Computational and Structural Biotechnology Journal 23 (2024) 1117–1128

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

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

Research Article

Cyclodextrins: Establishing building blocks for AI-driven drug design by determining affinity constants *in silico*

Amelia Anderson [a,b,c,*], Ángel Piñeiro [c], Rebeca García-Fandiño [b], Matthew S. O'Connor [a]

[a] Cyclarity Therapeutics, 8001 Redwood Blvd, Novato, CA 94945, USA
[b] Department of Organic Chemistry, Center for Research in Biological Chemistry and Molecular Materials, Santiago de Compostela University, CIQUS, Spain
[c] Soft Matter & Molecular Biophysics Group, Department of Applied Physics, Faculty of Physics, University of Santiago de Compostela, Spain

ARTICLE INFO

Keywords:
Cyclodextrin
Inclusion complex
Affinity constant
Molecular dynamics
Potential of mean force
Metadynamics

ABSTRACT

Cyclodextrins (CDs) are cyclic carbohydrate polymers that hold significant promise for drug delivery and industrial applications. Their effectiveness depends on their ability to encapsulate target molecules with strong affinity and specificity, but quantifying affinities in these systems accurately is challenging for a variety of reasons. Computational methods represent an exceptional complement to *in vitro* assays because they can be employed for existing and hypothetical molecules, providing high resolution structures in addition to a mechanistic, dynamic, kinetic, and thermodynamic characterization. Here, we employ potential of mean force (PMF) calculations obtained from guided metadynamics simulations to characterize the 1:1 inclusion complexes between four different modified βCDs, with different type, number, and location of substitutions, and two sterol molecules (cholesterol and 7-ketocholesterol). Our methods, validated for reproducibility through four independent repeated simulations per system and different post processing techniques, offer new insights into the formation and stability of CD-sterol inclusion complexes. A systematic distinct orientation preference where the sterol tail projects from the CD's larger face and significant impacts of CD substitutions on binding are observed. Notably, sampling only the CD cavity's wide face during simulations yielded comparable binding energies to full-cavity sampling, but in less time and with reduced statistical uncertainty, suggesting a more efficient approach. Bridging computational methods with complex molecular interactions, our research enables predictive CD designs for diverse applications. Moreover, the high reproducibility, sensitivity, and cost-effectiveness of the studied methods pave the way for extensive studies of massive CD-ligand combinations, enabling AI algorithm training and automated molecular design.

1. Introduction

Cyclodextrins (CDs) are cyclic oligosaccharides, usually formed from 6, 7, or 8 glucose units denoted as αCD, βCD, and γCD, respectively. With hydrophilic exteriors and hydrophobic interiors, CDs can encapsulate hydrophobic molecules in soluble inclusion complexes [1–3]. They can be chemically modified by substituting the hydroxyls with other functional groups, such as methyl, hydroxypropyl (HP), or sulfobutyl (SB). These modifications lead to distributions of structures with different numbers and locations of the substitutions which can improve the solubility of the CD and alter their properties in various ways [1,4–6]. The average number of functional groups in the distribution of chemically modified CDs is denoted as the degree of substitution/saturation (DS). These modifications can also modulate the selectivity for the encapsulation of certain molecules as well as their release rate [7–9], thus opening an opportunity to develop *ad hoc* applications. In pharmaceuticals, CDs are typically employed to enhance the solubility and stability of drugs, improving delivery and bioavailability [2,10–12]. Their applications extend to chemical reactions, environmental uses, food additives, and more [1,13–18]. Ongoing research explores novel applications such as for the treatment of atherosclerosis or removal of PFAS (a group of 'forever' chemicals linked to health problems) from water [19,20], showcasing their versatile potential. CD rings with 7 glucopyranoside units, βCDs, are widely used in industrial and drug delivery settings since their cavity is able to effectively host an interesting variety of molecules. In particular, it is known that these

* Correspondence to: Cyclarity Therapeutics, Center for Research in Biological Chemistry and Molecular Materials, Santiago de Compostela University, CIQUS, Spain.
E-mail address: amelia.anderson@cyclaritytx.com (A. Anderson).

https://doi.org/10.1016/j.csbj.2024.02.011
Received 8 December 2023; Received in revised form 13 February 2024; Accepted 14 February 2024
Available online 16 February 2024

medium-sized CDs form water soluble inclusion complexes with sterol molecules, which are related to a variety of health problems [11,19,21–27]. However, actually measuring the affinity constants between CDs and sterols is difficult and unreliable due to solubility problems, stoichiometric considerations, and the presence of a distribution of CD structures in the case of substituted CDs. Therefore, experimental affinity constants are often reported as "apparent" affinity constants, and there is no accepted experimental value for many of these systems, particularly when the guest is poorly soluble, as cholesterol is [19,28–30]. Here, we aim to establish an effective and efficient tool for obtaining reliable affinity constants for individual CD inclusion complexes using *in silico* methods. If this tool was validated, it could enable the automated production of association constants for many CD inclusion complexes. This would open the door for ML training on a system for which we have very little practical control over *in vitro*, but very precise control over *in silico*. Such a tool could revolutionize our understanding of CD inclusion complexation.

