Article

CGMD Platform: Integrated Web Servers for the Preparation, Running, and Analysis of Coarse-Grained Molecular Dynamics Simulations

Alessandro Marchetto [1,2,3,†], **Zeineb Si Chaib [2,3,†]**, **Carlo Alberto Rossi [1,†]**, **Rui Ribeiro [1]**, **Sergio Pantano [4]**, **Giulia Rossetti [2,5,6,*]** and **Alejandro Giorgetti [1,2,*]**

[1]   Department of Biotechnology, University of Verona, 37134 Verona, Italy; a.marchetto@fz-juelich.de (A.M.); carloalberto.rossi@studenti.univr.it (C.A.R.); ruipedro.fernandesribeiro@univr.it (R.R.)

[2]   Computational Biomedicine, Institute for Neuroscience and Medicine (INM-9) and Institute for Advanced Simulations (IAS-5), Forschungszentrum Jülich, 52425 Jülich, Germany; z.si.chaib@fz-juelich.de

[3]   Faculty of Mathematics, Computer Science and Natural Sciences, RWTH Aachen University, 52062 Aachen, Germany

[4]   Institut Pasteur de Montevideo, Mataojo 2020, 11400 Montevideo, Uruguay; spantano@pasteur.edu.uy

[5]   Jülich Supercomputing Center (JSC), Forschungszentrum Jülich, 52428 Jülich, Germany

[6]   Department of Hematology, Oncology, Hemostaseology, and Stem Cell Transplantation, RWTH Aachen University, 52062 Aachen, Germany

*   Correspondence: g.rossetti@fz-juelich.de (G.R.); alejandro.giorgetti@univr.it (A.G.)

†   These authors contributed equally to this work.

check for updates

Abstract: Advances in coarse-grained molecular dynamics (CGMD) simulations have extended the use of computational studies on biological macromolecules and their complexes, as well as the interactions of membrane protein and lipid complexes at a reduced level of representation, allowing longer and larger molecular dynamics simulations. Here, we present a computational platform dedicated to the preparation, running, and analysis of CGMD simulations. The platform is built on a completely revisited version of our *Martini coarsE gRained MembrAne proteIn Dynamics* (MERMAID) web server, and it integrates this with other three dedicated services. In its current version, the platform expands the existing implementation of the Martini force field for membrane proteins to also allow the simulation of soluble proteins using the Martini and the SIRAH force fields. Moreover, it offers an automated protocol for carrying out the backmapping of the coarse-grained description of the system into an atomistic one.

Keywords: CGMD; SIRAH; martini force field; backmapping

1. Introduction

Cellular processes lean on biological macromolecules and on interactions among them to serve as building blocks for large-scale functional complexes. These often contain many copies of the same or different biomolecules that aggregate through long-range interactions into functional suprastructures. These macromolecular complexes perform their function either in a soluble environment or embedded in the cell membrane. Furthermore, the way in which membrane proteins and their macromolecular complexes associate with and within the lipid bilayer may have an effect on the function of the protein itself. Thus, the availability of techniques that allow a systematic study of these lipids/membrane protein systems spanning large time- and space-scales is fundamental for the understanding of their functions.

Molecular dynamics (MD) simulations have emerged as a powerful tool to study biological systems at varying lengths and timescales. Indeed, all-atom (AA) molecular dynamics simulations

are being used by a wide scientific community as a routine to capture biomolecules' conformational dynamics and local motions. In addition, recent developments in algorithms that reduce the MD sampling [1–4], and in particular the new advancements in coarse-grained (CG) models, have paved the way to studying the dynamics of macromolecular complexes for timescales up to milliseconds [5]. CG models are a reduced representation of all-atom models that keep the crucial molecular features of the system. This reduced representation of atoms allows not only the simulation of large-scale biological systems, but they also enable faster sampling due to reduced degrees of freedom. As a consequence, CG simulations can cover longer timescales, and therefore provide access to the in-depth knowledge of dynamics of macromolecules, which is beyond the scope of AA simulations [6,7]. Indeed, CG models have been extensively used for studying the protein folding mechanism, structure prediction, protein-membrane systems, and aggregation [8]. For these reasons, the CG protein model development is constantly evolving (see for a review [9]), and their applications are widely used: plenty of studies that have exploited CG force fields, like Martini [10], SIRAH [11], OPEP [12], Shinoda-Devane-Klein (SDK) [13] and ELBA FFs [14], have been published [15], (for a recent review, see [8]).

In 2019, we presented the Martini CoarsE-gRained MembrAne proteIn Dynamics (MERMAID) web server [16] aimed at the preparation, running, and analysis of CG molecular dynamics (CGMD) simulations using the Martini force field tailored to membrane proteins. In this paper, we present a renewed and extended version of our previous server. MERMAID has now evolved into a comprehensive CGMD platform, dedicated to the preparation, running, and analysis of CGMD simulations using the Martini force field version 2.2 for both membrane and soluble proteins. In addition, our platform implements another CGMD force field, namely SIRAH, in an aqueous environment. Finally, we added a backmapping protocol, allowing the users to upload CG Martini structures to be backmapped in the appropriate standalone web server section.

