Chemico-Biological Interactions 407 (2025) 111372

Contents lists available at ScienceDirect

Chemico-Biological Interactions

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

Research paper

Conformation and binding of 12 Microcystin (MC) congeners to PPP1 using molecular dynamics simulations: A potential approach in support of an improved MC risk assessment

Sabrina Jaeger-Honz [a], Raymund Hackett [a], Regina Fotler [b], Daniel R. Dietrich [b,*], Falk Schreiber [a,c,**]

[a] Department of Informatics and Information Science, University of Konstanz, Germany
[b] Department of Biology, University of Konstanz, Germany
[c] Faculty of Information Technology, Monash University, Australia

ARTICLE INFO

Keywords:
Molecular dynamics simulation
Microcystin-LR
Microcystin
Conformation
In-silico Toxicology
Protein-ligand interaction
Interaction fingerprints

ABSTRACT

Microcystins (MCs) occur frequently during cyanobacterial blooms worldwide, representing a group of currently about 300 known MC congeners, which are structurally highly similar. Human exposure to MCs via contaminated water, food or dietary supplements can lead to severe intoxications with ensuing high morbidity and in some cases mortality. Currently, one MC congener (MC-LR) is almost exclusively considered for risk assessment (RA) by the WHO. Many MC congeners co-occur during bloom events, of which MC-LR is not the most toxic. Indeed, MC congeners differ dramatically in their inherent toxicity, consequently raising question about the reliability of the WHO RA and the derived guidance values. Molecular dynamics (MD) simulation can aid in understanding differences in toxicity, as experimental validation for all known MC congeners is not feasible. Therefore, we present MD simulations of a total of twelve MC congeners, of which eight MC congeners were simulated for the first time. We show that depending on their structure and toxicity class, MCs adapt to different backbone conformations. These backbone conformations are specific to certain MC congeners and can change or shift to other conformations upon binding to PPP1, affecting the stability of the binding. Analysis of the interactions with PPP1 demonstrated that there are frequently occurring patterns for individual MC congeners, and that published PPP interactions could be reproduced. In addition, common but also unique patterns were found for individual MC congeners, suggesting differences in binding behaviour. The MD simulations presented here therefore enhance our understanding of MC congener-specific differences and demonstrated that congener-specific investigations are prerequisite for allowing characterisation of yet untested or even unknown MC congeners, thereby allowing for a novel potential approach in support of an improved RA of microcystins in humans.

1. Introduction

Microcystin (MC) congeners are a group of structurally similar heptapeptide macrocycles [1], toxic to humans [2]. They are released during and post cyanobacterial blooms into the ambient water and thus can result in critical contamination of water, food and food-stuffs meant for human consumption as well as potentially acutely toxic contaminants of surface waters used for recreational purposes [2–6]. Today, about 300 MC congeners are known, with continuously new congeners being discovered [7]. They consist of structurally closely related cyclic peptides with a common overall structure (consensus sequence: *cyclo*-[D-Ala^1]-[L-X^2]-[β-D-MeAsp^3]-[L-Z^4]-[Adda^5]-[γ-D-Glu^6]-[Mdha^7]) [8,9].

Critical to MC toxicity are two components: toxicodynamic interaction of MCs with ser/thr phosphoprotein phosphatases (PPPs) [10–13] and toxicokinetics, e.g. cellular uptake [14–16]. The cellular uptake of MCs is predominantly mediated by organic anion polypeptides (OATPs) and is dependent on the type and expression level of OATPs in a certain cell, as well as by the affinity and transport capacity of each OATP for different MC congeners [14,15,17]. Once inside the cell, MCs act as

* Corresponding author.
** Corresponding author. Department of Informatics and Information Science, University of Konstanz, Germany.
E-mail addresses: daniel.dietrich@uni-konstanz.de (D.R. Dietrich), falk.schreiber@uni-konstanz.de (F. Schreiber).

https://doi.org/10.1016/j.cbi.2025.111372
Received 10 July 2024; Received in revised form 20 December 2024; Accepted 21 December 2024
Available online 7 January 2025

potent inhibitors of the catalytic subunits of PPPs, exhibiting varying inhibitory capabilities across different PPPs [10]. Nevertheless, for risk assessment (RA), up-to-date guidance documents of the World Health Organization (WHO) focuses nearly exclusively on the toxicity of MC-LR and does not consider other MC congeners [18] despite that a limited number of other congeners have been tested for toxicity [10,19,20]. Thereby, the current RA is primarily based on a 90-day toxicity study in mice in 1999 [21]. However, comparison between rodents and humans has shown that rodents are poor surrogates for humans due to differences in the type of OATPs, expressed across tissues and in the affinity and transport capacity of these OATPs for specific MC congeners [17] Given that MC-LR is not among the most toxic congeners in terms of human PPP inhibition [10], cellular uptake and cytotoxicity [16], the Fawell et al. study [21] appears to be outdated and inadequate as a basis for RA. Instead, it underscores species differences between rodents and humans, highlighting the inadequacy of the current WHO guidelines and the urgent need for a comprehensive revision [2,10].

With serine and threonine residues accounting for 86.4 % and 11.8 % of phosphorylation sites, respectively, PPPs comprise the largest class of human phosphatases [22]. The PPP family in humans includes seven main members: PPP1, PPP2A, PPP2B (Calcineurin), PPP4, PPP5, PPP6, and PPP7, all of which share structurally similar catalytic subunits [22]. Most of the PPPs are expressed ubiquitously, albeit at different levels in the various organs and tissues, e.g. PPP2B is mostly found in the brain, while PPP7 is highly restricted to the retina [23]. Among all phosphatases, PPP1 and PPP2A are particularly essential, performing over 90 % of protein dephosphorylation in eukaryotic cells and playing critical roles in cell cycle regulation, DNA maintenance, and glycogen metabolism [22,23]. As binding to the catalytic subunit and thus inhibition of ser/thr PPPs is the critical component of MC congener toxicodynamics [12,13] and our previous study demonstrated similarities between PPP1 and PPP2A in their inhibition profiles across 18 structurally different MC congeners [10], PPP1 was chosen as representative human PPP for this study here. Due to the structural variability of MCs, high cost of obtaining sufficient amounts and purity of individual congeners for testing via synthesis [24] or isolation from natural resources, appropriate toxicity testing of all the approximately 300 MC congeners is not feasible. On the other hand, computational methods as an alternative to *in vivo* and *in vitro* testing, can aid in exploring the toxins conformations and their interaction as ligands of different proteins and hence their influence on toxicokinetics and -dynamics. More specifically, the interactions of proteins with ligands as well as ligand conformation is critical for the understanding of binding characteristics and hence toxicity. Thus, employing 12 natural and *de novo* synthesised MC congeners, representing the known spectrum of hydrophobicities and molecular weights, and including both common and unusual modifications of the crucial Adda side chain, and for which *in vitro* data were available, allowed a more detailed investigation of the binding characteristics to PPP1 using computational methodology. Indeed, the naturally occurring MCs included the most common congeners MC-LR, MC-LF, MC-RR, MC-LY, and MC-YR, and thus showed the most interesting differences within their measured toxicity in inhibiting PPP1 and PPP5 *in vitro* [10]. In addition, specific modifications within the Adda side chain resulted in significant differences within the PPP binding and thus the inhibition capabilities of MC congeners [16].

Computational methods such as molecular dynamics (MD) simulations can aid in studying both: ligand conformation and interaction [25,26]. Computational studies on other aspects of MC congeners have been proposed, e.g., covalent binding of MCs to ser/thr PPP1 [27], investigating the adsorption mechanism of MCs at the water-mineral interface [28], metal binding selectivity and coordination [29], MC-LR/MC-RR and disinfection by-products [30,31] or MC-LR/MC-RR and conjugates [32,33] binding to PPP1 [30,31]or PPP2A [34], or investigating the sensitivity and binding mode of MCs to antibody binding [35,36].

The focus of the work presented here was to investigate one of the most studied experimental factors, also critical for RA, which is the toxicodynamics of PPP1 inhibition. Despite acknowledging that PPP2A is of equal importance in the toxicodynamics of MCs, the work presented here had to be restricted to PPP1 due to the limited amount of preexisting data available for PPP2A and simulation work and data processing involved. Studying the conformation of MC congeners (ligand) and their interaction with PPP1 has been limited, even though it is the best investigated factor in RA. For our use case, a crystal structure with PPP1 was only available for MC-LR [26]. Because of the large size of MC congeners with more than 35 heavy atoms, predicting the correct 3D structure to study interaction is difficult [15]. In addition, the shallow binding site of PPP1 [27–29], made it unlikely that binding configuration of docking approaches were sufficiently accurate, as docking and scoring are very sensitive to changes in 3D structure [15]. Accordingly, docking was used only to generate an initial starting structure for MD simulations.

To characterise the current knowledge regarding ligand conformation and interactions with PPP1 in more detail, a short overview of the published experimental and computational methods (e. g., nuclear magnetic resonance (NMR), docking, MD simulation) is provided: First efforts to understand the 3D structure of MC-LR were made in the early 1990s whereby the backbone structure was predicted to be planar [37,38]. Subsequent experimental studies with NMR spectroscopy by Rudolf-Böhner et al. [39] suggested, however, that MC-LY adopts a boat-like structure with the Adda and tyrosine side chain collapsing together, therefore restricting the movement of Adda. In contrast, MC-LR, where three conformational families of half-chair-like arrays were described, had a more compact structure than MC-LY. MC-LR, with its Adda and arginine side-chains facing outwards of the ring structure, consequently provided for more conformational freedom. Another NMR study by Bagu et al. [40] confirmed that Adda extends from the macrocyclic ring structure leading to conformational freedom but described MC-LR as a saddle-shaped structure. In 1995 a crystal structure of PPP1 with MC-LR bound was resolved [11] and interactions with PPP1 were studied. The conformation of MC-LR was confirmed as saddle-shaped. MC-LR was suggested to interact with the hydrophobic groove of PPP1 via the Adda-side chain to coordinate with metal ions and water molecules via the carboxylate groups of the Glu residue and carbonyl oxygen between the Adda and Glu residue of MC-LR. In addition, Tyr272 was suggested to interact with leucine, whereby Arg96 and Tyr134 form a hydrogen bond with the carboxylate group of Masp, and Cys273 to form a covalent bond via Mdha residue. In 1996 Trogen et al. [41] showed with NMR spectroscopy one conformational family of MC-LR backbone, which was saddle-shaped for water and DMSO. Based on the 3D structures generated, MD simulation as well as other techniques were used to study conformational exchanges. The MD simulation, run for 372 ps in DMSO, provided for only one conformational cluster.

In 1997 Bagu et al. [42] studied MC-LR and MC-LL in an NMR solution study. For both congeners, a similar saddle-shaped backbone and structure was identified, which was very similar to the crystal structure. MC-LL was docked into PPP1, resulting in a similar binding configuration as for MC-LR in the crystal structure. Therefore, Bagu et al. [42] assumed that both MC-LR and MC-LL have a similar structure and ligand interactions with PPP1, and that the high affinity towards PPP1 is caused by the rigid saddle-shaped conformation of the backbone, which does not change upon ligand binding. In addition, they described the Adda-side chain as a crucial anchor for correct placement in the binding site that facilitates interactions with other residues. In light of the fact that MC-LL did not lead to conformational changes, the authors concluded that other MC congeners will have a similar free and bound structure as MC-LR [42].

Lavigne et al. [43] docked an ensemble of MC-LR NMR structures into the catalytic centre of PPP1 to estimate the free energy of dissociation and calculate PPP1 properties, thereby allowing a comparison with the crystal structure of PPP1 with bound MC-LR. They concluded that numerous docking configuration exist that have a comparable

dissociation energy when compared to the crystal structure, therefore suggesting structural variability and conformational flexibility of the complex. The first known MD simulation study of MC-LR bound to PPP1 was carried out in the same year by Mattila et al. [44]. MC-LR with PPP1 was simulated for 600 ps which did not lead to major structural changes in the binding mode, just minor conformational rearrangements of individual residues. Indeed, four frequently formed hydrogen bonds, rather than three, were found in the crystal structures. The authors also identified other interactions in the crystal structures, e. g., metal coordination, hydrogen bonding with Arg96, Asp220, Arg221, and Gly274, an interaction with Asn124 and His125 via water molecules and an ionic interaction with Glu275, as well as coordination of His248, His66, Asp64 and Asp92. No significant interactions were found for leucine or arginine, while Adda ensured the correct orientation of the MC [44].

In conjunction with advancements in computational power, Jaeger-Honz et al. [45] revisited the MD simulation of MC-LR with PPP1, whereby three previously unstudied MC congeners (MC-LF, [*enantio*-Adda^5]MC-LF and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR) were simulated in addition to the previously studied MC-LR. All MC congeners were simulated in water as solvent (for a total of 45 ns) and in complex with PPP1 (for a total of 750 ns). The authors demonstrated that the PPP1 structure is stable during the simulation and that the MC-LR backbone can adopt at least two conformations in water, rather than the one saddle-shaped structure that was previously reported. In addition, the other MC congeners simulated can adapt to a saddle-shape structure but were also capable of adapting to conformations different to those known for MC-LR. As some MC congeners simulated were moderately or non-toxic compared to MC-LR and MC-LF, the authors concluded that different backbone conformations may explain the differences in tightness of PPP1 binding. A follow-up study by Jaeger-Honz et al. [46], analysing the previously published MD simulation data set of MC-LR and MC-LF in complex with PPP1 [45], showed that these simulations reproduced the interactions with well-known residues described in the literature. Moreover, they were able to demonstrate that the known residues' interactions never occur simultaneously, the existence of representative patterns of ligand interaction, and the presence of additional residues (Phe276, Val223 and Gln249) that are likely to be important for ligand binding.