The development and optimization of applications based on CDs depend on their ability to encapsulate molecules into their hydrophobic cavity. For certain uses, such as the extraction of harmful or undesirable molecules from organisms or from the environment, a high affinity constant and a very low release rate are required to achieve high specificity. A notable example is the use of the engineered sugammadex to capture the neuromuscular blocker rocuronium [31,32]. In contrast, employing CDs as excipients necessitates a substantial, though comparatively lower, affinity constant, along with lower specificity and a more flexible release rate. This distinction is critical; using a CD with a very high affinity for an active compound can impede its intended function by preventing its release. Therefore, the accurate quantification of CD-target complexation affinity constants and specificity to distinguish targets/off-targets is key for designing and optimizing CD-based applications.

Historically, this has been achieved through a rational approach, assisted by trial-and-error assays. However, AI algorithms represent a family of promising alternative methods for the same aim; the bottleneck of this approach is the availability of a large volume of high-quality data to effectively train models. Acquiring such data for CD systems is challenging due to several factors, including solubility issues of the compounds involved, the potential presence of impurities, and notably, the diversity in the number and location of substitutions in modified CDs, not to mention the difficulties in correctly analyzing measurements from different sources. Recently, a publicly accessible database that compiles affinity constants for the formation of inclusion complexes between relatively small molecules and CDs, gathered from existing literature, has been published [33]. This information has been obtained from different experimental methods published between the years 1963 and 2021 and includes mainly isothermal titration calorimetry (ITC) and various spectroscopic techniques. This publication represents an extraordinary effort to collect affinity data between CDs and ligands, but the sources are highly heterogeneous and difficult to control, reproduce, or scale. Important uncertainties are also present due to the different sources of the samples, experiments, and analysis methods, complicating analysis. Additionally, the employed techniques face challenges both in conducting effective experiments and in analyzing results. This is because complex protocols are often necessary, including competitive experiments with a reference compound, chemical labeling of molecules, or manipulating the solvent.

It is important to keep in mind the fact that none of these methods actually measure affinity constants directly - instead, they rely on the application of thermodynamic models coupled with numerical and statistical calculations. Best practices involve the global fitting of multiple data sets across different concentration ranges, and the employment of advanced mathematical methods, along with estimating reasonable uncertainty values for numerical parameters. The results typically yield apparent values or approaches useful for specific applications within a limited context, rather than direct thermodynamic affinities. Thus, truly direct laboratory measurements of affinity constants remain an unattainable ideal as of yet, but reasonable estimations are possible in many cases by applying robust multidimensional models to multiple experiments [34]. In this context, massive production of many thousand *in vitro* systems that include a large variety of CDs with specific types, numbers, and locations of substitutions, along with diverse ligands for the purpose of training AI-based algorithms, presents significant challenges in terms of complexity and cost.

The use of computational modeling represents an exceptional and increasingly realistic complement to *in vitro* assays for characterizing the interaction between molecules - particularly between CDs and small molecules that can potentially be encapsulated in their cavity. These methods can provide important insights into the mechanism of interaction between molecules, saving money, time, and resources by predicting how a molecule might act before synthesizing and/or testing it [35]. Additionally, they can be fully automated and connected to a database which would be fed with the results. The information contained in such a database would be highly uniform and it would be very easy to scale the production to reach many thousand complexes. The database would enable precise control of the CD structures, including the type, number, and location of substitutions, and cover a variety of targets, considering the ionic state of the molecules, the presence of cosolvents, or any other environmental factor. These data could then be easily utilized to train AI algorithms. Employing such a workflow could significantly enhance the design and optimization of CDs for specific applications. The bottleneck to developing a platform based on this computational workflow is, again, just as with *in vitro*, the availability of accurate and efficient methods to determine the affinity constant of the encapsulation event. Different computational approaches are widely employed for characterizing protein-ligand interactions in both academic research and practical applications, but their use in the characterization of CD systems is quite limited. A reason for this is the lack of specific software available to automatically perform these studies, since the strategies typically employed to build, execute, and analyze simulations of protein systems differ significantly from those that are more suitable for CDs [36–40]. To take advantage of computational simulations, careful parameterization of the involved molecules, efficient sampling algorithms, and well-designed analysis methods able to provide as much information as possible, are always required. Specifically, while CDs are often depicted as rigid, truncated cone structures, they are actually quite flexible both in solution and in solid form. Capturing this behavior is crucial for accurately describing interactions in computational studies of CD systems [41,42].