The input can be either an X-ray, NMR, or cryo-EM structure of a protein or models generated by different methods/web servers (for a review on membrane modeling servers, see [17]). In particular, when dealing with models of G-protein coupled receptors (GPCRs) we offer a direct link from our *GOMoDo* web server [18]. The platform funnels the users through several intuitive forms, guiding them, step-by-step, towards the preparation, running, and analysis of a protein CG system. The CGMD Platform implements not only a completely different and more user-friendly interface, but it also includes several novel features: (i) external web servers are no longer required for adding missing atoms within the structure; (ii) a new set of parametrized lipids, as well as different physiological membranes with pre-assembled compositions, are now available. In summary, in this new platform, the users gain the possibility of having a better follow-up of each process due to a constant iterative-mode throughout all the different stages of preparation and running of CGMD simulations. The 'CGMD Platform' is freely accessible at https://molsim.sci.univr.it/mermaid/begin.php.

2. Results and Discussion

In this section, we describe the different features of the platform, as well as three application cases.

2.1. CGMD Platform Architecture

The platform is organized as four different stand-alone web servers. These include: (i) the new version of MERMAID; (ii) the 'water Martini' web server dedicated to the simulation of soluble proteins using the Martini force field; (iii) a completely new web server for the preparation and running of CGMD simulations using the SIRAH force field in explicit solvent; (iv) a server dedicated to the backmapping of Martini CG representation of the protein and/or some lipids to an atomistic-level description of the system.

The workflows for the different web servers can be appreciated in Figures 1–3. Each of them can be divided into two different stages, which include the user interface (front-end) and the data retrieval, as well as the back end of the server.

2.1.1. MERMAID and Water Martini Web Servers

MERMAID and Water Martini Web Server workflows are shown in Figure 1.

- **Front-End**

  - *File upload*. The users can interact with a renewed web client interface to submit a protein structure. In particular, either a custom PDB structure or a structure automatically downloaded from the OPM Server (https://opm.phar.umich.edu/) [19] (for membrane proteins), or RCSB PDB for soluble proteins, can be submitted (Figure 1). The custom PDB file can be either an experimental or a modeled structure. If the experimental structure is an NMR-derived ensemble, the first conformer is automatically chosen using an in-house script. In the case of GPCRs, the user can employ any modeling program or server (for a review, see [17]). Alternatively, we offer a direct link from our *GOMoDo* web server [18] for the modeling of GPCRs. The models generated using methods other than *GOMoDo* have to be previously aligned along the *Z*-axis. This can be done manually or by using web servers like the PPM server (https://opm.phar.umich.edu/ppm_server). If the submitted structure contains missing atoms, they will be automatically added using an in-house script that implements the complete_pdb function of Modeller 9.25 [20]. At this point, the user is asked to register in the Modeller webpage (https://salilab.org/modeller/) and to add the corresponding Modeller license key in the appropriate field;

  - *Interactive preparation.* This interface allows the users to choose all the parameters for the simulation. The server suggests some default values, but expert users have the freedom to change them according to their needs or directly upload precompiled parameter files. Two versions of the Martini force field are offered, i.e., Martini22 and Elnedyn22. The users are then funneled through different panels. Each panel allows the choice of a large variety of parameters, including those related to the *martinize.py* [21] (Version 2.6) (charges of termini and chain breaks, disulfide bridges, and position restraints) and *insane.py* [22] (PBC type, distance between periodic images, box dimensions, lipid type and abundance in both upper and lower leaflets, etc.) python scripts. Counterions can also be chosen (Na^+, Ca^{2+}, and choline ($NC3^+$) as positive ions, Cl^- as negative ion) to neutralize the charge of the simulated system. Moreover, expert users can directly upload their own mdp files with custom-made parameters;

  - *Membrane selection*. The users can choose from among 57 already parametrized lipids to generate a custom-made membrane composition. On the other hand, the users can select one among several physiological-like membranes, where the known lipidic composition is already set up. The composition of the default membrane offered is the Golgi apparatus membrane [23]. Some of the offered membranes include: (a) Golgi membrane with a composition of: CHOL:18%, POPC:36%, POPE:21%, POSM:6%, POPS:6% and POPI:12%; (b) endoplasmic reticulum membrane: CHOL:8%, POPC:54%, POPE:20% and POPI:11%; (c) plasma membrane: CHOL:34%, POPC:23%, POPE:11%, POSM:17% and POPS:8% and (d) mitochondrial membrane: CDL:22%, POPC:37%, POPE:31% and POPI:6% [23]. Moreover, the user can "model" a customized membrane, varying the concentration and type of lipids. The choice of membrane composition can be done either for both leaflets of the membrane or considering the inner and external leaflets independently. The users have access to all the generated input and output files for each CGMD process at any time. Data can be accessed either from the bookmarked web link or directly from the server user directory by providing the required credentials in the MERMAID search bar, as indicated in the 'on-the-fly' documentation;

  - *Output and data retrieval*. Results can be viewed and downloaded for 2 weeks (renewal possible) by bookmarking the link, or alternatively by using the corresponding IDs. The full output of the preparation can be downloaded as a compressed archive file including the input,

output, and log files of all preparation and simulation steps. The downloaded files can be used to continue the CGMD simulations locally. Experienced users have the possibility of downloading the prepared system and tuning the parameters before running the simulation on their local computer. An array of trajectory analyses is available. These include the calculation of the Root Mean Square Deviation (RMSD), density, pressure, temperature, and gyration radius, among others. The corresponding plots are also visualized. This new version of MERMAID also allows for displaying the simulation run directly on the browser at three different speeds using the NGL Viewer [24].

- **Back-End**. After submitting all files and parameters, the web server creates a local user directory where all operations will be performed. During the initial setup of the system, the atomistic protein structure is converted into a CG representation with the help of the *martinize.py* python script [21]. In the case of membrane proteins, the CG structure is then embedded into a user-defined CG lipid membrane with the help of the python script *insane.py* [22]. Subsequently, the CG simulations are run within our servers. A typical CGMD protocol consists of four phases:

  ○ Minimization run;
  ○ Equilibration runs in two different ensembles, namely canonical ensemble (NVT) and isobaric–isothermal ensemble (NPT);
  ○ Production run continued in an NPT ensemble;
  ○ Analysis of all the trajectories produced during the simulation.