As only four MC congeners have been studied on a longer simulation timescale, the goal was to simulated eight new MC congeners in complex with PPP1 and water as solvent, in addition to the previously simulated MC-LR, MC-LF, [*enantio*-Adda^5]MC-LF and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR. To secure the experimental basis for all ensuing simulation studies, all MC congeners were analysed (identity, purity and quantity) using NMR [24] prior to their use in PPP1 inhibition assays. The aim of our MD simulations was to understand how conformation of the MC congener backbone and ligand interaction with PPP1 relate to different toxicity classes and ligand structural modifications. Therefore, MC congeners were selected based on structural variety and different toxicity classes

Table 1
MC congeners used in this study with their structural modifications, adapted from Altaner et al. [10], Zemskov et al. [24], and Fotler [16]. The hypervariable positions X and Z are denoted in the three-letter amino acid code.

MC Congener	X	Z	R_1	R_2	Adda
MC-LR	LEU	ARG	Methyl	H	normal
MC-LF	LEU	PHE	Methyl	H	normal
[*enantio*-Adda^5] MC-LF	LEU	PHE	Methyl	H	enantiomeric
[β-D-Asp^3, (*E*)-Dhb^7] MC-RR	ARG	ARG	Hydrogen	CH_3	normal
MC-RR	ARG	ARG	Methyl	H	normal
MC-LY	LEU	TYR	Methyl	H	normal
MC-YR	TYR	ARG	Methyl	H	normal
MC-LY(Prg)	LEU	TYR with propargyl	Methyl	H	normal
[Anda^5] MC-LY (Prg)	LEU	TYR with propargyl	Methyl	H	(2*S*,3*S*,4*E*,6*E*)-3-amino-2-methylnona-4,6-dienoic acid
[Amba^5] MC-LY (Prg)	LEU	TYR with propargyl	Methyl	H	(2*S*,3*S*)-3-amino-2-methylbutanoic acid
[Apha^5] MC-LF	LEU	PHE	Methyl	H	(2*S*,3*S*)-3-amino-2-methyl-7-phenylheptanoic acid
[Apda^5] MC-LF	LEU	PHE	Methyl	H	(2*S*,3*S*)-3-amino-2-methyl-10-phenyldecanoic acid

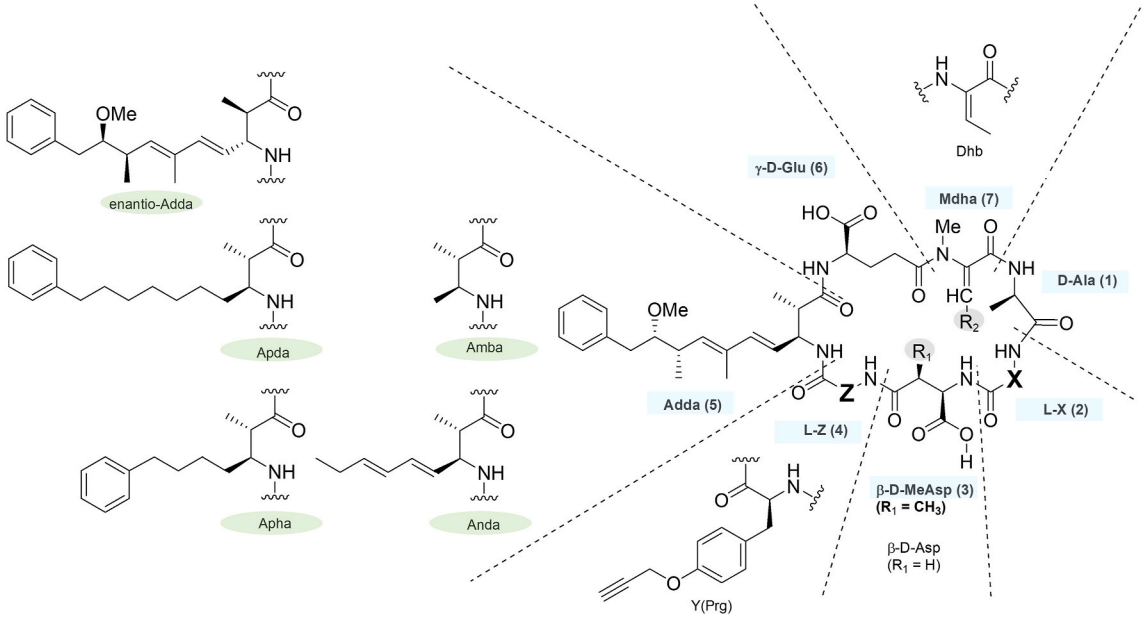

Fig. 1. Consensus structure of the MC congeners used in this study. Dashed lines represent individual amino acids and corresponding position, adapted from Altaner et al. [10]. Further details in Table 1.

(Fig. 1, Tables 1 and 2 [10,16]). Two different MD simulation approaches were chosen to study ligand conformations and interactions: 1) complex simulation of MC congeners with PPP1 for ligand conformations and interactions, and 2) solvent (water) simulation to analyse MC congener conformation and their properties. As the parameterization procedure differs between the previous study by Jaeger-Honz et al. [45] and this study, MD simulations of previously simulated MC congeners were repeated.

2. Methods

2.1. Structure preparation

The crystal structure of PPP1 was downloaded from the protein data bank (PDB code: 1fjm [11]). To prepare the structure for MD simulation, hydrogens were added with Avogadro at a pH 7.4 [47] and the structure corrected with the SWISS-Model server [48–54] by adding 61 unresolved atoms. Subsequently, Molprobity [55,56] was used to add hydrogens and thereby adjusting the protonation, analyse the orientation of asparagine, glutamine and histidine residues, and perform the recommended flips. The structure was then capped at the *N*- and *C*-terminus with an acetyl group and *N*-methyl amide group, respectively, by using PyMol [57].

MC congeners used for this study were selected from a wide range of PPP1 IC_{50} values. The MC-congener specific PPP1 IC_{50} values and structural differences are provided in Fig. 1, Tables 1 and 2. To achieve three-dimensional structures for MD simulation, MC-LR was taken from PDB structure 1fjm by deleting the bond between Cys273 of PPP1 and MC-LR. Hydrogen atoms for pH 7.4 were added with Avogadro [47]. The other MC congener structures were generated by modifying the MC-LR structure with PyMol [57] accordingly. UCSF Chimera was used for structure minimisation with 5000 steps and 500 conjugate steps with a step size of 0.01. This procedure removed structural clashes and resulted in an energy minimised structure [58]. The mol2 file was then reformatted with Avogadro [47].

2.2. Docking

For MD simulation, a 3D complex structure of MC congeners with PPP1 must be provided. For MC-LR a complex structure with PPP1 was available in the PDB. For all other MC congeners, a 3D complex structure with PPP1 was generated by docking. An initial structure was built for each MC congener based on the X-ray structure of MC-LR. Since conformations may differ from MC-LR, solvent MD simulation was used to generate a more diverse set of structures. The solvent simulation was set up as described in Section 2.3. On each solvent (water) simulation, a GROMACS cluster analysis [59] with a cut-off of 0.3 nm was performed and the average middle structures of the first two clusters was extracted. Therefore, a total of three different structures were available for docking. To generate a structure suitable for docking with AutoDock Vina (v1.1.2) [60], bond order and connectivity were corrected, and explicit polar hydrogens were added with PyMOL [57]. The MC congener structures were then minimised with UCSF Chimera [58] as described in Section 2.1 and Avogadro [47] was used to reformat the mol2 file format.

Autodock Tools (v1.5.7) [61] was used to generate a pdbqt file of PPP1 and charges for manganese ions and water were corrected manually. A grid box was centred around the binding site of MC-LR at 91.226, 23.163 and 24.424 with dimensions of 30 × 22.5 × 18.75 °A, in the x, y and z directions, respectively. The side chains of MC congeners were set flexible by using torsion trees and defining bond rotation and flexibilities for side chains. The exhaustiveness was set to 16. The selection of the binding configuration of a given MC congener with PPP1 was based on: 1) the highest affinity value, and 2) when two identical affinities for two structures were present, an orientation similar to the conformation of MC-LR observed in the X-ray crystal structure was preferred, as Adda is a strong anchor of MC-LR in the binding site and strongly influences binding of MC congeners [62]. Table 2 delineates the structures upon which the selected configuration was based on and were subsequently used for further MD simulation.

2.3. Molecular dynamics simulation

The GROMACS version 2020.2 was used to perform MD simulations and Docker images [64] were downloaded from NVIDIA [65]. CHARMM36 force-field as published in July 2020 [66–72] was used for parameterization, which had to be adjusted for the PPP1 structure. Due to capping of PPP1 with *N*-methyl amide and acetylation caps (Section 2.1), compatibility issues were incurred with GROMACS. Therefore, the atom nomenclature was changed for these residues, as suggested in GROMACS User discussions [59]. In addition, parameters for manganese ions are missing in CHARMM36 forcefield. However, previous MD simulations on MC congeners by Jaeger- Honz et al. [45] used parameters of magnesium ions. As the latter are only slightly smaller than

Table 2
Data set of MC congeners with the structures on which the selected AutoDock Vina [60] configuration was based, with their respective PPP1 IC_{50} values [10,16] and binding energy.

MC Congener		Selected Structure	Binding Energy (kcal/mol)	IC_{50} [nM]	Cl_{95} [nM]	Toxicity Class
MC-LR	1	–	–	0.3	0.2–0.4	Toxic
MC-LY	6	Cluster 1	−8.2	0.8	0.7–0.9	Toxic
MC-LF	2	X-ray	−8.2	2.0	1.5–2.6	Toxic
MC-YR	7	X-ray	−7.2	1.2	–	Toxic
MC-RR	5	Cluster 1	−7.5	1.5	1.3–1.8	Toxic
MC-LY(Prg)	8	Cluster 2	−8.5	1.7	1.3–2.2	Toxic
[*β*-D-Asp^3,(*E*)-Dhb^7]MC-RR	4	X-ray[a]	−6.9	62.0	51.7–74.3	Moderately Toxic
[Apda^5]MC-LF	12	X-ray	−7.2	1280	1058–1926	Non-toxic
[Anda^5]MC-LY(Prg)	9	Cluster 2	−7.5	1724	1434–2072	Non-toxic
[Apha^5]MC-LF	11	Cluster 1	−8.1	138,261	–	Non-toxic
[Amba^5]MC-LY(Prg)	10	Cluster 1	−8.7	520,817	449,800–603,048	Non-toxic
[*enantio*-Adda^5]MC-LF	3	X-ray[a]	−7.4	x	–	Non-toxic

Not applicable, crystal structure available.
X-ray: MC congener structure based on X-ray structure of MC-LR has the highest affinity.
Cluster 1: MC congener structure derived with cluster analysis from solvent (water) MD simulation that has the highest affinity. Cluster 1 is the average structure from the largest populated cluster.
Cluster 2: MC congener structure derived with cluster analysis from solvent (water) MD simulation that has the highest affinity. Cluster 2 is the average structure from the second largest populated cluster.
x: IC_{50} > 5 μM [10] and >250 μM of the *enantio*-cyclo[(D)Ala-Leu-(D)Asp-Phe-(L-β-homo-Phe)-(D)Glu-Gly], with a shortened ADDA side-chain [63].
[a] Identical affinity for two configurations, correct orientation preferred.

manganese ions and have similar coordination preferences [45,73], the parameters for manganese ions were set to $\epsilon = 0.277000$ kJ and $\sigma = 0.01234280$ Å [74], which led to a stable coordination of manganese ions in the PPP1 binding site. To parameterise MC congeners for MD simulation, the docked configurations or crystal structure was loaded with PyMOL [57] and bonds corrected before saving to mol2 file format with Avogadro. Then, atoms and bonds were sorted in ascending order [75] and the CHARMM General Force Field (CGenFF) version 2.4.0 [70,72] was used to generate MC congener topology files and convert them to the GROMACS file format.

To model periodic boundary conditions, a dodecahedron was used with a distance of 1.0 nm between the solute and the box. The water solvent model applied was the TIP3P water model [76]. Na^+ and Cl^- were used to neutralise the system and a Verlet cut-off-scheme [77] was applied with Particle-Mesh Ewald (PME) electrostatics [78]. The cut-off for van der Waals interactions was 1.2 nm with a leap-frog integrator and LINCS [79] to constrain bonds to hydrogens. The steepest descent algorithm with 50,000 steps and a step size of ≤ 0.01 with a tolerance of < 10.0 kJ mol^{-1} was used to minimise the energy. Position restraints of 1.000 kJ mol^{-1} nm^2 on all heavy atoms in all axes were applied to all simulations with an MC congener. To stabilise the temperature, an NVT equilibration was performed for 100 ps with a time step of 2 fs with randomised velocity rescaling [80] to 300 K. The Maxwell distribution at 300 K was used to generate initial random velocities. Afterwards, the pressure was stabilised using the Berendsen isotropic pressure coupling [81] to 1.0 bar with NPT equilibration, which was performed for 100 ps with a time step of 2 fs.