A large variety of computational methods able to provide useful information on inclusion complexes is available. Unbiased molecular dynamics simulations (MD) follow the dynamic behavior of an initial conformation of the studied system. This technique is useful to explore the vicinity of the local minima, but the sampling out of this region is not guaranteed. A number of biased methods have been developed to overcome this limitation, also aiming to determine the free energy profile or Potential of Mean Force (PMF). The PMF is a function of a reduced number of collective variables (CVs) able to describe the process of interest, in this case, the formation of inclusion complexes. From PMF profiles it is possible to obtain the most stable structure of the complex, as well as the standard Gibbs energy (ΔG^0) corresponding to its formation. Examining PMF functions reveals the energy barriers and wells that contribute to the interaction mechanism. Metadynamics is a biased molecular dynamics method that allows for the incorporation of specific forces into the system, making it possible to observe events that are typically highly improbable; the likeliness of those events can then be rationalized using PMF, based on the amount of energy necessary for the events to occur in the simulation [43,44]. Metadynamics has already been successfully applied to get PMF profiles of relatively complex CD systems, even considering membranes and higher-order stoichiometries [34,45].

In this study, we have adapted and tested different approaches based

on metadynamics simulations to simulate and evaluate CD-based inclusion complexes. In order to test these methods, 1:1 βCD-sterol systems were employed. Cholesterol (CHOL) and its oxidized form 7-ketocholesterol (7KC) were used as ligands for βCD structures with different type, number and location of substitutions. The position of the substitutions was selected based on steric hindrance and the acidities of the CD OH groups. The difference between 7KC and native CHOL is minuscule: just one additional atom in 7KC, an oxygen forming a keto group at position 7 (Fig. 1). This difference, small as it is, is apparently key to the development of many diseases such as atherosclerosis and Alzheimer's [46]. Thus, unlike cholesterol, which is essential for various physiological roles including cell membrane composition and hormone precursor functions, 7KC has disastrous effects in living organisms [46]. This is the main reason to look for methods to selectively extract 7KC from cells without significantly affecting the concentration of CHOL. Encapsulation of the oxidized sterol by CDs has been proposed for this end [19].

In this strategy, specificity for the encapsulation of 7KC over CHOL is crucial. Experimental evidence obtained from turbidity studies suggests that, in contrast to the native βCD, HPβCD exhibits specific behavior to capture 7KC over CHOL [19,47]. The detailed examination of 1:1 βCD-sterol systems featuring CHOL and 7KC serves as a preliminary model for our broader goal, which is to develop accurate and efficient computational methods that can be implemented in an automated workflow to produce affinity constants for thousands of inclusion complexes. These results enable training machine learning algorithms to design optimal modified CDs for specific purposes.

2. Methods

2.1. Description of the systems

PMF profiles were obtained to characterize the binding of CHOL and 7KC with native βCD, DS3 HPβCD, and SBβCD with DS3 and DS7 (Fig. 1). Regarding the modified CDs, the two DS3 structures, both with HP substitutions and SB substitutions, have these modifications at the exact same positions on the CD molecule.

2.2. System parameters

Both the initial coordinates for the βCD and the corresponding topologies were built using our own tools [19]. Initial coordinates and topology files for the target molecules (CHOL and 7KC) were obtained from the ATB server [48]. To ensure consistency with the force field, manual adjustments were made to the parameters, preventing unreasonable interaction constants for bonds and angles. After the βCD and ligand were built and parameterized, the ligand was introduced into the βCD cavity with its principal axis parallel or antiparallel to the symmetry axis of the CD ("up" and "down" orientations, respectively). The up orientation corresponds to the complex with the tail of the sterol molecule pointing to the large face of the βCD while the down orientation corresponds to the tail of the sterol molecule pointing to the small face (Fig. 2). For each of these complexes, twelve different structures, with the position of the ligand incrementally shifted along the symmetry axis of the CD by steps of 1 or 2 Å, both in the positive and in the negative directions, were built. Each of these structures was inserted in a 7x7x7 nm^3 dodecahedral simulation box, which was then filled with approximately 7800 SPC waters. The water molecules overlapping the solute molecules were removed and Na^+ ions were added to neutralize the total charge, when needed. Simulations at constant temperature (298 K) and pressure (1 bar) using the v-rescale thermostat [49] and the Parrinello-Rahman barostat [50] with coupling times of 0.1 ps and 0.5 ps, respectively, were performed in all cases. The long-range interactions were determined using the PME (Particle Mesh Ewald) [51,52] algorithm with a grid spacing of 0.15 nm and a direct-space cut-off of 1.2 nm. The same cut-off was employed for the short-range interactions. A timestep of 2 fs was employed in all cases using the leap-frog integrator algorithm [53] to integrate the equation of motion. All bonds were constrained using the SETTLE algorithm [54] for water and the LINCS [55] algorithm for the CDs and sterol molecules. These simulations were executed using the GROMACS [56,57] engine using