The suggested default parameters are the ones recommended by Martini developers, based on extensive testing [25]: for example, it is advisable to treat coulomb interactions using a reaction-field, as this gives slightly better results at a negligible extra computational cost. A straight cutoff can be used in Martini simulations, with a cutoff distance of 1.1 nm. Good temperature control can be achieved with the velocity rescale (V-rescale) thermostat, using a coupling time constant of at least 0.5 ps. For bilayer systems, the pressure coupling should be semi-isotropic. However, not all systems have been tested and it is recommended that the users perform their own tests.
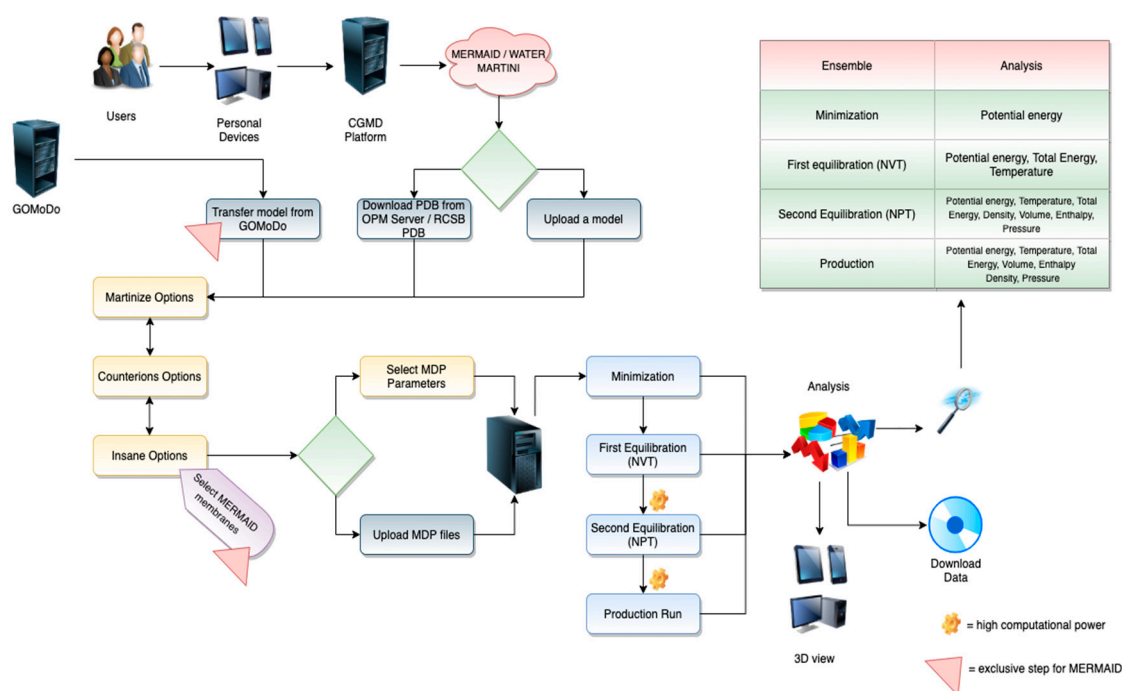

Figure 1. The Martini CoarsE-gRained MembrAne proteIn Dynamics (MERMAID) and Water Martini workflow.

2.1.2. SIRAH Web Server

SIRAH Web Server workflow is shown in Figure 2.

- **Front-End.** With this web server, the users have the possibility of preparing, running, and analyzing CG simulations using the SIRAH 2.2 force field [26] (Figure 2). This feature allows for using a completely independent approach for running CGMD simulations.

  ○ *File upload.* The users must provide a PDB file containing the correct protonation state for each residue to map at a proper pH the SIRAH CG beads. Since hydrogen nomenclature is not entirely standardized across different software packages, we strongly suggest the use of the pdb2pqr web server [27] (http://server.poissonboltzmann.org/);

  ○ Notice that, if the system contains disulfide bridges, the cysteine residue names involved in disulfide bridges have to be manually edited from "CYS" to "CYX" in the PDB file provided as input. Similarly, to simulate a cysteine in thiolate state, the residue name must be changed from "CYS" to "CYM". Protonation states of aspartates and glutamates can be set to neutral by editing the residue name to ASH or GLH, respectively;

  ○ *Simulation Parameters.* The server shows the precompiled MD parameters (mdp) files for the simulation. In this case, the parameters are fixed and visible in read-only mode;

  ○ *Running.* After the submission of the file, the user's job is queued and its status can be monitor at any time during the simulation;

  ○ *Output and data retrieval.* As for the MERMAID web server, the results are stored for 2 weeks. The full output of the preparation can be downloaded as a compressed archive file including the input, output, topologies, and log files of all the preparation and simulation steps. The downloaded files can be used to continue the CGMD simulations locally. The offered analysis tools are the same as for the water Martini case. It can also display the simulation run directly on the browser at three different speeds (1×, 2×, and 5×) using NGL Viewer [24].