Different simulation types were set up for the production run: 1) MC congeners in water, 2) PPP1 in water and 3) MC-congener bound to PPP1 in water. From here onwards, we will refer to the MC congener in water simulation as the "*solvent simulation*", PPP1 in water as the "*apo simulation*" and MC-congener bound to PPP1 in water as the "*complex simulation*". For solvent simulation, three replicates were set up, as these simulations ran fast. For the first replicate, the docked structure of the complex was extracted and placed in solvent as the starting point. This simulation was used to get a GROMACS cluster analysis with the gromos method (pairwise distances) and a cut-off of 0.25 nm. The central representative structure of the two most populated clusters were then used as the starting point for replicate 2 and 3. Due to longer running time, one replicate was set up for "*apo*" and "*complex simulation*", which was based on the crystal structure (available for MC-LR and PPP1) or docked complexes (other MC congeners, see Table 2). All simulations were run with a step size of 2 fs and the output logged every 10 ps. Parrinello-Rahman barostat [82] and randomised velocity rescaling [80] were used. The simulation types and their total simulation times were summarised in Table 3.

2.4. Analysis

To analyse well-equilibrated trajectories from MD simulation, the first 5 ns for solvent or 30 ns for apo and complex simulations were cut off from the trajectory. Subsequently, trajectories were centred on the protein structures and a least-squares fit applied to remove rotational and translational motions, as well as periodic boundary conditions. GROMACS functions were used to analyse the root-mean-square deviation (RMSD), volume, radius of gyration, and solvent accessible surface area for PPP1 and MC congener. Principal Component Analysis (PCA) [83] was applied to analyse the backbone of PPP1 and MC congeners.

Table 3
Summary of employed MD simulations and total simulation time.

Simulation Type	System	Simulation Time
Solvent	MC-congener	105 ns
Apo	PPP1	530 ns
Complex	PPP1–MC-congener	530 ns

The eigenvalues and eigenvectors for PPP1 backbone analysis were derived from the apo simulation and projected on all complex simulations. Normal Mode Analysis was performed in PyMOL [57,84]. This resulted in images (so-called porcupine plots) of the protein backbone with arrows visualising the movement and dynamics along the PCA derived eigenvectors 1 and 2 based on the first and last extreme conformation sampled during the simulation along the eigenvectors. The eigenvalues and eigenvectors for the MC congener backbone analyses were derived from all solvent simulations and projected on the respective complex simulation. Eigenvector 1 and 2 of the PCA projection were visualised with Mol2vec (v0.1) helper function [85], where eigenvectors were shown as scatter plots and were complemented by density histograms to analyse the distribution of the data. To analyse which conformations were at the peaks of the distribution, a 3D density mesh grid was calculated over the data points with scikit-image (v.0.22.0) [86]. For complex simulations, every 50th and for solvent simulations, every 5th data point was used, and the smoothing parameter k was set to 25 and 20, respectively. The structure closest to the individual maximum was then used to determine the representative structure. The representative structures of the solvent and complex simulations were marked as circles and triangles, respectively, on the 2D scatter plot derived with the Mol2vec helper function. The 3D structures of representative MC congener backbones were aligned with Wizard Pair Fitting in PyMOL [57] with the MC-LR backbone observed in the X-ray crystal structure for visualisation.

To calculate further properties and to visualise the data, Python programming language [87] was used. NumPy (v1.21.2) [88] was used to calculate mean values and standard deviations over all replicates. To read and plot xvg files obtained by GROMACS, Jscatter (v1.2.7.2) [89] was used with Matplotlib (v3.5.1) [90]. Each 25th or 2nd data point was plotted for the complex and solvent simulations to reduce visual clutter in the time series visualisation.

To further analyse interactions with PPP1, an Interaction Fingerprint (IFP) analysis, as proposed by Jaeger-Honz et al. [46], was performed. For initial IFP calculation, MDAnalysis (v2.4.0) [91,92] was used to read trajectories and select atoms. In MDAnalysis, Mn^{2+} ions do not have a van der Waals radius assigned. As the parameters of Mn^{2+} for MD simulation were replaced with those of Mg^{2+}, the van der Waals radius of Mg^{2+} was assigned to Mn^{2+}. The IFP for each frame in the complex-simulations was calculated using ProLIF (v1.0.0) [93] and RDKit (v2021.03.5) [94]. All available interactions were used for calculations. However, test runs showed that van der Waals contact calculation between MC congeners and PPP1, H_2O and Mn^{2+} resulted in all interactions being replaced with a van der Waals contact instead of more specific ones. Therefore, van der Waals contact was only analysed between MC congeners and Mn^{2+}, as this interaction should be more unspecific. The resulting IFPs were then further processed with IFPAggVis (v0.1) [46] to aggregate and visualise the results of the IFP analysis. The sliding window filter (x_1 filter) was set to 1.5 %, and the occurrence filter (x_2 filter) to 20 % to stay within the recommended time window of approximately 1.5 ns for an interaction to be detected as present. To estimate similarity within an MC congener simulation, the number of differences between the IFPs was calculated. To estimate similarity between different MC congener simulations, the inverse Rogers–Tanimoto dissimilarity was calculated. As different interactions occurred between different simulations, it was not feasible to calculate the number of differences and interactions, and accordingly, they were combined and set to 0 when not present. As we simulated all MC congeners with the same protein (i.e., PPP1), all MC congeners could theoretically all interact with the same residues. Therefore, it is a valid approach to set this interaction to 0 when absent. This procedure results in a vector of the same length, which is crucial for calculating similarity. A similarity threshold (x) of $x \geq 0.90$ for an IFP was considered as identical, $0.80 \leq x < 0.90$ as similar, and $x \leq 0.50$ as dissimilar an IFP. The interactions are represented as glyphs, see Fig. 2.

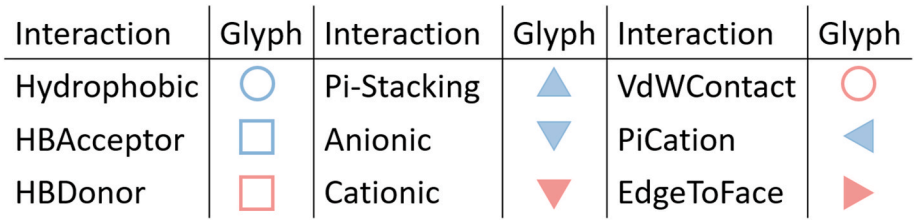

Fig. 2. Interactions encoded by glyphs for network representation of IFPs, VdW = van der Waals.

3. Results and discussion

3.1. Structural stability

To estimate the structural stability of PPP1 and MC congeners during the simulation, the volume, solvent-accessible surface area, radius of gyration and root mean squared deviation were calculated. These properties helped to estimate the structural compactness, its accessibility to solvent and overall changes upon MC congener binding to PPP1.

3.1.1. Stability of PPP1

Overall, PPP1 had a high structural stability (Table S1) independent of MC congener binding. The volume of PPP1 in the apo simulation had a mean value of 60.97 ± 0.56 $\frac{nm}{S^3N}$. In comparison, for complex simulation with MC congeners, PPP1 had a mean value ranging between 60.33 ± 0.52 and 61.42 ± 0.79 $\frac{nm}{S^3N}$, thus in a similar range as the apo simulation volume. Moreover, the time series visualisation (not shown) did not show any major changes or jumps in volume, therefore suggesting the absence of a larger structural change in PPP1 upon MC congener binding.

Similarly, the solvent accessible surface area of PPP1 displayed no changes in the time series visualisations and provided for comparable mean values of PPP1 apo simulation (145.03 ± 2.70 $\frac{nm}{S^2N}$) and complex simulations (mean values ranging between 139.44 ± 2.36 and 145.93 ± 4.53 $\frac{nm}{S^2N}$).

The radius of gyration for PPP1 was also very stable for all simulations, with mean values ranging between 1.87 and 1.91 nm and standard deviations ranging between 0.01 and 0.02 nm.

The RMSD values of PPP1 backbone (Fig. 3 and Fig. S2d) increased for the apo and most MC-congener-complex simulations in comparison to the beginning of the simulation (reference structure). Irrespective of the latter, the RMSD was considered as highly stable because the values were overall low. For the apo simulation the mean value was 0.27 ± 0.03 nm, while for complex simulations the mean values ranged between 0.19 and 0.28 nm with standard deviations between 0.01 and 0.07 nm.

This comparatively high structural stability of PPP1 was already previously reported by Jaeger-Honz et al. [45] for MD simulations with four MC congeners. The calculated properties, however, were overall slightly lower than the values presented here and could have resulted from the different force fields employed. The increased RMSD was most likely the result of RMSD calculations against a reference structure or the start of the simulation, which was based on the crystal structure of PPP1. We therefore hypothesised that some structural rearrangement or further equilibration of the structure could have occurred, as we also noted this trend in the apo simulation. Structural rearrangement or further equilibration of the structure may have resulted from a deletion of the covalent bond of MC-LR in the binding sites, potentially allowing for minor rearrangements, as discussed in more detail in Section 3.2.1.

3.1.2. Stability of MC congeners

Time series visualisation of volume, radius of gyration and solvent-accessible surface area showed minor changes for solvent and complex simulation (data not shown), suggesting high MC congener stability. Our findings suggested that differences in size, compactness and solvent accessibility are likely caused by the different structural modifications (i.e., hyper-variable amino acids at position 2 and 4 or modification of the Adda-side chain) of MC congeners.

Indeed, as all three calculated properties (size, compactness and solvent accessibility) are related, similar trends were observed (Table S2). For these properties, an increase in mean values and standard deviations was observed when comparing solvent to complex simulation, except for MC-LR, MC-LF and MC-LY, where we found similar values, and [Anda^5]MC-LY(Prg) and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR, where we observed a decrease. For all MC congeners, the mean values of complex simulation were nevertheless still within the range of the standard deviation of the mean value of the solvent simulation, thereby suggesting no significant differences of the mean values observed in the complex and the solvent simulation. The latter thus corroborated earlier reports suggesting that the structure of MC-LR does not change when in solvent compared to binding to PPP1 [11,41,44]. Accordingly, the differences we observed in mean values and standard deviation between individual MC congeners, are most likely explained by their structure rather than their toxicity class, which indeed was derived from the binding properties to PPP1. The latter findings stand in contrast to previous simulations of Jaeger-Honz et al. [45], where no increase in the calculated properties (size, compactness and solvent accessibility) was

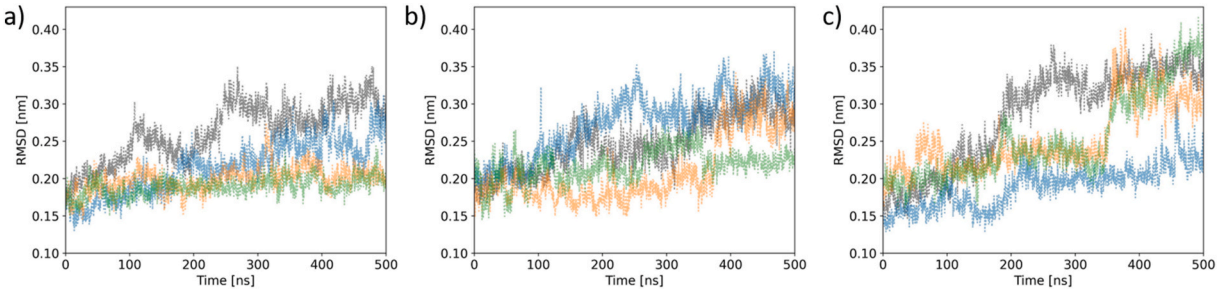

Fig. 3. PPP1 backbone RMSD in complex with different MC congeners. The MC congeners shown are: a) PPP1–MC-LR (grey), PPP1–[β-D-Asp^3,(*E*)-Dhb^7]MC-RR (blue), PPP1–MC-RR (orange) and PPP1–MC-YR (green), b) PPP1–MC-LF (grey), PPP1–[*enantio*-Adda^5]MC-LF (blue), PPP1–[Apha^5]MC-LF (orange) and PPP1–[Apda^5]MC-LF (green), and c) PPP1–MC-LY (grey), PPP1–MC-LY(Prg) (blue), PPP1–[Anda^5]MC-LY(Prg) (orange) and PPP1–[Amba^5]MC-LY(Prg) (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

reported for MC-LF and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR when binding to PPP1. However, when comparing the mean values of these properties of four previously simulated MC congeners (MC-LR, MC-LF, [β-D-Asp^3, (*E*)-Dhb^7]MC-RR, and [*enantio*-Adda^5]MC-LF) in this simulation, they all reached similar or slightly higher mean values when compared to the simulations by Jaeger-Honz et al. [45], suggesting that these small differences were most likely caused by the use a different force field.