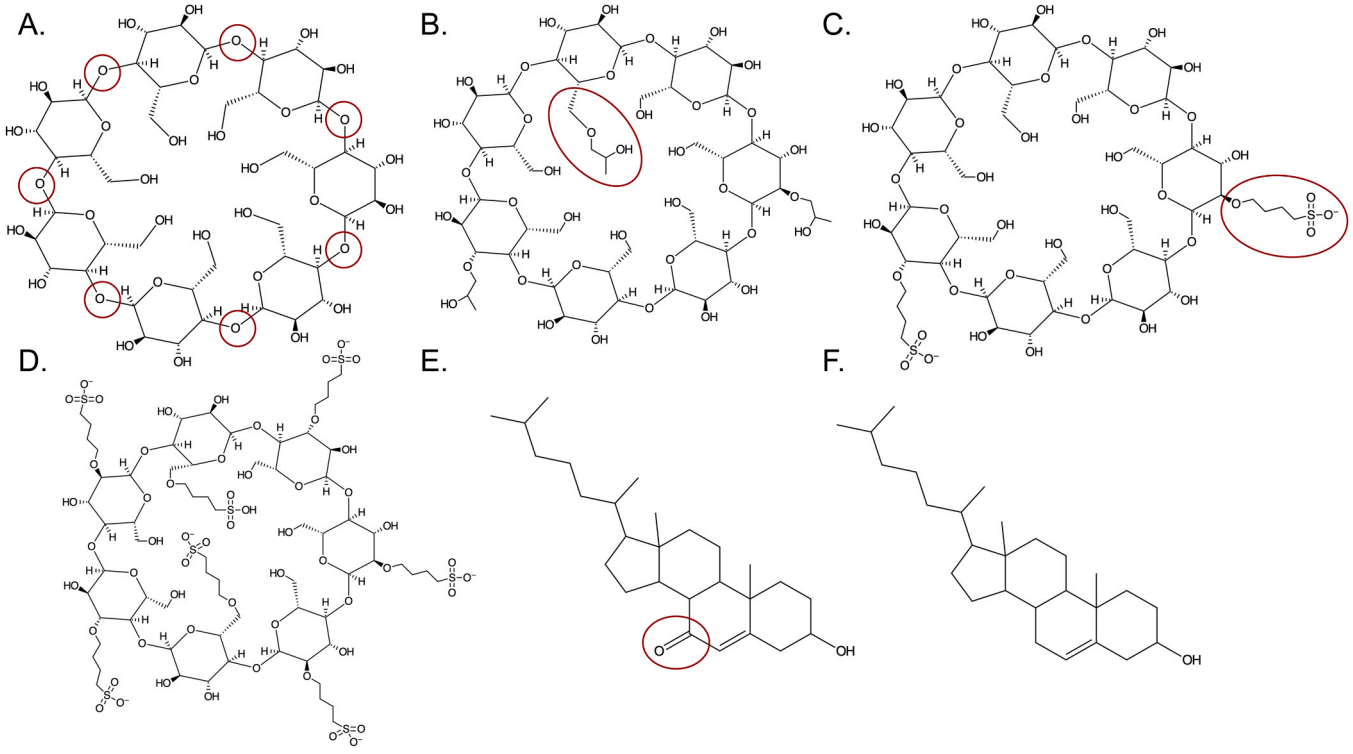

Fig. 1. 2D structures of the molecules to be simulated. (A) Native βCD (O4 atoms circled in red) (B) DS3 HPβCD (one HP group circled in red) (C) DS2 SBβCD (one SBE group circled in red) (D) DS7 SBβCD (E) 7KC (oxygen at position 7 circled in red) (F) Cholesterol.

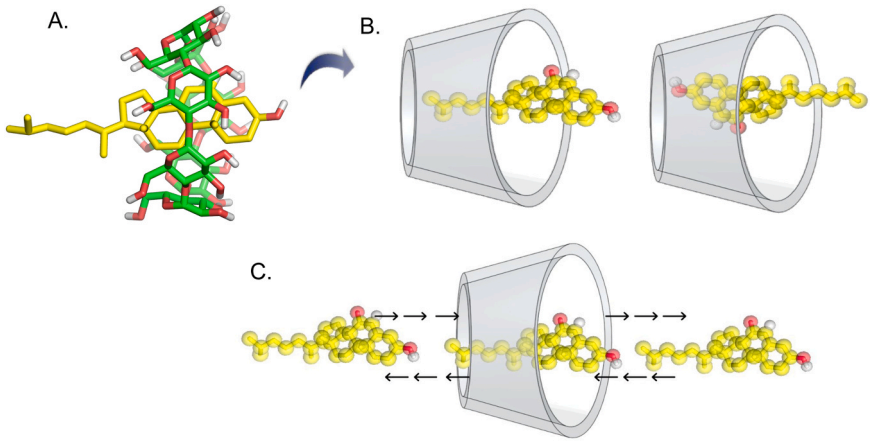

Fig. 2. (A) 3D molecular structure of the bCD (green carbons) - 7KC (yellow carbons) inclusion complex. (B) Simplified cone representation of the bCD (where the smaller side is the primary face and the larger side is the secondary face) showing the up (left) and down (right) orientations of the inclusion complex. The target is shown in yellow. (C) Illustration of metadynamics simulations, where the target is guided through the cavity of the CD in each direction and in each orientation.