- **Back-end.** The protonated model is converted to CG using SIRAH Tools [28] and solvated using pre-stabilized boxes of the WatFour (WT4) CG water model and electrolytes [29]. Each run can be divided into two: solvation and addition of counterions, and the five molecular dynamics steps (two minimization steps, two equilibration steps, and one production step). The mdp files are displayed during the preparation. These parameters were extensively tested and should allow a smooth and fully automated preparation of the system in explicit solvent.
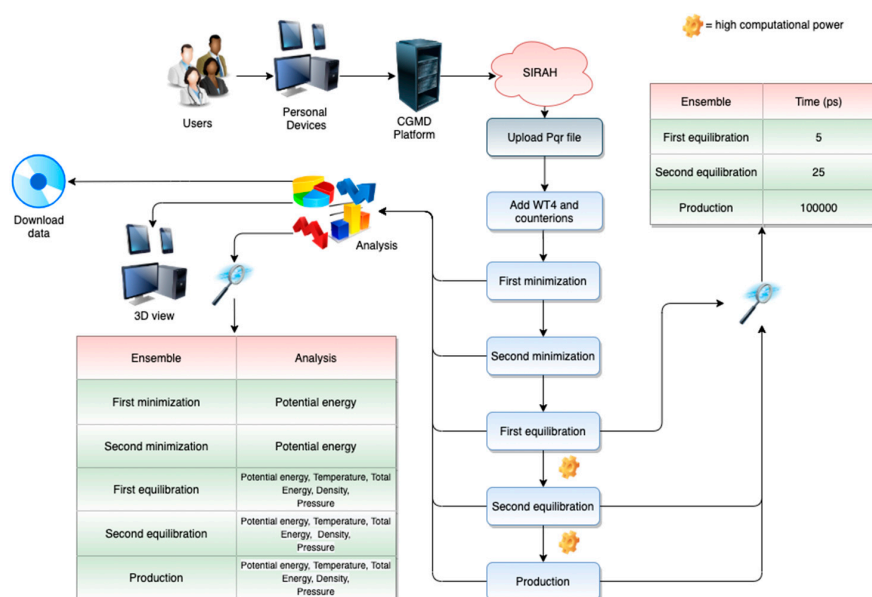

Figure 2. The SIRAH web server workflow.

2.1.3. Backmapping Web Server

The implicit loss of resolution of CG representations is a limiting factor when trying to interpret the details of the simulations. Indeed, atomistic level details, such as specific contacts, are the key for understanding molecular recognition and specific intra/intermolecular details. The process of retrieving atomistic details from a CG representation is known as reverse transformation, inverse mapping, or backmapping. There are several different backmapping protocols that follow two different steps, i.e., generation of an atomistic structure based on the CG coordinates, and relaxation step of the generated atomistic structure [30]. Here, we implement the backward program [31] (Figure 3).

- **Front-End**. The backmapping procedure can be reached from an independent menu. The users can backmap a protein in water from the Martini force field to Amber [32], Charmm36 [33] or Gromos [34] force fields. For membrane systems, the following lipids are supported: CHOL, DOPC, DOPE, DOPG, DOPS, DPPC, POPC, POPE, and POPG. Slipids' force field [35] topologies are used, and consequently the associated protein is backmapped to the amber force field [32];
- **Back-end.** MERMAID backmapping allows reconstructing the protein from a CG to AA representation using the backward program [31]. The latter consists of three scripts and a number of CG to atomistic mapping definition files. For a description of the backmapping procedure, see [31].
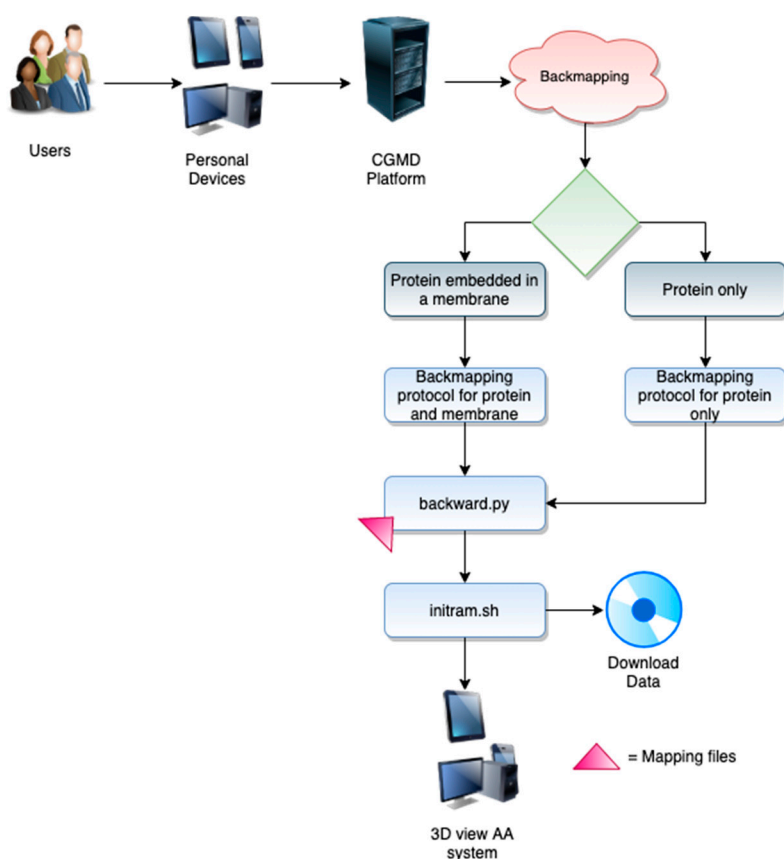

Figure 3. The backmapping workflow.

2.2. Documentation

Extensive documentation for each of the steps and parameters is offered through the use of dynamic pop-ups generated using the Bootstrap Tour tools (https://bootstraptour.com/) locally installed within our machine. The availability of a dynamic, always-accessible guide through the entire process

allows the users to be documented "on the fly", avoiding the opening of new pages or movement along long manual pages.