Previous structural investigations assumed that the hydrophobic hypervariable amino acid phenylalanine would result in a more coil-like structure in the solvent when compared to complex simulation, thereby shielding this residue from water [45]. In accordance with the assumption, the mean properties of MC-LF were more constant, with small increases for radius of gyration and solvent accessible surface area. Consequently, the findings reported here corroborate earlier structural investigations, as the structure did lose some compactness when binding to PPP1. In contrast, [*enantio*-Adda^5]MC-LF increased in volume and solvent-accessible surface area upon binding to PPP1, thus suggesting that [*enantio*-Adda^5]MC-LF was more exposed to solvent than its toxic entmer MC-LF.

The more hydrophilic MC congeners ([β-D-Asp^3,(*E*)-Dhb^7]MC-RR, MC-RR) had higher mean values for all three properties in the solvent and complex simulations, when compared to the hydrophobic MC congeners (e. g., MC-LF, MC-LY). The more hydrophilic MC congeners also adopted a more open structure in the solvent due to the positively charged amino acid arginine that allowed for interaction with surrounding water molecules. MC-RR had an even more open structure when bound to PPP1, resulting from the arginines at positions 2 and 4 which, in conjunction with the increased positive charge, led to a more open and water-accessible conformation.

MC congeners with modified Adda-side chains ([Amba^5]MC-LY (Prg), [Anda^5]MC-LY(Prg), [Apha^5]MC-LF and [Apda^5]MC-LF) had overall lower mean values for all properties in comparison to MC congeners without modified Adda-side chain (MC-LY(Prg), MC-LF). This was expected for [Amba^5]MC-LY(Prg), [Anda^5]MC-LY(Prg), and [Apha^5]MC-LF as the side chain of residue 5 is smaller compared to Adda and therefore should result in lower mean values of the three calculated properties. However, [Apda^5]MC-LF had a size comparable to Adda but appeared to have a more compact folding, which could have resulted from the removal of side chains and double bonds in comparison to Adda.

The all-atom RMSD indicates a similar trend (Table S2, Figs. S4 and S5). For some MC congeners ([*enantio*-Adda^5]MC-LF, [β-D-Asp^3,(*E*)-Dhb^7]MC-RR, MC-LY, MC-LY(Prg), [Anda^5]MC-LY(Prg), and [Apha^5]MC-LF) there was an increase in all-atom RMSD in the complex simulation (when compared to solvent simulation), which was also clearly visible in the all-atom RMSD time series visualisation (Fig. S5). A shift in the [*enantio*-Adda^5]MC-LF congener structure in the complex simulation was also noted when compared to the reference structure at the beginning of the determination. With the expected exception of [*enantio*-Adda^5]MC-LF, all MC congeners were coordinated within the binding site of PPP1. Accordingly, we assumed that the initial binding configuration determined by docking for [*enantio*-Adda^5]MC-LF did not fit well, therefore suggesting that some structural rearrangements must have occurred. In contrast, the all-atom RMSD for MC-LR, MC-LF, MC-RR, MC-YR, [Amba^5]MC-LY(Prg) and [Apda^5]MC-LF decreased from solvent to complex simulation. While this was expected for MC-LR, due to the fact that the X-ray crystal structure was used as initial configuration for MD simulation where a good fit in the binding site is ensured, the initial docking configuration provided for a good fit for the other MC congeners as only very small structural rearrangements were needed.

3.2. Conformational space analysis

Principal component analysis (PCA) was used to analyse the conformational space of apo-PPP1 and MC congener backbone individually, and as a PPP1–MC complex. Due to the high structural stability of PPP1, normal mode vector analysis was performed to visualise extreme movement along the individual eigenvectors derived by PCA. The different conformations derived from the PCA analyses also represented different energy states [95] of the different conformations.

3.2.1. PCA of the PPP1 backbone

The movements of PPP1 along the eigenvectors, of the complex simulations, i.e., when bound to MC congeners, were overall smaller, but scattered across the same conformational space (e.g., Fig. S6), and therefore less distributed than the movements of apo-PPP1 along the Eigenvectors of the apo simulation. The extreme movement of the PPP1 backbone, determined with normal mode vector analysis, resulted in porcupine plots that visualised the movement along the individual eigenvectors. The apo simulation (Fig. 4a and b) provided for the most extreme PPP1 backbone movement. The PPP1–MC congener complex simulations for all other MC congeners provided for similar but less extensive movements in the PPP1 backbone and therefore shorter normal mode arrows (data not shown). Movement along eigenvector 1 demonstrated the outward bending of the A′ α-helix loop of the *N*-subdomain to the 1′ β-sheet of the *N*-subdomain (Fig. 4c) at the bottom of PPP1 structure. Similarly, at the top of the PPP1 structure, the loop extending from the β-sheet 8 of the *C*-subdomain to the β-sheet 9 of the *C*-subdomain had some twisting movement which led to a horizontal instead of vertical orientation towards the binding site. On the top right side of the structure behind the H-α-helix of the C subdomain, the loop stretched at the same time towards the top. Movement along eigenvector 2 (Fig. 4b) is similar to movement along eigenvector 1 located at the top and bottom parts of PPP1, as well as some movement of the *C*-terminus along eigenvector 2. In conclusion, PCA and normal mode analysis, in agreement with the calculated structural properties, indicated a very stable structure of PPP1 with only minor changes due to MC congener binding and no major structural rearrangements of the binding site.

3.2.2. PCA of the MC congener backbone

The MC congener backbone structure was also analysed using PCA (Fig. 5 and Figs. S7–S12) and showed comparable results to those reported in the literature [11,39–45].

Indeed, in the early 1990s, the backbone conformation of **MC-LR** was first predicted to be planar [37,38]. Subsequently, the backbone of MC-LR was described as saddle-shaped conformation in water [11, 39–42] and in the complex with PPP1 [11,44]. Lavigne et al. [43] described that various different structures of MC-LR can exist and bind to PPP1 and that the complex was flexible in solution. Later MD simulations could identify a second conformational cluster in solvent simulation which had a planar shape as well as saddle-shaped structures in complex and solvent simulations [45]. Corroborating the latter findings, we identified two major conformations in solvent that were saddle-shaped (orange circle, Fig. 5a) and planar (cyan circle, Fig. 5a), and a saddle-shaped conformation when bound to PPP1 (orange triangle, Fig. 5a and Fig. S7a).

For MC congeners (MC-LF, [*enantio*-Adda^5]MC-LF and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR) calculated backbone conformations were reported by Jaeger-Honz et al. [45]. While, for MC-LF the previous simulations showed two conformational clusters in solvent — one saddle-shaped and one slightly more stretched conformation, we derived one saddle-shaped conformation from the solvent (orange circle) and complex simulation (orange triangle) in this simulation, which was structurally almost identical with the conformations derived for MC-LR (Fig. 5e and Fig. S7b). The latter finding does not stand in conflict with the previous simulations, as Jaeger-Honz et al. [45] noted that the second, stretched conformation was presumed unstable, as it was very sparsely populated.

In contrast to the toxic MC-LF, the backbone structures of **[*enantio*-Adda^5] MC-LF** were planar, irrespective of the simulation type (Fig. 5f and Fig. S8a). Apparently, the side-chain stereocenters in the Adda^5 residue influence the overall backbone structure and may explain the difference in toxicity. Jaeger-Honz et al. [45] described the backbone

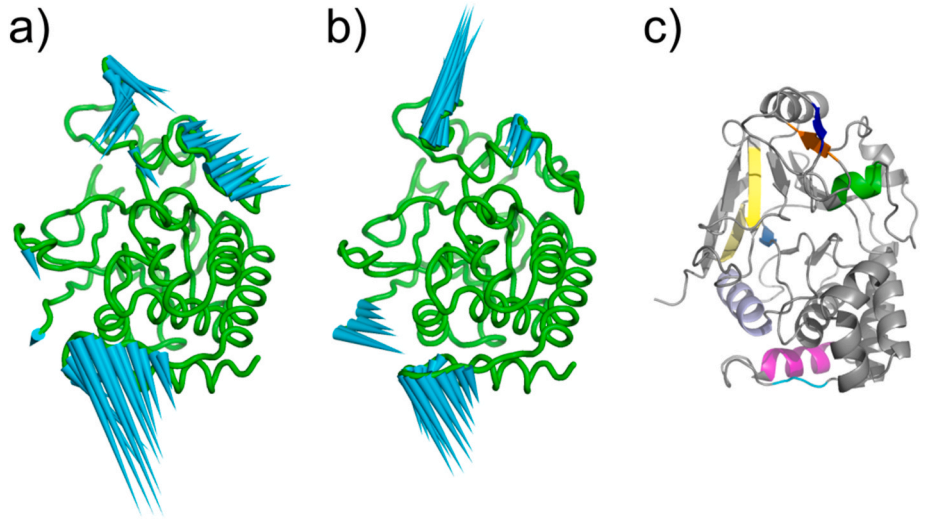

Fig. 4. Porcupine plot of the backbone movement of PPP1 and its structural elements. a) and b) Porcupine plots of movement in PPP1 backbone of the apo simulation as projected in extreme conformations along eigenvectors 1 and 2, respectively, derived by PCA. Arrows indicate movements above 4 Å. c) Structural elements of PPP1 (image generated with PyMol [57]). Pink: A′ α-helix of the *N*-subdomain, cyan: β-sheet 1′ of the *N*-subdomain, green: H-α-helix of the *C*-subdomain, dark blue: β-sheet 8 of sheet 3 in the *C*-subdomain, orange: β-sheet 9 of sheet 3 in the *C*-subdomain, yellow: β-sheet 12 in sheet 2 of the *C*-subdomain, pale yellow: β-sheet 13 of the *C*-subdomain, marine: β-sheet 2 of the *N*-subdomain of sheet 1, and light blue: α-helix B in the *N*-subdomain. Structural element notation was taken from Goldberg et al. [11]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

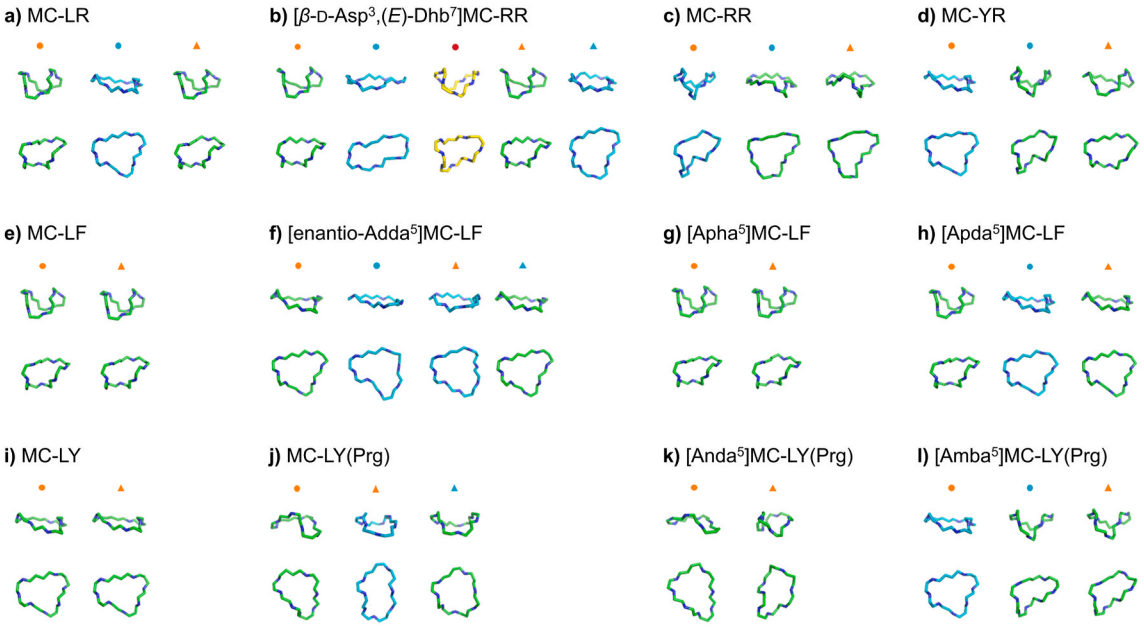

Fig. 5. Representative conformations of MC congener backbone derived after analysis of the 2D principal component projection from solvent (circles) and complex (triangles) MD simulation in side and top view (see Figs. S7–S12). The colouring of the molecular structures should help to distinguish the individual structures within a complex or solvent simulation.

conformation for solvent and complex as a stretched backbone, but not planar as shown here. The latter discrepancy most likely stemmed from the fact that structure processing was more detailed in the simulation reported here.

Similar to previous simulations [45] and to the simulations with [*enantio*-Adda^5]MC-LF, simulations of the moderately toxic **[β-D-Asp^3, (*E*)-Dhb^7] MC-RR** provided for a widely distributed scatter plot in the 2D projection (Fig. 5b and Fig. S8b). A saddle-shaped conformation was found for solvent and complex (orange circle and triangle), as well as a planar structure for solvent and complex (cyan circle and triangle). For the planar structure, the shape between solvent and complex simulation was slightly different. The cyan triangle of complex simulation occupied a conformational space, which appeared to be only accessible when binding to PPP1. Accordingly, it is suggested that [β-D-Asp^3,(*E*)-Dhb^7] MC-RR will undergo a conformational change when binding to PPP1. A third conformation was detected in the solvent simulation (red circle) which was more skewed and V-shaped compared to the other previously discussed MC backbone structures and was never reported before in the literature.

