Protein Structure Prediction*; Kihara, D., Ed.; Springer New York: New York, NY, USA, 2014; pp. 1–15.
55. Shen, M.Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2009**, *15*, 2507–2524. [[CrossRef](#)] [[PubMed](#)]
56. Bino, J.; Sali, A. Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res.* **2003**, *31*, 3982–3992. [[CrossRef](#)]
57. Baker, F.N.; Porollo, A. CoeViz: A web-based tool for coevolution analysis of protein residues. *BMC Bioinform.* **2016**, *17*, 119. [[CrossRef](#)] [[PubMed](#)]
58. Larson, S.M.; Di Nardo, A.A.; Davidson, A.R. Analysis of covariation in an SH3 domain sequence alignment: Applications in tertiary contact prediction and the design of compensating hydrophobic core substitutions. *J. Mol. Biol.* **2000**, *303*, 433–446. [[CrossRef](#)] [[PubMed](#)]
59. Lomize, A.L.; Pogozheva, I.D.; Mosberg, H.I. Anisotropic solvent model of the lipid bilayer. 2. energetics of insertion of small molecules, peptides, and proteins in membranes. *J. Chem. Inf. Model.* **2011**, *51*, 930–946. [[CrossRef](#)] [[PubMed](#)]
60. DeLano, W.L. *The PyMOL Molecular Graphics System, Version 1.7.4.5*; Edu Schrödinger, LLC: New York, NY, USA, 2018.
61. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
62. Halgren, T.A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389. [[CrossRef](#)] [[PubMed](#)]
63. Jambeck, J.P.M.; Lyubartsev, A.P. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J. Phys. Chem. B* **2012**, *116*, 3164–3179. [[CrossRef](#)] [[PubMed](#)]
64. Jambeck, J.P.M.; Lyubartsev, A.P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **2012**, *8*, 2938–2948. [[CrossRef](#)] [[PubMed](#)]
65. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
66. Wang, J.M.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)] [[PubMed](#)]
67. Bayly, C.I.; Cieplak, P.; Cornell, W.; Kollman, P.A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The resp model. *J. Phys. Chem.* **1993**, *97*, 10269–10280. [[CrossRef](#)]
68. Wang, J.M.; Cieplak, P.; Kollman, P.A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074. [[CrossRef](#)]

69. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 09, Revision A.02*; Gaussian, Inc.: Wallingford, CT, USA, 2016.
70. Petersson, G.A.; Al-Laham, M.A. A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms. *J. Chem. Phys.* **1991**, *94*, 6081–6090. [[CrossRef](#)]
71. Petersson, G.A.; Bennett, A.; Tensfeldt, T.G.; Al-Laham, M.A.; Shirley, W.A. A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *J. Chem. Phys.* **1988**, *89*, 2193–2218. [[CrossRef](#)]
72. Sousa da Silva, A.W.; Vranken, W.F. ACPYPE-antechamber python parser interface. *BMC Res. Notes* **2012**, *5*, 367. [[CrossRef](#)] [[PubMed](#)]
73. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092. [[CrossRef](#)]
74. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472. [[CrossRef](#)]
75. Hunenberger, P. Thermostat algorithms for molecular dynamics simulations. *Adv. Polym. Sci.* **2005**, *173*, 105–147. [[CrossRef](#)]
76. Parrinello, M.; Rahman, A. Polymorphic transitions in single-crystals—A new molecular-dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190. [[CrossRef](#)]
77. Zimmermann, K. Oral: All purpose molecular mechanics simulator and energy minimizer. *J. Comput. Chem.* **1991**, *12*, 310–319. [[CrossRef](#)]
78. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–18. [[CrossRef](#)] [[PubMed](#)]
79. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [[CrossRef](#)] [[PubMed](#)]
80. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W.F.; Mark, A.E. Peptide folding: When simulation meets experiment. *Angew. Chem. Int.* **1999**, *38*, 236–240. [[CrossRef](#)]
81. Paramo, T.; East, A.; Garzón, D.; Ulmschneider, M.B.; Bond, P.J. Efficient characterization of protein cavities within molecular simulation trajectories: *Trj\_cavity*. *J. Chem. Theory Comput.* **2014**, *10*, 2151–2164. [[CrossRef](#)] [[PubMed](#)]
82. Caliendo, R.; Rossetti, G.; Carloni, P. Local fluctuations and conformational transitions in proteins. *J. Chem. Theory Comput.* **2012**, *8*, 4775–4785. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).