The tutorial with examples (see Section 2.3) offers pre-calculated systems that can be easily accessed by the user at any time. These cases show all the possible calculations that can be performed through the CGMD platform.

2.3. Application Cases

2.3.1. Application Case: The Martini Force Field

As an application case, we present the preparation and running of the *Rhodobacter sphaeroides* translocator protein (RsTSPO), also known as the peripheral benzodiazepine receptor (PDB accession code: 4UC3). TSPO is an 18 KDa membrane protein conserved across the three domains of life [36] with a vast spectrum of roles ranging from an environmental sensor to a functional bioregulator of apoptosis, autophagy, inflammation, along with cholesterol and porphyrin transport [37]. Recently, we presented a CGMD study of this receptor aimed at the characterization of the impact of cholesterol in the formation of multimeric assemblies [38]. In this paper, we claimed that cholesterol affects the size and rigidity of the bacterial translocator proteins to a lesser extent than the mammalian protein. Moreover, an overabundance of cholesterol causes a decrease in the number of contacts at the subunit–subunit interface of the RsTSPO and *Bacillus cereus* TSPO (BcTSPO) systems. For this reason, the role of the sterol could be potentially significant, although the study of these bacterial proteins is not conclusive. Aiming to illustrate the capability of this server, we offer the users the possibility of preparing, running, and analyzing the bacterial TSPO simulation. This application case has also been run, and it is freely accessible here: https://molsim.sci.univr.it/mermaid/public_html/membrane/mrstspo1/run/analysis.php. This system was prepared following the workflow presented in Figure 1 and simulated in a 70% POPG and 30% CHOL model membrane (offered by our CGMD Platform). A total of 305 Na$^+$ and 307 Cl$^-$ was added to neutralize the system. A 1.5 nm distance between periodic images was applied. The simulation was conducted in four steps (minimization, NVT, NPT, and production) using default parameters. After ~20 ns, the structure was equilibrated, as shown from the RMSD plot. The radius of gyration plot reflects how the protein remained compact during the simulation. Additional plots provided on the analysis page offer tools to assess the behavior of the simulation, which were stable in all the ensembles. Indeed, taking all the analyses together, the user has a system that is prepared to be run in a similar fashion to the one in [38]. During and after the simulation, the users are constantly informed of the status of their calculation through a progress bar positioned at the top of each analysis page, which is always accessible. The backmapping procedure was used at the end of the MD following the steps depicted in the workflow in Figure 3. The final *backmapped.gro* structure is freely accessible in our public Github Repository: https://github.com/JavaScript92/CGMD_Platform. A "Demo Case" button is available from the preparation page to allow an automatic upload of the PDB and running this application case. The reference input file can be downloaded from the same repository.

2.3.2. Application Case: The Water Martini

As an application case, we present the preparation and running of the structure of the orthorhombic form of the hen egg-white lysozyme at 1.5 Å resolution (PDB accession code: 1AKI), which is accessible here: https://molsim.sci.univr.it/mermaid/public_html/water/wusecase/run/analysis.php. This system was prepared following the workflow in Figure 1 and simulated in a 15 nm^3 cubic box. 305 Na$^+$ and 313 Cl$^-$ were added to neutralize the system. A 1.5 nm distance between periodic images was applied. The simulation was conducted in four steps (minimization, NVT, NPT, and production) using default parameters. After ~2 ns, the structure was equilibrated, as shown from the RMSD plot. Besides, the expected trend in the radius of gyration plot reflects how the protein remained compact during the simulation. The charts within the analysis page allowed us to assess the stability and convergence of the first 100 ns of simulation. These results indicate that the simulation is ready to begin a production

run. For the sake of completeness, to provide an estimation concerning the time gained between an AA and Martini CG simulation carried out in the same system (1AKI) in a water environment, a CG representation, consisting of 29,726 beads, is ~100 times faster than an AA simulation containing 90,273 atoms. Both AA and CG simulations were simulated using the same computational resources (one node and 14 OpenMP threads).

2.3.3. Application Case: The SIRAH Force Field

As an application case of the SIRAH force field [11], we present and run the Binary Complex of Restriction Endonuclease HinP1I with its Cognate DNA (PDB accession code: 2FL3), following the workflow in Figure 2. The results for this demo case can be freely accessed here: https://molsim.sci.univr.it/mermaid/public_html/sirah/s2FL3demo/run/analysis.php. This system was prepared as described in Section 2.1.2 and simulated in explicit solvent. The octahedron box was neutralized adding CG K^+ and Cl^- ions, using the *gmx genion* tool, providing a physiologic ionic strength of 0.15 M. The simulation was performed in five steps: two minimization runs (5000 steps each), two equilibration runs (5 and 25 ps, respectively), and one production run (100 ns). After and during the MD simulation, the users are constantly informed on the status of their simulation through the progress bar positioned at the top of each analysis page, which is always accessible during the running. Within our CGMD Platform, several analyses can be carried out, including 20 dynamic plots (temperature, pressure, potential energy, etc.), a 3D visualization for each MD step, and an "on-the-fly" 3D visualization of the production ensemble. The documentation explains how to retrieve the corresponding jobs. At the end of the simulation, a zip archive containing all the necessary molecular structures (in gro format), topologies, trajectories, and calculated properties (in xvg format) can be downloaded from the analysis page. Finally, the users can run this example by uploading the file 2fl3.pqr available here: https://github.com/JavaScript92/CGMD_Platform.