the GROMOS54a7 forcefield [58] and PLUMED 2.8. [59–61].

2.3. Metadynamics Simulations

Coupled metadynamics simulations of the twelve slightly different complexes built for each system and relative orientation of the ligand in the CD cavity were executed in parallel. Simulations with the up and down relative orientations of the ligand were conducted separately.

The trajectories were carried out for 300 ns with walls at −2.5 and 2.5 nm between the sterol and the βCD to prevent the sterol from leaving the simulation box and crossing the periodic boundary. The final free energy profile leading to the formation of the complexes, as well as their optimal structure, was obtained by well tempered metadynamics simulations. Many tests were conducted to determine the ideal parameters for the PLUMED metadynamics simulation, and the best are presented here. In these simulations, Gaussian Height = 0.1 kJ/mol, σ = 0.1 nm, biasfactor = 15 at 298 K, and a grid spacing of 0.02 nm between −3.4 and 3.4 nm. The CV was the displacement of the center of mass of the target along the symmetry axis of the CD (ξ) which is calculated from the plane formed by the O4 atoms in the CD (Fig. 1A). The following restraints were devised to focus the investigation on inclusion complex structures: permitting the ligand movement solely along the axis perpendicular to the plane of the O4 atoms of the CD and constraining the relative orientation of the ligand to the up or down orientation (Fig. 2). The O4 atoms of the CD were selected as reference points for distance measurements and defining reference planes and axes in the restraints, due to their relatively static nature compared to the other atoms in the CD. All restraints in these simulations were implemented using harmonic potentials with a constant force of 1000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-2}$. Each metadynamics simulation was repeated four times with slightly different distributions of the starting positions to check for reproducibility.

2.4. Potential of mean force (PMF)

The PMF profiles were determined from the sum of the gaussian functions generated by the 12 walkers corresponding to each relative orientation of the target molecule with each CD during the metadynamics trajectories, using the *sum_hills* tool of PLUMED. The resulting PMF profiles were post-processed to determine the corresponding standard binding constant, which for a bimolecular interaction (A+B ↔ AB) can be determined from the partition function of each chemical species (Z_i):

$$\mathrm{K_A} = \frac{Z_{AB}/V}{(Z_A/V)(Z_B/V)} \tag{1}$$

where V is the volume of the simulation box. Neglecting the internal degrees of freedom of A and B, the corresponding Hamiltonian only accounts for the kinetic energy:

$$H_i = \sum_{j=1}^{3} \frac{p_j^2}{2m_i} \tag{2}$$

Where p_j is the j component of the linear momentum and m_i is the mass of the molecule i. The partition function corresponding to this Hamiltonian is:

$$Z_i = V\left(\frac{2\pi{\cdot}m_i{\cdot}K_BT}{h^2}\right)^{3/2} \tag{3}$$

With K_B representing the Boltzmann constant and h the Plank constant. The Hamiltonian corresponding to the complex would be:

$$H_{AB} = \sum_{j=1}^{3} \frac{\left(p_j^{com}\right)^2}{2(m_A+m_B)} + \sum_{j=1}^{3} \frac{\left(p_j^{R}\right)^2}{2\mu} + E(r) \tag{4}$$

p_j^{com} being the j component of the linear momentum of the center of mass (COM) of the complex, p_j^R the j component of the linear momentum of the relative movement between A and B in the complex, $\mu = \frac{m_A{\cdot}m_B}{(m_A+m_B)}$, and $E(r)$ is the potential energy of interaction between A and B as a function of the relative coordinates between both molecules. The corresponding partition function is:

$$Z_{AB} = V\left(\frac{2\pi \quad (m_A+m_B) \quad K_B \quad T}{h^2}\right)^{3/2} \cdot \left(\frac{2\pi \quad \mu \quad K_B \quad T}{h^2}\right)^{3/2} \int e^{-E(r)/RT} dv \tag{5}$$

where R and T are the gas constant and the temperature, respectively and dv is the volume element expressed in terms of the r coordinates. Replacing Eqs. (2)–(5) in Eq. (1), we obtain:

$$\mathrm{K_A} = \int e^{-E(r)/RT} dv \tag{6}$$

Given the symmetry of the CD monomer and the restraints applied to the relative movement of the ligand with respect to the macromolecule, the interaction energy $E(r)$ and the volume element dv are naturally expressed in cylindrical coordinates (φ, ρ, z). Since the rotations around the CD symmetry axis are not restrained, and the interaction energy is

assumed to be independent on the corresponding angle, the volume element can be directly integrated over φ (from 0 to 2π). The distance from the ligand to the axis perpendicular to the O4 atoms that pass throughout their COM is restrained with a harmonic potential, so it is not exactly constant, but the PMF is assumed to be independent of this parameter. Thus, the volume element is also integrated over ρ from 0 to the maximum radius of the cylinder where the ligand can move (ρ_{max}). Importantly, the simulations were not performed in the standard state, so a correction should be performed for this aim. This correction is performed by dividing the volume of each molecule in the simulation box by the volume available for one molecule at 1 M concentration (1660 Å^3). After simplifying this implies an additional factor of Avogadro's number (N_A) in the expression for K_A. Taking into account all these approaches and considerations, the final expression for the association constant would be:

$$K_A = \pi N_A \rho_{max}^2 \int e^{-E(\xi)/RT} d\xi \quad (7)$$

where the cylindrical coordinate z was replaced by the collective variable ξ employed in our simulations and $E(\xi)$ can be replaced by the PMF. Using this equation, the binding constant for each system was computed by numerical integration of the PMF along the CV. For our calculations, a constant value of 0.4 nm was taken for ρ_{max}. In order to estimate the uncertainty of the K_A values, multiple integrations were considered for each PMF profile, with different integration limits in the unbound region. The values of K_A obtained in this way were then transformed into ΔG^0 by using the classical relationship: $\Delta G^0 = -RT \ln(K_A)$. Then, the ΔG^0 values obtained in this way for each PMF profile were clusterized with the k-means algorithm with the optimal number of clusters determined using the Silhouette [62] method. The final ΔG^0 value, together with the corresponding standard errors were taken as the average and standard deviation of the ΔG^0 values of the most populated cluster. Then, the corresponding association constant was obtained as $K_A = \exp(-\Delta G^0/RT)$ and its uncertainty determined by standard error transmission. This procedure allows us to automatically eliminate outliers without any human bias. For the calculation of average values we decided to work in the ΔG space due to the exponential nature of the association constants, whose direct averaging can be heavily skewed by values that differ significantly in magnitude. However, energies are additive, a property that makes it well suited to operations such as summation and averaging. By converting K_A to ΔG^0 values, one can obtain a more representative average, which can then be transformed back into the K_A space for interpretation. This method ensures that the derived average value is not disproportionately influenced by constants that are significantly larger or smaller than the others, providing a more accurate and meaningful representation of the dataset. This approach resulted in a fully automated PMF method, providing a quantitative approach for obtaining association constants for the formation of 1:1 inclusion complexes using a reasonable amount of computational resources. The final values of K_A and the corresponding ΔG^0 were obtained using a Python script that uses the pandas, [63,64] numPy, [65] matplotlib, [66] and kmeans1d [67,68] libraries.

The PMF calculation was performed two times: once using the entire distribution of the CV (−2 to 2 nm, the "All Method"), and once considering only the wide face of the βCD (CV from 0 to 2 nm only, the "Wide Face Method"). It is expected that the targets will tend to form complexes via the wide face of the βCD, and sampling only half of the βCD will lead to a faster and still accurate result. In addition to association constants, from PMF profiles we can also get the energy wells and barriers of the encapsulation process as a function of the distance between the center of mass of both molecules and the most stable structures for each complex. By using these systems, we hope to gain insight into the effect of each substitution type on the complex.

2.5. Statistical analysis

The ΔG^0 values obtained from the PMF profiles using both integration methods ("All" and "Wide-Face") for the four metadynamics simulations of each CD + sterol combination and for both orientations of each system were plotted to check for reproducibility and systematic differences. A total of 128 ΔG^0 values were collected in this way. Then several statistical tests were applied to check for the normality of the uncertainty distributions (Shapiro-Wilk [69] test) and for significant differences between the values (t-test, Mann-Whitney [70] test, paired t-test, or Wilcoxon [71] test, depending on the case) obtained using different integration methods, considering relative orientations of the ligand in the CD cavity, different ligands for the formation of the complexes, and also to compare the uncertainties obtained using the two integration methods. The average values of the repeats under the same conditions were obtained after removing possible outliers (those with a z-score > 1.5). The standard deviation was taken as a measurement of the reproducibility for these repeats. Bland-Altman [72] plots were also generated to highlight possible trends in the differences of ΔG^0 value obtained using different integration methods, relative orientations, and ligands.