MC-RR (Fig. 5c and Fig. S9a), differing from [β-D-Asp^3,(*E*)-Dhb^7]MC-RR in residues 3 and 7 is a toxic MC congener. The simulation clusters in the 2D projection were less scattered and better defined and resulted in two structures that could be identified for solvent simulation: a major representative with V-shaped backbone (orange circle) and a minor

representative (cyan circle) with a more planar structure with a downward bend to the right, which is surrounded by the complex cluster. The major structure of the complex cluster (orange triangle) was similar in shape to the minor conformational structure of the solvent (cyan circle). However, the structures are completely different to the ones reported for [*β*-D-Asp^3,(*E*)-Dhb^7]MC-RR, suggesting that minor modifications of side chains 3 and 7 highly influenced the shape of the overall backbone structure.

Two solvent clusters and a large, widely distributed complex cluster could be detected for **MC-YR** (Fig. 5d and Fig. S10a). The large solvent cluster had a boat-like structure (orange circle), which is similar to the structure observed for MC-LY (*vide infra*). The minor solvent cluster was V-shaped (cyan circle), similar to the structure observed for solvent cluster of MC-RR (orange circle in Fig. 5c). The backbone structure of MC-YR found in the complex simulation was also saddle-shaped, similar to MC-LR. Therefore, a conformational shift appeared to occur upon binding of MC-YR to PPP1.

MC-LY is a toxic MC congener with well overlapping clusters for solvent and complex simulation (Fig. 5i and Fig. S9b). One structural representative could be identified for each simulation type that looked overall very similar amongst solvent and complex simulations. The structure was mostly planar, but had skewed parts at both ends of the ring, resulting in a boat conformation. The latter is in agreement with the solvent simulation (cyan circle), but more bent down to the right. In contrast to the other toxic MC congeners MC-LR and MC-LF and moderately toxic [*β*-D-Asp^3,(*E*)-Dhb^7]MC-RR, the backbone structure of MC-LY was reported in an experimental study from 1994, where the structure of MC-LY in DMSO was determined with NMR spectroscopy [39]. Interestingly, MC-LY and MC-LF differ in only one hydroxy group but were completely different in their backbone conformation (boat shape vs. saddle shape), although both are toxic MC congeners.

For **MC-LY(Prg)** (Fig. 5j and Fig. S10b) the solvent structure was shaped like an inverted U (orange circle). In contrast, both structures for complex simulation were skewed upwards towards both ends of the MC congener backbones, leading to a boat-like or U-shaped structure (orange and cyan triangle). The small modification in comparison to MC-LY resulted in different backbone conformations, especially for solvent simulation.

[Anda^5] MC-LY(Prg) is modified at the Adda-side chain compared to MC-LY(Prg) and is a non-toxic MC congener (Fig. 5k and Fig. S11a). The solvent conformation identified was also an inverted U-shape (orange circle), but not as bent as MC-LY(Prg). In contrast, the reported conformation detected in the complex simulation, was highly skewed (orange triangle) and more curved than the saddle-shaped backbone structures described above (*vide supra*). We therefore concluded that modification of the Adda-side chain influenced the backbone conformation. Therefore, binding to PPP1 may be altered and thus could explain the differences in binding to PPP1 when compared to the toxic MC-LY(Prg).

[Amba^5] MC-LY(Prg) is also modified at the Adda residue (whole side chain removed) in comparison to MC-LY(Prg) and a non-toxic MC congener (Fig. 5l and Fig. S11b). Two conformational clusters were identified for the solvent simulations. One planar ring (orange circle) similar to the structures of solvent simulations already identified for MC-YR (orange circle), MC-LR (blue circle) and both structures of MC-LY (orange circle and triangle). The second conformation observed was V-shaped, similar to the solvent structures of MC-YR (cyan circle) and MC-RR (orange circle). The structure determined for complex simulation was also V-shaped (orange triangle). Accordingly, [Amba^5]MC-LY (Prg) is not presumed to undergo conformational change upon binding to PPP1.

[Apha^5] MC-LF is modified at the Adda-side chain when compared to MC-LF. In contrast to MC-LF, it is a non-toxic MC congener (Fig. 5g and Fig. S12a). Both structures determined for solvent and complex simulation were saddle-shaped, comparable to the structure of MC-LF. The modification of Adda to Apha^5 apparently did not influence the backbone conformation of the MC congener but had an effect on its toxicity.

[Apda^5] MC-LF (Fig. 5h and Fig. S12b) is also modified at the Adda-side chain when compared to MC-LF and is also a non-toxic MC congener. The major conformation identified for solvent simulation was also saddle-shaped (orange circle) as was the case for many other MC congeners. Nevertheless, we were able to identify a second conformation for solvent structures, as well as complex simulation, which was more planar and slightly skewed at both ends of the structures (cyan circles and triangles). The latter structure was similar to those of the solvent simulations for MC-YR (orange circle), MC-LR (blue circle), solvent and complex simulation of MC-LY (orange circle and triangle) and [Amba^5] MC-LY(Prg) (orange circle). Consequently, the modification of the Adda-side chain in the backbone did lead to a conformational shift towards a more open or planar structure upon binding to PPP1.

In summary, the saddle-shaped structure could be identified for several MC congeners (i.e., MC-LR, MC-LF, [*β*-D-Asp^3,(*E*)-Dhb^7]MC-RR, [Apha^5]MC-LF) in both solvent and complex simulations, whereas for [Apda^5]MC-LF it was only observed in the solvent simulation. The modification of MC congeners at the Adda-side chain can result in a change of the overall conformation of the macrocyclic ring structure compared to their unmodified toxic MC congeners (i.e., MC-LF and MC-LY(Prg)), e.g. for [*enantio*-Adda^5]MC-LF, [Anda^5]MC-LY(Prg), [Anda^5] MC-LY(Prg) and [Amba^5]MC-LY(Prg). As the latter MC congeners are all non-toxic, one could presume that the changes in the Adda-side chain and thus backbone structure could be critical for toxicity. However, these conformations are not necessarily unique to non-toxic MC congeners and can be also adopted by toxic MC congeners. Indeed, a previous study by Bagu et al. [42] suggested that MC-LR and MC-LL have saddle-shaped structures with minor differences and proposed that docking suggested similar binding sites and conformations of MC-LL to MC-LR. They therefore proposed that all other MC congeners will have a similar structure in solution and in complex to PPP1. We found differences in conformational behaviour of MC congeners, largely dependent on the amino acids at the hypervariable positions 2 and 4 but also of other structural modifications, such as [*β*-D-Asp^3,(*E*)-Dhb^7]. In addition, we were able to demonstrate that some MC congeners undergo conformational changes upon binding to PPP1 (e.g., MC-YR, [Apda^5]MC-LF, MC-LY(Prg), [Anda^5]MC-LY(Prg) and [Amba^5]MC-LY(Prg)). However, conformational changes upon binding to PPP1 is not prerequisite (e.g., MC-LR, MC-LF, MC-LY). Moreover, multiple conformations can exist for individual MC congeners and hence may influence their binding behaviour.

3.3. Interaction fingerprints analysis

Although visual inspection of MD simulations suggests that [*enantio*-Adda^5]MC-LF moves out of PPP1 binding site (i.e. is not binding to PPP1), while all other MC congeners remained stable in the binding site, the conformational analysis of MC congeners also demonstrated that conformations of toxic and non-toxic MC congeners partially overlap. Thus, conformational analyses in and of itself will not allow distinction of toxic from non- or moderately toxic MC congeners. Consequently, interaction fingerprints (IFPs) were employed to investigate the interactions of MC congeners with PPP1 in more detail.

3.3.1. IFP comparison within MC congener simulation

To allow evaluation of IFP development within one simulation, a combined visualisation of IFP similarity, number of differences, occurrence of individual IFP and identical IFP was carried out (Fig. 6 and Fig. S13).

To enable identification of the most important representative IFP, occurrence visualisation of individual IFPs can be helpful, allowing determination of the most frequent IFPs, their appearance within the simulation, and their distribution over time. Indeed, matrix visualisation helped to identify regions of higher (dark blue) and lower similarity

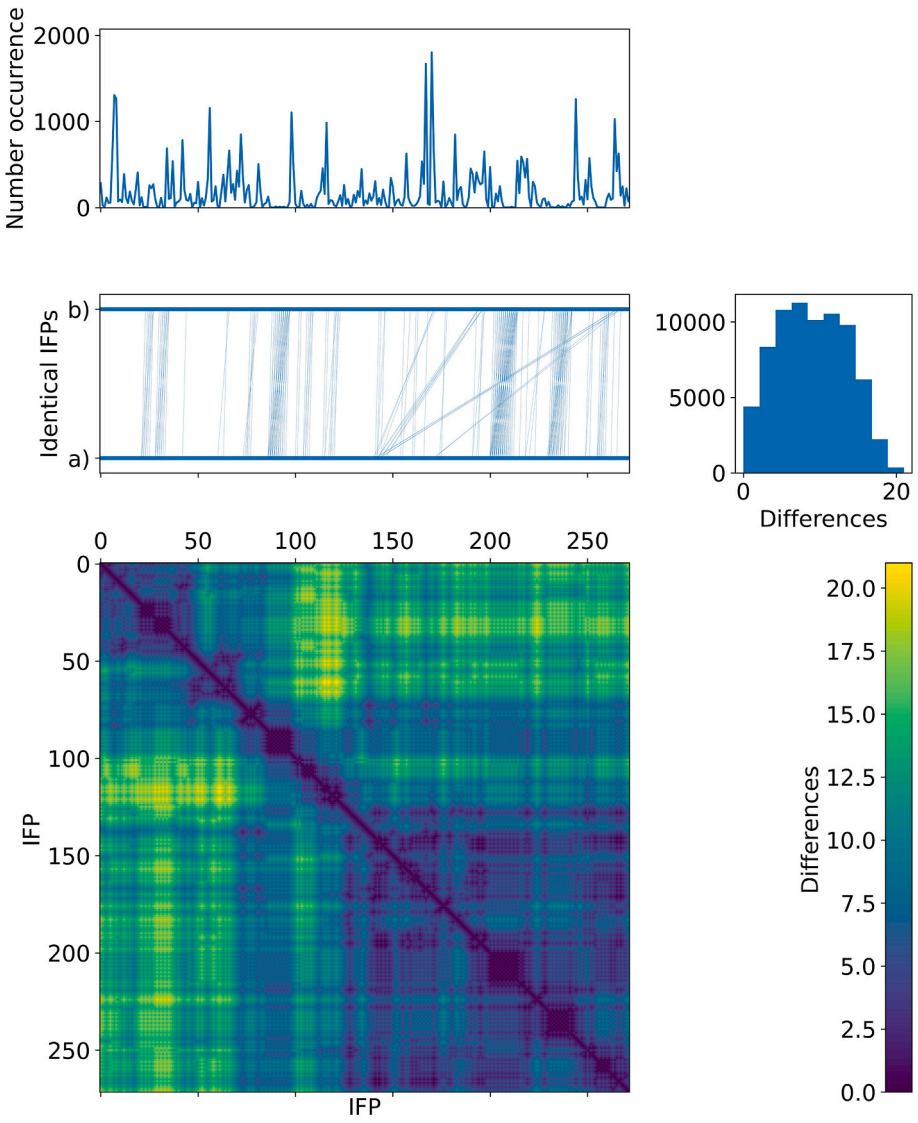

Fig. 6. Comparison of IFP similarity within one MD simulation of MC-LR. The occurrence and identical IFP (connected by vertical lines between a and b) are shown as a line plot. The occurrence plot shows how many individual IFPs were summarised, as they occurred immediately after each other and had identical interactions. The number of differences to all IFPs are shown as a histogram (top right), and as colour in the matrix visualisation. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(yellow). As changes occurred and accumulated over time, locally defined regions of higher similarity were identifiable, largely depending on the MC congener. IFPs that are revisited during the simulation, can be identified with the line plot via vertical connection of identical IFPs.

It is important to note that for some non-toxic MC congeners (i.e., [*enantio*-Adda^5]MC-LF, [Anda^5]MC-LY(Prg) and [Apha^5]MC-LF), very few frequently occurring representative IFPs were identified, while the other moderately toxic and toxic MC congeners (i.e., MC-LR (Fig. 6), MC-LF, MC-RR, [β-D-Asp^3,(*E*)-Dhb^7]MC-RR, MC-LY, MC-YR, MC-LY (Prg), [Amba^5]MC-LY(Prg), and [Apda^5]MC-LF) had many frequently occurring representative IFPs. Moreover, the similarity matrix showed very few higher order arrangements for [*enantio*-Adda^5]MC-LF, MC-RR, and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR, while for other MC congeners many more clusters or higher order arrangements were visible.