2.4. Usage Statistics

The previous version of the MERMAID web server has been extensively used by different research groups around the world. As of today (31 October 2020), over 100 calculations were performed on our web server, including Covid-19 proteins, ATPases, and several other membrane receptors. For the current version of the CGMD Platform, our group has run several test cases for each of them. For the sake of comprehension, some application cases are reported here: https://molsim.sci.univr.it/mermaid/page/applicationCases.php. Other test files are freely accessible from our Github repository: https://github.com/JavaScript92/CGMD_Platform.

3. Materials and Methods

The CGMD Platform web-interface is equipped with HTML 5 (https://www.w3.org/html/), PHP version 7.4.3 (http://www.php.net/), JavaScript (https://www.javascript.com/ and https://www.highcharts.com/), Bash shell (https://www.gnu.org/software/bash/), jQuery 3.4.1 (https://jquery.com/), and Bootstrap 4.4.1 (https://getbootstrap.com/). For the submitted jobs, the web server prepares the input files on our local cluster. All simulations were run in a Linux Ubuntu Server 20.04.1 LTS machine with 12 cores and 32GB of RAM. Results were post-processed (which includes analysis of CGMD trajectory, construction of various data plots, visualization of the production ensemble trajectories) and displayed graphically by using Plotly.js (https://plot.ly).

The current version of CGMD Platform implements: (i) a local version of Gromacs 2019.3 [39] to perform MD simulations and carry out the trajectories analysis as well as the backmapping procedure; (ii) a locally installed version of the Dictionary of Protein Secondary Structure (DSSP) [40] to get the geometrical properties of the secondary structure required for running the Martini force field version 2.2; (iii) all the programs needed for preparing the files for the Martini force field, including the *martinize.py* script [21] for coarse-graining the protein structures, and the *insane.py* script [22] for embedding the protein in the membrane; (iv) the version 2.2 of the SIRAH force field along with

SirahTools [28] to prepare a CG version of the solute (downloaded from http://www.sirahff.com); (v) Plotly (http://gnuplot.info/) for data plotting; (vi) NGL viewer [24] to provide an interactive 3D molecular viewer of the molecular dynamics simulation embedded in the web page. Moreover, several Gromacs packages were employed: *gmx energy* tool was used to generate the temperature, pressure, total and potential energy, volume, and enthalpy plots; *gmx density* tool was used to calculate the density of all components of the system; *gmx trjconv* tool was used to convert the trajectories to gro format and remove water as well as lipids from the structures displayed on the browser; *gmx rms* and *gmx gyrate* were applied for the calculation of the RMSD of the protein (the same group was used for the least-squares fit, using the output structure of the NPT equilibration as a reference) and radius of gyration of the protein, respectively.

Finally, since GPCRs share the same topology, the modelled proteins coming from *GOMoDo* are structurally aligned against the Nociceptin receptor in its inactive state, which has been used as a reference structure (RCSB PDB Code: 4EA3, downloaded from the OPM Server), using a locally installed version of LovoAlign [41] to help to obtain a proper orientation in the membrane.

4. Conclusions

Understanding the mechanisms and dynamics underlying the function of big macromolecular complexes and membrane proteins embedded in their physiological lipidic environments has always been a challenge for researchers, as the time and space scales needed to gain a comprehensive understanding are very difficult to uncover.

CGMD techniques have been shown to be useful for gaining deeper insights into the mechanisms underlying the function of these systems at time/space-scales that cannot be studied at an atomistic resolution. Due to increasing interest and the availability of more powerful algorithms and hardware, in recent years, several web servers have been developed with the scope to prepare and/or perform different resolution levels of MD simulations, i.e., CHARMM-GUI [42], CABS-flex [43], locPREFMD [44], MDWeb [45], PREFMD [46], ProBLM [47], SMOG [48], UNRES [49], Vienna-PTM [50], MM/CG web server [51]. In particular, the existing web servers provide a web-based graphical user interface to generate various molecular simulation systems and input files to facilitate the usage of common simulation techniques. They are widely used to prepare the input files needed to be run on local clusters, and, in some of them, to run short simulations. Nevertheless, they are not specialized in membrane proteins and they allow neither the selection of different membranes nor the running of CGMD simulations. Our newly developed computational platform implements, together with the previously described features, several novel elements to aid in the preparation and running of CG simulations, such as the availability of a backmapping service, and the possibility of choosing between the Martini or the SIRAH CG force field for carrying out the system preparation. Moreover, the user is constantly guided through the entire procedure by a series of dynamic help pop-ups on each of the pages. This allows more direct and self-aware interaction with the server. Besides, the system offers a set of analyses that can be carried out on the fly to follow the evolution of fundamental features, such as temperature, pressure, and different energy terms during the simulation. Last but not least, it offers the possibility of choosing from among 57 different lipids for custom membrane building, or the availability of ready-to-use realistic models for the most common types of physiological membranes.

Membrane composition and membrane protein oligomerization play a biologically relevant role in cell function. The new platform could prospectively be used for efficiently setting up CGMD simulations of the same protein embedded in different lipidic environments, as well as comparative studies of membrane proteins in different multimeric states, by capitalizing on the low computational cost of CGMD.