3. Results

3.1. General description of the PMF profiles

PMF profiles were generated for all systems studied by metadynamics simulationswith 300 ns-long trajectories for each of the 12 walkers to ensure simulation convergence. This convergence is demonstrated by the overlapping curves in the PMF profile evolution over time, as shown in Figs. 3–4 and in the Supplementary Material. The simulations sampled the displacement of the target along the symmetry axis of the CD ring, with restraints on the relative orientation and on the distance to this axis (see Methods for details). Consequently, two regions with negligible interaction in the PMF are expected at both negative and positive long-distance displacements between the CD and the ligand. In these "unbound" regions, the PMF profile should level off, indicating no change in interaction energy. This behavior was observed in most systems, with some exceptions showing significant differences in the height of the PMF at opposite sides of the CD. These discrepancies contribute to the uncertainties in the calculated ΔG^0 values. Typically, one or more energy barriers are observed for the ligand entry into the CD cavity from the narrower side, presenting a sampling challenge. Conversely, calculating ΔG^0 values might be simpler when integrating the pathway from the bound to the unbound region via the wider cavity side, usually devoid of significant energy barriers. K_A and corresponding ΔG^0 values for all systems were calculated using these two approaches, as shown in the Supplementary Material.

All employed CDs generally exhibit a marked preference, evidenced by more negative ΔG^0 values, for the down orientation. Specifically, the native βCD monomer shows its lowest energy at exactly 0.748 nm for both sterols in the down orientation, and around 0.4 nm in the up orientation. The positions of these minima are more varied for modified CDs, but they remain fairly consistent for both sterols within each CD and orientation pair, with one notable exception for HPβCD with cholesterol in the up orientation, which shows the absolute minimum at − 1.05 nm and secondary minima at approximately 0.5 and 1 nm. While the PMF profiles for both sterols with a given CD are similar, the presence of two local minima of varying depth for both sterols in the up orientation influences these results. Overall, across all systems, the roughness of the energy profile is much more pronounced in the region corresponding to the narrow face of the CD than the wide face, making the former more challenging to sample than the latter. Additionally, it seems that the larger the number of substitutions the more sophisticated the PMF profile is, with more and higher energy barriers for SBβCD (DS7) than for SBβCD (DS3) and a more complex energy profile for both modified CDs with DS3 than for the native βCD. (Fig. 5).

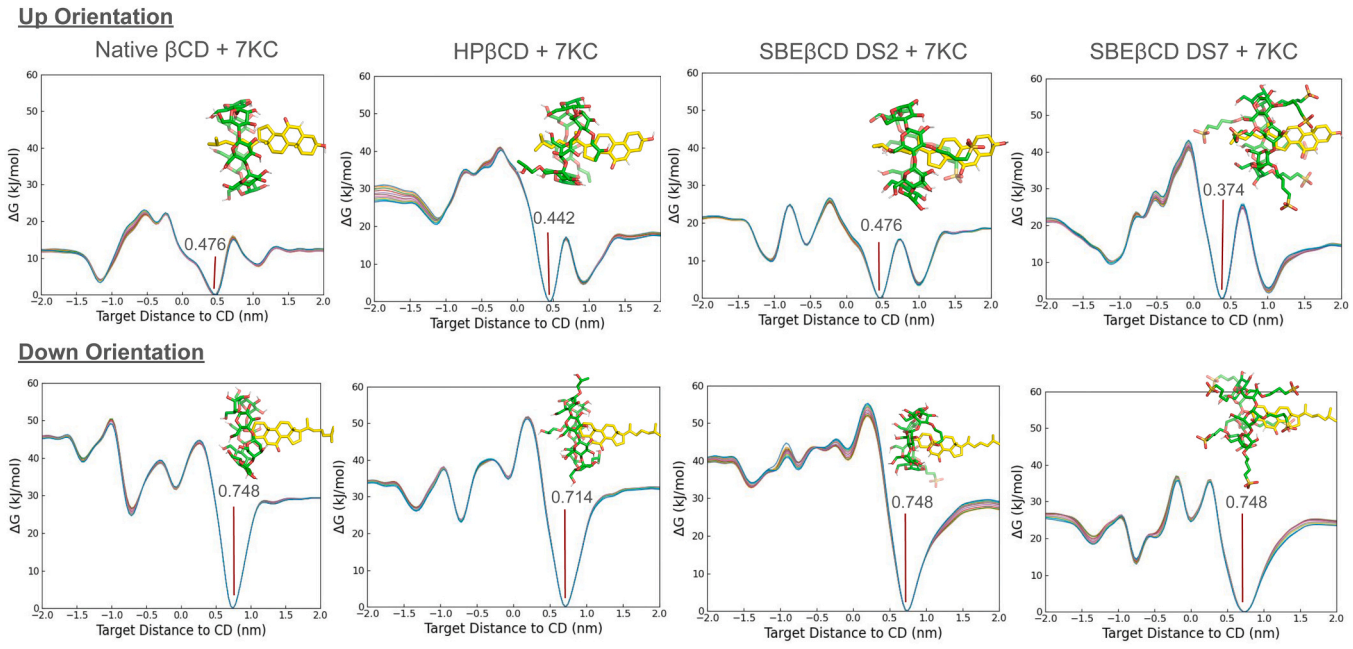

Fig. 3. PMF profiles for the four βCDs with 7KC in the up and down orientations, together with the structures of the complexes and the CV value corresponding to the energy minimum. The K_A and ΔG^0 values calculated from both the All and Wide Face method, as well as the additional simulation replicates, are available in the Supplementary Material.