To further analyse the interactions, we calculated the minimum, mean, and maximum of interactions for each IFP per MC congener (Table 4 and Fig. S14). The comparison of the mean of the number of interactions for each IFP of toxic to moderately or non-toxic MC congeners demonstrated that the mean number of interactions of toxic MC congeners were higher when compared to their corresponding non- or moderately toxic structurally similar MC congeners (i.e., MC-RR versus [β-D-Asp^3,(*E*)-Dhb^7]MC-RR, MC-LF versus [*enantio*-Adda^5]MC-LF, MC-LY(Prg) versus [Anda^5]MC-LY(Prg) and [Amba^5]MC-LY(Prg)). Overall,

Table 4
Minimum, mean, and maximum number of interactions detected per IFP of MC congeners simulated with PPP1.

MC Congener	Minimum	Mean	Maximum	Toxicity Class
MC-LR	9	14.25	22	Toxic
[β-D-Asp^3,(*E*)-Dhb^7]MC-RR	3	8.28	18	Moderately toxic
MC-RR	13	17.30	22	Toxic
MC-YR	11	15.64	24	Toxic
MC-LF	12	16.36	21	Toxic
[*enantio*-Adda^5]MC-LF	5	8.96	14	Non-toxic
[Apha^5]MC-LF	8	13.65	19	Non-toxic
[Apda^5]MC-LF	10	15.67	24	Non-toxic
MC-LY	6	14.41	20	Toxic
MC-LY(Prg)	9	15.84	28	Toxic
[Anda^5]MC-LY(Prg)	7	11.40	16	Non-toxic
[Amba^5]MC-LY(Prg)	5	11.36	16	Non-toxic

the mean of interactions for non and moderately-toxic MC congeners tended to be lower than for toxic congeners, except for the non-toxic congeners [Apha[5]]MC-LF and [Apda[5]]MC-LF (Fig. S1). Indeed, when compared to its toxic structurally similar MC-LF, the mean IFP values of [Apha[5]]MC-LF and [Apda[5]]MC-LF were lower, albeit close to the 16.36 of MC-LF.

As [Apda[5]]MC-LF merely differs from MC-LF by the missing side-chain of Adda while maintaining the overall size, a similar number of interactions would suggest that the size of the Adda-side chain influenced the overall backbone structure and thus also the number of interactions. Our PCA projections, however, suggested that the backbone conformations of MC-LF was saddle-shaped while [Apda[5]]MC-LF, a non-toxic MC congener, was suggested to be planar for complex simulations comparable to the non-toxic [*enantio*-Adda[5]]MC-LF.

Similar trends were observed for the number of interactions in the major representative IFP (Fig. 7 and Fig. S15, except for [Apha[5]]MC-LF, MC-LY(Prg), [Anda[5]]MC-LY(Prg), and [Amba[5]]MC-LY(Prg), where the number of interactions were similar to their respective mean value.

It is noteworthy that the interactions of at least two out of three most frequent IFPs were similar for all MC congeners. For [*enantio*-Adda[5]]MC-LF, we could determine that two were identical (Figs. S15m and n), leading to one very frequently occurring interaction pattern with only 7 interactions. Although individual MC congeners showed different interaction patterns, we were able to identify several critical residues for nearly all MC congeners (exception: [*enantio*-Adda[5]]MC-LF and [*β*-D-Asp[3](*E*)-Dhb[7]]MC-RR). Indeed, Fontanillo and Köhn [62]), already described interactions with residues of PPP1: hydrogen bonds (Arg96, Tyr134, and water (indirect coordination to manganese ions)), replacement of interaction with water (Asn124, His125, Ile130, Tyr134 and Trp206), hydrophobic interactions (Cys127, Ile130, Ile133, Trp206, Tyr272 and Gly274), and a covalent bond (Cys273). Interestingly, the observed interacting residues never occurred together at the same time in one representative IFP. We therefore concluded that different interaction patterns exist within each of the MC congener simulations when complexed to PPP1. In addition, individual residues were not limited to one type of interaction, thus potentially allowing identification of multiple interactions that can be formed with MC congeners.

3.3.2. *IFP comparison amongst MC congener simulations*

To compare the similarity of simulations amongst different MC congeners, interactions were merged to generate a vector with the same length to calculate similarity. The evaluation of similarity was then changed to the inverse Rogers–Tanimoto dissimilarity with an "identical IFP" threshold of $x \geq 0.90$, "similar IFP" between $0.80 \leq x < 0.90$, and $x \leq 0.50$ as "dissimilar IFP" (see Fig. 8 and Table S3).

Line plots to visualise which IFPs were similar or identical are shown in Fig. 9 and Figs. S16–S33. While evaluating identical and similar IFPs with the highest occurrence in the simulations, we found that for almost all MC congeners (except for some frequent IFP of [*enantio*-Adda[5]]MC-LF and [*β*-D-Asp[3],(*E*)-Dhb[7]]MC-RR), we could identify residues known to interact from Fontanillo and Köhn [62]).

Most MC congeners shared similar IFPs with MC-LR. However, identical IFPs could only be found with MC-LR. The latter was most likely due to the starting configuration of the MD simulation, as in most cases it was close to that of MC-LR. In addition, MC-LR and MC-LF had a high number of similar IFPs themselves (32.03 %), which was previously also described in Jaeger-Honz et al. [46]. Moreover, corroborating earlier work by Jaeger-Honz et al. [46], we were able to identify some or all of the interactions (Phe276, Val223, Gln249) that were proposed as important for many MC congeners and could also identify additional residues (His66, Val250, Tyr272) that we propose as being important for binding to PPP1. Indeed, substitution of Phe276 and mutation of Tyr272 was suggested to reduce the ability of PPP1 to bind MC congeners [96, 97].

We were able to identify some MC congeners (MC-LF, MC-LY, [Apha[5]] MC-LF) that have a high proportion of similar IFP with MC-LR (32.02 %, 35.30 % and 40.39 %, respectively) and also have large areas of overlap in the line plot, indicating similar IFP patterns. For MC-LR and MC-LF, Fig. 9b and **c** show the two similar IFPs with the highest occurrence within the simulation. In comparison to other IFPs, both of them occurred frequently for both MC congeners, but did not belong to the most frequently occurring IFP. For MC-LR and MC-LY, a small proportion of identical (0.71 %) and a high proportion of similar IFPs (35.36 %) were identified and overlapped in the line plot as well. However, the individual IFP had a rather low occurrence with a lower number of interactions (11 and 10, respectively) and were therefore less frequent. For MC-LR and [Apha[5]]MC-LF, we were able to identify a small proportion of identical (3.22 %) and the highest proportion of similar IFPs (40.39 %) across the whole data set. The selected similar representative IFPs occurred frequently in both simulations and had 11 and 14 interactions with PPP1. Both MC congeners belonged to different toxicity classes, thereby providing an unexpected result. As we identified a similar saddle-shaped conformation of [Apha[5]]MC-LF when compared to MC-LR, it may be that the conformation of [Apha[5]]MC-LF did indeed favour a similar binding mode to MC-LR. Further simulations are therefore needed to assess this in more detail.

For other MC congeners, little, or some overlap between certain areas of the simulation could be observed. The similarity values ranged between 10.40 % and 30.49 %. Depending on the simulations that were compared, we could show that for some IFPs a very low number of interactions were retrieved ([*enantio*-Adda[5]]MC-LF and [*β*-D-Asp[3],(*E*)-Dhb[7]]MC-RR, [*β*-D-Asp[3],(*E*)-Dhb[7]]MC-RR and MC-LY). Therefore, the high proportion of similar IFPs might be an artefact caused by the merging of two IFP sets. In other simulations, we could identify some

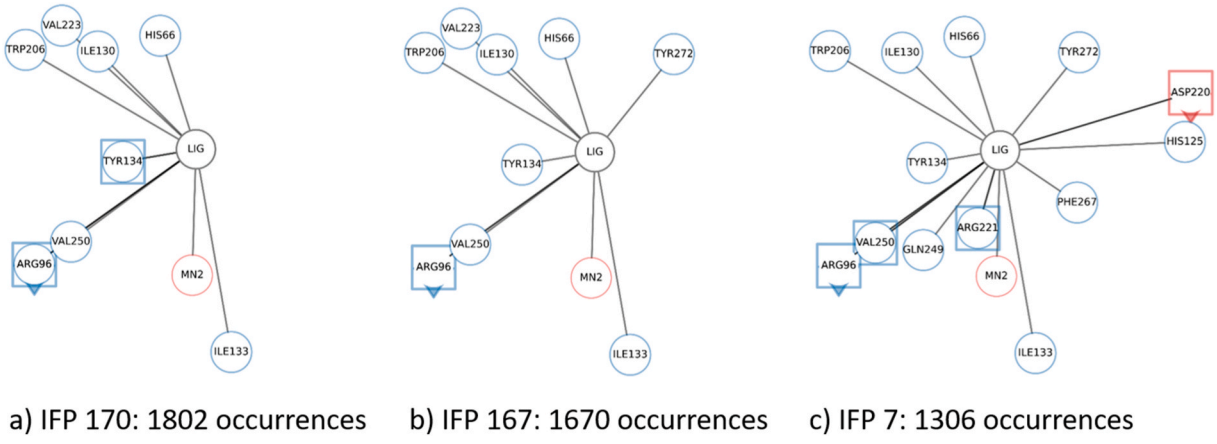

Fig. 7. Three most frequent occurring IFP for MC-LR.

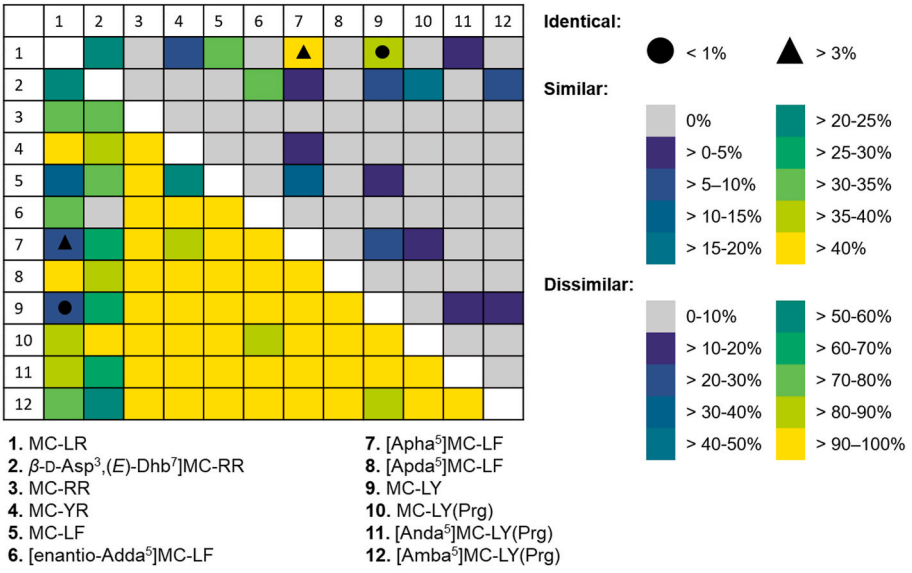

Fig. 8. Summary of identical, similar and dissimilar IFPs of MC congeners interacting with PPP1. The number of IFPs in the respective category had been calculated as percentages which are encoded as indicated in the figure. The upper right and lower left of the diagonal represent similar and dissimilar IFP, respectively. Identical IFP are indicated by glyphs.

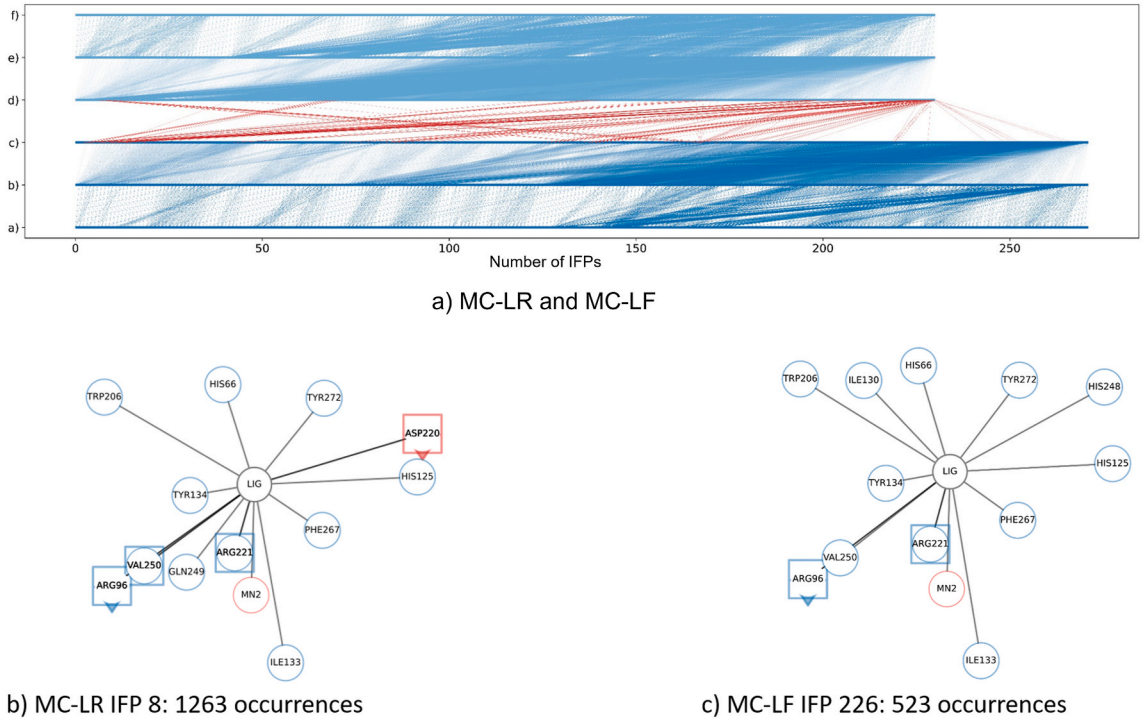

Fig. 9. Similarity calculation between IFPs of different simulations. a) MC-LR is shown in dark blue, MC-LF in light blue. Between a–b and e–f, identical IFPs within the same simulation are shown. Between b–c and d–e, similar IFPs within the same simulation are shown. Between c–d similar IFPs between simulations are shown as red lines. The most frequently occurring similar IFPs are shown in b) and c). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

similar IFP that had rather low occurrence within a simulation (e. g., MC-LR and [β-D-Asp^3,(*E*)-Dhb^7]MC-RR, [β-D-Asp^3,(*E*)-Dhb^7]MC-RR and MC-LY(Prg)). Also, MC-LF and [Apha^5]MC-LF shared a number of similar IFPs (14.10 %), and a small area within the MC-LF region did map to a larger area in [Apha^5]MC-LF simulation. Therefore, both MC congeners appeared to share some binding patterns, as they are also structurally very similar, but both of the MC congeners had their unique interaction patterns. As similarity may be dependent on simulation time and repetition, longer or higher numbers of simulations would also refine the grouping into identical, similar or dissimilar IFPs.