To summarize, our novel CGMD Platform offers a user-friendly interface available to all users without any login requirement. The CGMD Platform is the only on-line protocol specifically designed for preparing and running CGMD simulations online: it guides the user step-by-step, therefore enormously reducing the usually time-consuming step of the system-set-up, even for non-expert users. It supports

the preparation of complex systems, like membrane protein/lipids, easing the laborious process of membrane assemblies but still allowing them to be completely customizable. In addition, it offers a wide spectrum of highly specialized tools for simulation analyses.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Evangelista Falcon, W.; Ellingson, S.R.; Smith, J.C.; Baudry, J. Ensemble docking in drug discovery: How many protein configurations from molecular dynamics simulations are needed to reproduce known ligand binding? *J. Phys. Chem. B* **2019**, *123*, 5189–5195. [CrossRef]
2. Phillips, J.C.; Hardy, D.J.; Maia, J.D.; Stone, J.E.; Ribeiro, J.V.; Bernardi, R.C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130. [CrossRef] [PubMed]
3. Gonçalves, M.A.; Santos, L.S.; Prata, D.M.; Peixoto, F.C.; da Cunha, E.F.; Ramalho, T.C. Optimal wavelet signal compression as an efficient alternative to investigate molecular dynamics simulations: Application to thermal and solvent effects of MRI probes. *Theor. Chem. Acc.* **2017**, *136*, 15. [CrossRef]
4. Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159–184. [CrossRef] [PubMed]
5. Joshi, S.Y.; Deshmukh, S.A. A review of advancements in coarse-grained molecular dynamics simulations. *Mol. Simul.* **2020**, 1–18. [CrossRef]
6. Pluhackova, K.; Wassenaar, T.; Böckmann, R. Membrane Biogenesis. In *Methods in Molecular Biology*; Rapaport, D., Herrmann, J.M., Eds.; Springer: Berlin, Germany, 2013; pp. 85–101.
7. Harpole, T.J.; Delemotte, L. Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations. *Biochim. Et Biophys. Acta (Bba)-Biomembr.* **2018**, *1860*, 909–926. [CrossRef] [PubMed]
8. Singh, N.; Li, W. Recent advances in coarse-grained models for biomolecules and their applications. *Int. J. Mol. Sci.* **2019**, *20*, 3774. [CrossRef] [PubMed]
9. Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A.E.; Kolinski, A. Coarse-grained protein models and their applications. *Chem. Rev.* **2016**, *116*, 7898–7936. [CrossRef]
10. Marrink, S.J.; Risselada, H.J.; Yefimov, S.; Tieleman, D.P.; De Vries, A.H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824. [CrossRef]
11. Darré, L.; Machado, M.R.; Brandner, A.F.; González, H.C.; Ferreira, S.n.; Pantano, S. SIRAH: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *J. Chem. Theory Comput.* **2015**, *11*, 723–739. [CrossRef]
12. Barroso da Silva, F.L.s.; Sterpone, F.; Derreumaux, P. OPEP6: A new constant-pH molecular dynamics simulation scheme with OPEP coarse-grained force field. *J. Chem. Theory Comput.* **2019**, *15*, 3875–3888. [CrossRef] [PubMed]
13. Shinoda, W.; DeVane, R.; Klein, M.L. Zwitterionic lipid assemblies: Molecular dynamics studies of monolayers, bilayers, and vesicles using a new coarse grain force field. *J. Phys. Chem. B* **2010**, *114*, 6836–6849. [CrossRef] [PubMed]

14. Orsi, M.; Essex, J.W. The ELBA force field for coarse-grain modeling of lipid membranes. *PLoS ONE* **2011**, *6*, e28637. [CrossRef] [PubMed]

15. Takahashi, K.; Oda, T.; Naruse, K. Coarse-grained molecular dynamics simulations of biomolecules. *Aims Biophys.* **2014**, *1*, 1–15. [CrossRef]

16. Damre, M.; Marchetto, A.; Giorgetti, A. MERMAID: Dedicated web server to prepare and run coarse-grained membrane protein dynamics. *Nucleic Acids Res.* **2019**, *47*, W456–W461. [CrossRef] [PubMed]

17. Busato, M.; Giorgetti, A. Structural modeling of G-protein coupled receptors: An overview on automatic web-servers. *Int. J. Biochem. Cell Biol.* **2016**, *77*, 264–274. [CrossRef]

18. Sandal, M.; Duy, T.P.; Cona, M.; Zung, H.; Carloni, P.; Musiani, F.; Giorgetti, A. GOMoDo: A GPCRs online modeling and docking webserver. *PLoS ONE* **2013**, *8*, e74092. [CrossRef]

19. Lomize, M.A.; Lomize, A.L.; Pogozheva, I.D.; Mosberg, H.I. OPM: Orientations of proteins in membranes database. *Bioinformatics* **2006**, *22*, 623–625. [CrossRef]

20. Šali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [CrossRef]

21. de Jong, D.H.; Singh, G.; Bennett, W.D.; Arnarez, C.; Wassenaar, T.A.; Schäfer, L.V.; Periole, X.; Tieleman, D.P.; Marrink, S.J. Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **2013**, *9*, 687–697. [CrossRef]

22. Wassenaar, T.A.; Ingólfsson, H.I.; Böckmann, R.A.; Tieleman, D.P.; Marrink, S.J. Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144–2155. [CrossRef]

23. Casares, D.; Escribá, P.V.; Rosselló, C.A. Membrane lipid composition: Effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. *Int. J. Mol. Sci.* **2019**, *20*, 2167. [CrossRef] [PubMed]

24. Rose, A.S.; Hildebrand, P.W. NGL Viewer: A web application for molecular visualization. *Nucleic Acids Res.* **2015**, *43*, W576–W579. [CrossRef] [PubMed]

25. De Jong, D.H.; Baoukina, S.; Ingólfsson, H.I.; Marrink, S.J. Martini straight: Boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **2016**, *199*, 1–7. [CrossRef]