For a high number of comparisons of MD simulations carried out, the number of similar IFPs achieved was between 0.16 and 5.65 %, which is rather low. These MD simulations either provided for a low number of interactions thereby leading to the same effect as described above, or a medium number of interactions, which is in general lower than their respective mean value. In addition, the most common IFPs tended to

occur infrequently within a given simulation, such that the similar IFPs did not represent a frequent interaction pattern and were therefore likely to be less relevant (Figs. S19, S20, S22, S27, S28, S29, S32, and S33).

MC-RR and [Apda^{5}]MC-LF were the only MC congeners that did not share any similar IFP with other simulations. For MC-RR, the conformation identified when bound to PPP1 was unique and could not be found for any of the other MC congeners. Also, the conformation of [Apda^{5}]MC-LF when bound to PPP1 was relatively planar, in comparison to MC-LF, even though only side chains of Adda were removed. This seems to cause a conformational change and also a change in interactions, leading to unique IFPs.

In summary, we were able to show that changes in interactions accumulate over time and that for some MC congeners frequently occurring IFPs were suggested by modelling that can help to identify important interaction patterns. The analysis of the three most frequent IFPs per MC congener showed that the interactions between at least two of the three were similar for all MC congeners, and that residues, known to interact from Fontanillo and Köhn [62]), never occur simultaneously. In addition, the analysis of the average of the interactions per individual IFP showed that toxic MC congeners had a higher mean value than their respective structurally similar moderately or non-toxic MC congeners, and that the Adda-side chain modification may influence the average value. Consequently, average interactions per individual IFP allows to distinguish toxic from moderately or non-toxic MC congeners and thus may provide a better basis for the toxicodynamic part of risk assessment. Indeed, being able to distinguish problematic (containing predominantly toxic MC congeners) from non-problematic (containing predominantly moderately or non-toxic MC congeners) cyanobacterial blooms based on an improved RA model may have major health and economic implications in the near future [98].

4. Conclusion

We carried out MD simulations for twelve MC congeners, eight of which were simulated for the first time. We could show that the PPP1 structure is stable upon MC congener binding, and that size, compactness and solvent accessibility of MC congeners were dependent on their structural modification.

The conformational analysis of MC congener backbones suggested that MC congeners partly adopted the saddle-shaped conformation previously reported for MC-LR, but that MC-congener-specific conformations (i.e., planar, inverted U-shape, V-shape or boat conformation) also exist. In addition, MC congeners can in some cases adopt different conformations in solvent and complex. Modification of the Adda-side chain can influence the backbone conformation as well. Therefore, we conclude that the backbone conformation influences the stability of the PPP1 binding but is unlikely to be the sole reason for binding.

Analysis of the IFP patterns of MC congeners led to the conclusion that for individual MC congeners, frequently occurring interaction patterns are present that could reproduce interactions with residues of PPP1 described previously, but never appear to occur simultaneously. Non-toxic congeners had a lower number of interactions per IFP compared to their respective structurally similar toxic MC congeners. In addition, comparison across MD simulations showed similarities in binding patterns between some MC congeners, but overall MC congeners also had their own unique binding patterns. These patterns were not necessarily shared with other MC congeners, although we were able to identify common IFPs representatives and residues common to many IFP of different MC congeners.

We therefore conclude that MD simulations can aid in understanding the dynamics of MC congener binding to PPP1, as well as the conformational space of individual MC congeners. The simulations presented here indicate that a congener-specific investigation is necessary as, in contrast to most previous publications, MC congeners do not adopt a single conformation. MD simulations could also help to investigate other MC congener structures that so far have not been characterised, and therefore provide for an important step towards a robust MC congener-specific risk assessment. In addition, MD simulations with other protein phosphatases (e.g. PPP2A) may help to understand the selectivity profiles. Thus, the protocol presented here can serve as a starting point for further simulations with more MC congeners or different protein phosphatases to develop strategies to even better discriminate MC congener toxicity.

CRediT authorship contribution statement

Sabrina Jaeger-Honz: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Raymund Hackett:** Writing – review & editing, Software, Methodology, Formal analysis. **Regina Fotler:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Daniel R. Dietrich:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Data curation, Conceptualization. **Falk Schreiber:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Availability of data and materials

The data of the MD simulation as well as the structural data of the PCA analysis of the MC congeners are available at Zenodo [99]. The remaining analysis and scripts supporting the conclusions of this article are available upon request.

Funding

This work was supported by the DFG (Project-ID 251654672 – TRR 161), and the Arthur und Aenne Feindt Foundation, Hamburg, Germany.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank SilcsBio, LLC for providing access to CGenFF program version 2.4.0 for use with CGenFF version 4.4.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cbi.2025.111372.

Data availability

Data will be made available on request.

References

[1] E.M. Driggers, S.P. Hale, J. Lee, N.K. Terrett, The exploration of macrocycles for drug discovery-an underexploited structural class, Nat. Rev. Drug Discov. 7 (2008) 608–624.
[2] D.R. Dietrich, S. Hoeger, Guidance values for microcystins in water and cyanobacterial supplement products (blue-green algal supplements): a reasonable or misguided approach? Toxicol. Appl. Pharmacol. 203 (2005) 273–289.
[3] S.M. Azevedo, W.W. Carmichael, E.M. Jochimsen, K.L. Rinehart, S. Lau, G.R. Shaw, G.K. Eaglesham, Human intoxication by microcystins during renal dialysis treatment in Caruaru-Brazil, Toxicology 181–182 (2002) 441–446.
[4] Y. Grosse, R. Baan, K. Straif, B. Secretan, F. El Ghissassi, V. Cogliano, W.H.O.I.A.f. R.o.C.M.W. Group, Carcinogenicity of nitrate, nitrite, and cyanobacterial peptide toxins, Lancet Oncol. 7 (2006) 628–629.

[5] S. Pouria, A. de Andrade, J. Barbosa, R.L. Cavalcanti, V.T. Barreto, C.J. Ward, W. Preiser, G.K. Poon, G.H. Neild, G.A. Codd, Fatal microcystin intoxication in haemodialysis unit in Caruaru, Brazil, Lancet 352 (1998) 21–26.

[6] M. Yuan, W.W. Carmichael, E.D. Hilborn, Microcystin analysis in human sera and liver from human fatalities in Caruaru, Brazil 1996, Toxicon 48 (2006) 627–640.

[7] N. Bouaicha, C.O. Miles, D.G. Beach, Z. Labidi, A. Djabri, N.Y. Benayache, T. Nguyen-Quang, Structural Diversity, Characterization and Toxicology of Microcystins, Toxins (Basel) 11 (2019).

[8] W.W. Carmichael, V. Beasley, D.L. Bunner, J.N. Eloff, I. Falconer, P. Gorham, K. Harada, T. Krishnamurthy, M.J. Yu, R.E. Moore, et al., Naming of cyclic heptapeptide toxins of cyanobacteria (blue-green algae), Toxicon 26 (1988) 971–973.

[9] K.L. Rinehart, K. Harada, M. Namikoshi, C. Chen, C.A. Harvis, M.H.G. Munro, J. W. Blunt, P.E. Mulligan, V.R. Beasley, et al., Nodularin, microcystin, and the configuration of Adda, J. Am. Chem. Soc. 110 (1988) 8557–8558.

[10] S. Altaner, S. Jaeger, R. Fotler, I. Zemskov, V. Wittmann, F. Schreiber, D. R. Dietrich, Machine learning prediction of cyanobacterial toxin (microcystin) toxicodynamics in humans, ALTEX 37 (2020) 24–36.

[11] J. Goldberg, H.B. Huang, Y.G. Kwon, P. Greengard, A.C. Nairn, J. Kuriyan, Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1, Nature 376 (1995) 745–753.

[12] C. MacKintosh, K.A. Beattie, S. Klumpp, P. Cohen, G.A. Codd, Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1 and 2A from both mammals and higher plants, FEBS (Fed. Eur. Biochem. Soc.) Lett. 264 (1990) 187–192.

[13] R.W. MacKintosh, K.N. Dalby, D.G. Campbell, P.T. Cohen, P. Cohen, C. MacKintosh, The cyanobacterial toxin microcystin binds covalently to cysteine-273 on protein phosphatase 1, FEBS (Fed. Eur. Biochem. Soc.) Lett. 371 (1995) 236–240.

[14] A. Fischer, S.J. Hoeger, K. Stemmer, D.J. Feurstein, D. Knobeloch, A. Nussler, D. R. Dietrich, The role of organic anion transporting polypeptides (OATPs/SLCOs) in the toxicity of different microcystin congeners in vitro: a comparison of primary human hepatocytes and OATP-transfected HEK293 cells, Toxicol. Appl. Pharmacol. 245 (2010) 9–20.

[15] W.J. Fischer, S. Altheimer, V. Cattori, P.J. Meier, D.R. Dietrich, B. Hagenbuch, Organic anion transporting polypeptides expressed in liver and brain mediate uptake of microcystin, Toxicol. Appl. Pharmacol. 203 (2005) 257–263.

[16] R. Fotler, Toxicological Profile of Microcystins: Determination of the Toxicokinetic and Toxicodynamic Mechanisms of Microcystins in Human Cells, University of Konstanz, 2024. http://nbn-resolving.de/urn:nbn:de:bsz:352-2-28s9cke9kuei6Konstanz.

[17] D. Feurstein, J. Kleinteich, A.H. Heussner, K. Stemmer, D.R. Dietrich, Investigation of microcystin congener–dependent uptake into primary murine neurons, Environ. Health Perspect. 118 (2010) 1370–1375.

[18] WHO, Guidelines for drinking-water quality, in: Incorporating the First and Second Addenda, fourth ed., 2022. Geneva.

[19] D. Garibo, C. Flores, X. Ceto, B. Prieto-Simon, M. Del Valle, J. Caixach, J. Diogene, M. Campas, Inhibition equivalency factors for microcystin variants in recombinant and wild-type protein phosphatase 1 and 2A assays, Environ. Sci. Pollut. Res. Int. 21 (2014) 10652–10660.

[20] S.J. Hoeger, D. Schmid, J.F. Blom, B. Ernst, D.R. Dietrich, Analytical and functional characterization of microcystins [Asp3]MC-RR and [Asp3,Dhb7]MC-RR: consequences for risk assessment? Environ. Sci. Technol. 41 (2007) 2609–2616.

[21] J.K. Fawell, R.E. Mitchell, D.J. Everett, R.E. Hill, The toxicity of cyanobacterial toxins in the mouse: I Microcystin-LR, Hum. Exp. Toxicol. 18 (1999) 162–167.

[22] Y. Shi, Serine/threonine phosphatases: mechanism through structure, Cell 139 (2009) 468–484.

[23] P.T.W. Cohen, Overview of protein serine/threonine phosphatases, in: J.n. Ariño, D.R. Alexander (Eds.), Protein Phosphatases, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, pp. 1–20.

[24] I. Zemskov, S. Altaner, D.R. Dietrich, V. Wittmann, Total synthesis of microcystin-LF and derivatives thereof, J. Org. Chem. 82 (2017) 3680–3691.

[25] Y. Chen, J. Kirchmair, Cheminformatics in natural product-based drug discovery, Mol. Inform. 39 (2020) e2000171.

[26] A. Hospital, J.R. Goni, M. Orozco, J.L. Gelpi, Molecular dynamics simulations: advances and applications, Adv. Appl. Bioinform. Chem. 8 (2015) 37–47.