26. Machado, M.R.; Barrera, E.E.; Klein, F.; Sónora, M.; Silva, S.; Pantano, S. The SIRAH 2.0 Force field: Altius, fortius, citius. *J. Chem. Theory Comput.* **2019**, *15*, 2719–2733. [CrossRef]

27. Dolinsky, T.J.; Nielsen, J.E.; McCammon, J.A.; Baker, N.A. PDB2PQR: An automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*, W665–W667. [CrossRef]

28. Machado, M.R.; Pantano, S. SIRAH tools: Mapping, backmapping and visualization of coarse-grained models. *Bioinformatics* **2016**, *32*, 1568–1570. [CrossRef]

29. Darré, L.; Machado, M.R.; Dans, P.D.; Herrera, F.E.; Pantano, S. Another coarse grain model for aqueous solvation: WAT FOUR? *J. Chem. Theory Comput.* **2010**, *6*, 3793–3807. [CrossRef]

30. Roel-Touris, J.; Bonvin, A.M. Coarse-grained (hybrid) integrative modeling of biomolecular interactions. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1182–1190. [CrossRef]

31. Wassenaar, T.A.; Pluhackova, K.; Böckmann, R.A.; Marrink, S.J.; Tieleman, D.P. Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models. *J. Chem. Theory Comput.* **2014**, *10*, 676–690. [CrossRef]

32. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct. Funct. Bioinform.* **2006**, *65*, 712–725. [CrossRef]

33. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.; Mittal, J.; Feig, M.; MacKerell, A.D., Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ1 and χ2 dihedral angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273. [CrossRef] [PubMed]

34. van Gunsteren, W.F.; Billeter, S.; Eising, A.; Hünenberger, P.; Krüger, P.; Mark, A.; Scott, W.; Tironi, I. Biomolecular simulation: The GROMOS96 manual and user guide. *Vdf Hochschulverlag AG Der ETH ZürichZürich* **1996**, *86*, 1–1044.

35. Jämbeck, J.P.; 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.

36. Bonsack, F.; Sukumari-Ramesh, S. TSPO: An evolutionarily conserved protein with elusive functions. *Int. J. Mol. Sci.* **2018**, *19*, 1694. [CrossRef] [PubMed]

37. Fan, J.; Lindemann, P.; GJ Feuilloley, M.; Papadopoulos, V. Structural and functional evolution of the translocator protein (18 kDa). *Curr. Mol. Med.* **2012**, *12*, 369–386. [PubMed]

38. Si Chaib, Z.; Marchetto, A.; Dishnica, K.; Carloni, P.; Giorgetti, A.; Rossetti, G. Impact of Cholesterol on the Stability of Monomeric and Dimeric Forms of the Translocator Protein TSPO: A Molecular Simulation Study. *Molecules* **2020**, *25*, 4299. [CrossRef]

39. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [CrossRef]

40. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolym. Orig. Res. Biomol.* **1983**, *22*, 2577–2637. [CrossRef]

41. Martínez, L.; Andreani, R.; Martínez, J.M. Convergent algorithms for protein structural alignment. *BMC Bioinform.* **2007**, *8*, 1–15. [CrossRef]

42. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [CrossRef] [PubMed]

43. Kuriata, A.; Gierut, A.M.; Oleniecki, T.; Ciemny, M.P.; Kolinski, A.; Kurcinski, M.; Kmiecik, S. CABS-flex 2.0: A web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res.* **2018**, *46*, W338–W343. [CrossRef] [PubMed]

44. Feig, M. Local protein structure refinement via molecular dynamics simulations with locPREFMD. *J. Chem. Inf. Model.* **2016**, *56*, 1304–1312. [CrossRef]

45. Hospital, A.; Andrio, P.; Fenollosa, C.; Cicin-Sain, D.; Orozco, M.; Gelpí, J.L. MDWeb and MDMoby: An integrated web-based platform for molecular dynamics simulations. *Bioinformatics* **2012**, *28*, 1278–1279. [CrossRef] [PubMed]

46. Heo, L.; Feig, M. PREFMD: A web server for protein structure refinement via molecular dynamics simulations. *Bioinformatics* **2018**, *34*, 1063–1065. [CrossRef] [PubMed]

47. Kimmett, T.; Smith, N.; Witham, S.; Petukh, M.; Sarkar, S.; Alexov, E. ProBLM web server: Protein and membrane placement and orientation package. *Comput. Math. Methods Med.* **2014**, *2014*, 838259. [CrossRef]

48. Noel, J.K.; Levi, M.; Raghunathan, M.; Lammert, H.; Hayes, R.L.; Onuchic, J.N.; Whitford, P.C. SMOG 2: A versatile software package for generating structure-based models. *PLoS Comput. Biol.* **2016**, *12*, e1004794. [CrossRef]

49. Czaplewski, C.; Karczyńska, A.; Sieradzan, A.K.; Liwo, A. UNRES server for physics-based coarse-grained simulations and prediction of protein structure, dynamics and thermodynamics. *Nucleic Acids Res.* **2018**, *46*, W304–W309. [CrossRef]

50. Margreitter, C.; Petrov, D.; Zagrovic, B. Vienna-PTM web server: A toolkit for MD simulations of protein post-translational modifications. *Nucleic Acids Res.* **2013**, *41*, W422–W426. [CrossRef]

51. Schneider, J.; Ribeiro, R.; Alfonso-Prieto, M.; Carloni, P.; Giorgetti, A. Hybrid MM/CG Webserver: Automatic set up of molecular mechanics/coarse-grained simulations for human G protein-coupled receptor/ligand complexes. *Front. Mol. Biosci.* **2020**, *7*, 232. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.