[27] S.R. Pereira, V.M. Vasconcelos, A. Antunes, Computational study of the covalent bonding of microcystins to cysteine residues-a reaction involved in the inhibition of the PPP family of protein phosphatases, FEBS J. 280 (2013) 674–680.

[28] A.L. Pochodylo, T.G. Aoki, L. Aristilde, Adsorption mechanisms of microcystin variant conformations at water-mineral interfaces: a molecular modeling investigation, J. Colloid Interface Sci. 480 (2016) 166–174.

[29] A.L. Pochodylo, A.R. Klein, L. Aristilde, Metal-binding selectivity and coordination dynamics for cyanobacterial microcystins with Zn, Cu, Fe, Mg, and Ca, Environ. Chem. Lett. 15 (2017) 695–701.

[30] Z. Hu, X. Wang, S. Zhang, W. Zong, Research on the discrepant inhibition mechanism of microcystin-LR disinfectant by-products target to protein phosphatase 1, Environ. Sci. Pollut. Res. Int. 28 (2021) 45586–45595.

[31] W. Zong, Q. Wang, S. Zhang, Y. Teng, Y. Du, Regulation on the toxicity of microcystin-LR target to protein phosphatase 1 by biotransformation pathway: effectiveness and mechanism, Environ. Sci. Pollut. Res. Int. 25 (2018) 26020–26029.

[32] W. Zong, X. Wang, Y. Du, S. Zhang, Y. Zhang, Y. Teng, Molecular mechanism for the regulation of microcystin toxicity to protein phosphatase 1 by glutathione conjugation pathway, BioMed Res. Int. (2017) 9676504.

[33] W.S. Zong, S.H. Zhang, Q. Wang, Y. Teng, Y.Z. Liu, Y.G. Du, Evaluation of the direct and indirect regulation pathways of glutathione target to the hepatotoxicity of microcystin-LR, BioMed Res. Int. (2018) 5672637.

[34] H. Yu, Y. Xu, J. Cui, W. Zong, Mechanism for the potential inhibition effect of microcystin-LR disinfectant by-products on protein phosphatase 2A, Toxins (Basel) 14 (2022).

[35] Y. Liu, S. Liu, C. Xu, M. Lin, Y. Li, C. Shen, Y. Liang, X. Sun, D. Wang, P. Lu, X. Liu, Epitopes prediction for microcystin-LR by molecular docking, Ecotoxicol. Environ. Saf. 227 (2021) 112925.

[36] D.A. McPartlin, C. Murphy, J. Fitzgerald, H. Ma, F. Regan, R.J. O'Kennedy, Understanding microcystin-LR antibody binding interactions using in silico docking and in vitro mutagenesis, Protein Eng. Des. Sel. 32 (2019) 533–542.

[37] T. Lanaras, C.M. Cook, J.E. Eriksson, J.A. Meriluoto, M. Hotokka, Computer modelling of the 3-dimensional structures of the cyanobacterial hepatotoxins microcystin-LR and nodularin, Toxicon 29 (1991) 901–906.

[38] C. Taylor, R.J. Quinn, R. McCullock, R. Nishiwaki-Matsushima, H. Fujiki, An alternative computer model of the 3-dimensional structural of microcystin-LR and nodularin rationalising their interactions with protein phosphatases 1 and 2A, Bioorg. Med. Chem. Lett 2 (1992) 299–302.

[39] S. Rudolph-Bohner, D.F. Mierke, L. Moroder, Molecular structure of the cyanobacterial tumor-promoting microcystins, FEBS Lett. 349 (1994) 319–323.

[40] J.R. Bagu, F.D. Sonnichsen, D. Williams, R.J. Andersen, B.D. Sykes, C.F. Holmes, Comparison of the solution structures of microcystin-LR and motuporin, Nat. Struct. Biol. 2 (1995) 114–116.

[41] G.B. Trogen, A. Annila, J. Eriksson, M. Kontteli, J. Meriluoto, I. Sethson, J. Zdunek, U. Edlund, Conformational studies of microcystin-LR using NMR spectroscopy and molecular dynamics calculations, Biochemistry 35 (1996) 3197–3205.

[42] J.R. Bagu, B.D. Sykes, M.M. Craig, C.F. Holmes, A molecular basis for different interactions of marine toxins with protein phosphatase-1. Molecular models for bound motuporin, microcystins, okadaic acid, and calyculin A, J. Biol. Chem. 272 (1997) 5087–5097.

[43] P. Lavigne, J.R. Bagu, R. Boyko, L. Willard, C.F. Holmes, B.D. Sykes, Structure-based thermodynamic analysis of the dissociation of protein phosphatase-1 catalytic subunit and microcystin-LR docked complexes, Protein Sci. 9 (2000) 252–264.

[44] K. Mattila, A. Annila, T.T. Rantala, Metal Ions Mediate the Binding of Cyanobacterial Toxins to Human Protein Phosphatase 1, University of Oulu, Oulu, 2000.

[45] S. Jaeger-Honz, J. Nitschke, S. Altaner, K. Klein, D.R. Dietrich, F. Schreiber, Investigation of microcystin conformation and binding towards PPP1 by molecular dynamics simulation, Chem. Biol. Interact. 351 (2022) 109766.

[46] S. Jaeger-Honz, K. Klein, F. Schreiber, Systematic analysis, aggregation and visualisation of interaction fingerprints for molecular dynamics simulation data, J. Cheminf. 16 (2024) 28.

[47] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, J. Cheminf. 4 (2012) 17.

[48] P. Benkert, M. Biasini, T. Schwede, Toward the estimation of the absolute quality of individual protein structure models, Bioinformatics 27 (2011) 343–350.

[49] M. Bertoni, F. Kiefer, M. Biasini, L. Bordoli, T. Schwede, Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology, Sci. Rep. 7 (2017) 10480.

[50] S. Bienert, A. Waterhouse, T.A. de Beer, G. Tauriello, G. Studer, L. Bordoli, T. Schwede, The SWISS-MODEL Repository-new features and functionality, Nucleic Acids Res. 45 (2017) D313–D319.

[51] N. Guex, M.C. Peitsch, T. Schwede, Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective, Electrophoresis 30 (Suppl 1) (2009) S162–S173.

[52] G. Studer, C. Rempfer, A.M. Waterhouse, R. Gumienny, J. Haas, T. Schwede, QMEANDisCo-distance constraints applied on model quality estimation, Bioinformatics 36 (2020) 1765–1771.

[53] G. Studer, C. Rempfer, A.M. Waterhouse, R. Gumienny, J. Haas, T. Schwede, QMEANDisCo-distance constraints applied on model quality estimation, Bioinformatics 36 (2020) 2647.

[54] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303.

[55] V.B. Chen, W.B. Arendall 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G. J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta. Crystallogr. D. Biol. Crystallogr. 66 (2010) 12–21.

[56] C.J. Williams, J.J. Headd, N.W. Moriarty, M.G. Prisant, L.L. Videau, L.N. Deis, V. Verma, D.A. Keedy, B.J. Hintze, V.B. Chen, S. Jain, S.M. Lewis, W. B. Arendall 3rd, J. Snoeyink, P.D. Adams, S.C. Lovell, J.S. Richardson, D. C. Richardson, MolProbity: more and better reference data for improved all-atom structure validation, Protein Sci. 27 (2018) 293–315.

[57] L.L.C. Schrödinger, Pymol: the PyMOL Molecular Graphics System, 2018, version 2.4.0.

[58] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera-a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.

[59] GROMACS, Gromacs User Discussions, 2022.

[60] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.

[61] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A. J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791.
[62] M. Fontanillo, M. Kohn, Microcystins: synthesis and structure-activity relationship studies toward PP1 and PP2A, Bioorg. Med. Chem. 26 (2018) 1118–1126.
[63] M. Fontanillo, I. Zemskov, M. Häfner, U. Uhrig, F. Salvi, B. Simon, V. Wittmann, M. Köhn, Synthesis of highly selective submicromolar microcystin-based inhibitors of protein phosphatase (PP)2A over PP1, Angew. Chem. 128 (2016) 14191–14195.
[64] D. Merckel, Docker: lightweight linux containers for consistent development and deployment, Linux J. 2 (2014).
[65] NVIDIA, Gromacs 2020.2 Container, 2022.
[66] R.B. Best, X. Zhu, J. Shim, P.E. Lopes, J. Mittal, M. Feig, A.D. Mackerell Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles, J. Chem. Theor. Comput. 8 (2012) 3257–3273.
[67] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmuller, A.D. MacKerell Jr., CHARMM36m: an improved force field for folded and intrinsically disordered proteins, Nat. Methods 14 (2017) 71–73.
[68] I. Soteras Gutierrez, F.Y. Lin, K. Vanommeslaeghe, J.A. Lemkul, K.A. Armacost, C. L. Brooks 3rd, A.D. MacKerell Jr., Parametrization of halogen bonds in the CHARMM general force field: improved treatment of ligand-protein interactions, Bioorg. Med. Chem. 24 (2016) 4812–4825.
[69] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D. Mackerell Jr., CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, J. Comput. Chem. 31 (2010) 671–690.
[70] K. Vanommeslaeghe, A.D. MacKerell Jr., Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing, J. Chem. Inf. Model. 52 (2012) 3144–3154.
[71] K. Vanommeslaeghe, E.P. Raman, A.D. MacKerell Jr., Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges, J. Chem. Inf. Model. 52 (2012) 3155–3168.
[72] W. Yu, X. He, K. Vanommeslaeghe, A.D. MacKerell Jr., Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations, J. Comput. Chem. 33 (2012) 2451–2468.
[73] C.W. Bock, A.K. Katz, G.D. Markham, J.P. Glusker, Manganese as a replacement for magnesium and zinc: functional comparison of the divalent ions, J. Am. Chem. Soc. 121 (1999) 7360–7372.
[74] O. Allner, L. Nilsson, A. Villa, Magnesium ion-water coordination and exchange in biomolecular simulations, J. Chem. Theor. Comput. 8 (2012) 1493–1502.
[75] J.A. Lemkul, Script to Sort Atoms and Bonds in Ascending Order, 2022.
[76] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[77] S. Páll, B. Hess, A flexible algorithm for calculating pair interactions on SIMD architectures, Comput. Phys. Commun. 184 (2013) 2641–2650.
[78] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[79] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[80] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101.
[81] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[82] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
[83] M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, the GROMACS development team, GROMACS User Manual Version 2016, 2018.
[84] S.M. Law, Modevectors, 2024.
[85] S. Jaeger, S. Fulle, S. Turk, Mol2vec: unsupervised machine learning approach with chemical intuition, J. Chem. Inf. Model. 58 (2018) 27–35.
[86] S. van der Walt, J.L. Schonberger, J. Nunez-Iglesias, F. Boulogne, J.D. Warner, N. Yager, E. Gouillart, T. Yu, c. scikit-image, scikit-image: image processing in Python, PeerJ 2 (2014) e453.
[87] D.J.F.L. van Rossum G., Python Reference Manual, Centrum voor Wiskunde en Informatica, Amsterdam, 1995.
[88] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. Del Rio, M. Wiebe, P. Peterson, P. Gerard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature 585 (2020) 357–362.
[89] R. Biehl, Jscatter, a program for evaluation and analysis of experimental data, PLoS One 14 (2019) e0218789.
[90] J.D. Hunter, Matplotlib: a 2D graphics environment, Comput. Sci. Eng. 9 (2007) 90–95.
[91] R.J. Gowers, M. Linke, J. Barnoud, T. Reddy, M.N. Melo, S.L. Seyler, J.J. Domanski, D.L. Dotson, S. Buchoux, I.M. Kenney, O. Beckstein, MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations, Proc. 15th Python Sci. Conf. (2016) 98–104.
[92] N. Michaud-Agrawal, E.J. Denning, T.B. Woolf, O. Beckstein, MDAnalysis: a toolkit for the analysis of molecular dynamics simulations, J. Comput. Chem. 32 (2011) 2319–2327.
[93] C. Bouysset, S. Fiorucci, ProLIF: a library to encode molecular interactions as fingerprints, J. Cheminf. 13 (2021) 72.
[94] G. Landrum, Rdkit: Open-Source Cheminformatics software,Release 2021.03.5, 2021.
[95] G.G. Maisuradze, A. Liwo, H.A. Scheraga, Principal component analysis for protein folding dynamics, J. Mol. Biol. 385 (2009) 312–329.
[96] M.P. Egloff, P.T. Cohen, P. Reinemer, D. Barford, Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate, J. Mol. Biol. 254 (1995) 942–959.
[97] A.J. Zhang, G. Bai, S. Deans-Zirattu, M.F. Browner, E.Y. Lee, Expression of the catalytic subunit of phosphorylase phosphatase (protein phosphatase-1) in Escherichia coli, J. Biol. Chem. 267 (1992) 1484–1490.
[98] D.P. Hamilton, S.A. Wood, D.R. Dietrich, J. Puddick, Costs of Harmful Blooms of Freshwater Cyanobacteria, Cyanobacteria, 2014, pp. 245–256.
[99] S. Jaeger-Honz, R.E. Hackett, R. Fotler, D.R. Dietrich, F. Schreiber, Molecular Dynamics Simulations of 12 Microcystin Congeners in Solvent and Complex with PPP1, Zenodo, 2024.