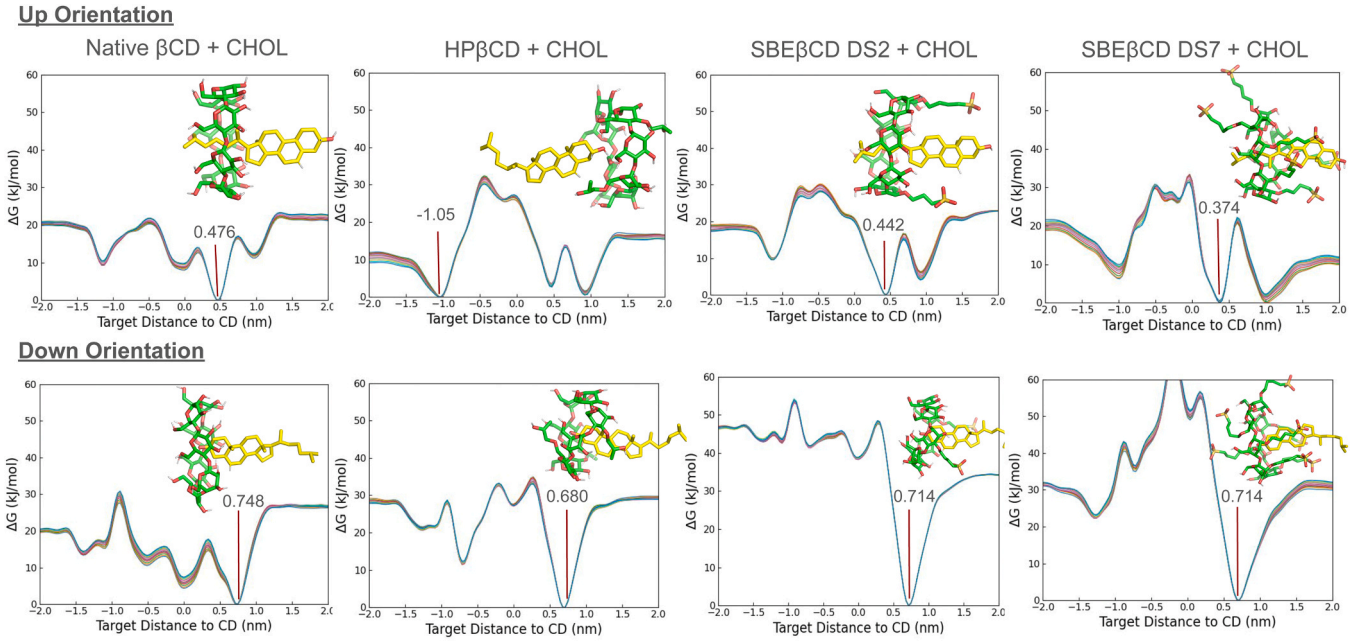

Fig. 4. PMF profiles for the four βCDs with cholesterol in the up and down orientations, together with the structures of the complexes and the CV value corresponding to the energy minimum. The K_A and ΔG^0 values calculated from both the All and Wide Face method, as well as the additional simulation replicates, are available in the Supplementary Material.

To facilitate comparison of results from various combinations of employed variables (relative orientation, integration method, and ligand name), the average values for each specific combination, along with their standard deviations, were calculated (Fig. 6 and Table 1). As mentioned above, the outliers (values with z-score > 1.5) were excluded for these calculations (a total of 11 ΔG values). In general, there is a good correlation observed between the two integration methods. Additionally, clear trends are evident in these results, although some are more subtle. A notable example is the difference between the results for the up and down orientations, with the down orientation consistently showing significantly stronger affinity across all CDs. This trend is reproducible with both integration methods. However, the specificity of the CDs for each sterol is less clear. While the native βCD in the up orientation appears to be more specific for CHOL and HPβCD, also in the up orientation, is more specific for 7KC, the overall impact of these differences is limited due to the much lower significance of the up orientation compared to the down orientation. It is important to note that the overall association constant for a given CD is the sum of the association constants from both the up and down orientations. The association constants are proportional to the exponential of ΔG^0 ($K_A = \exp$

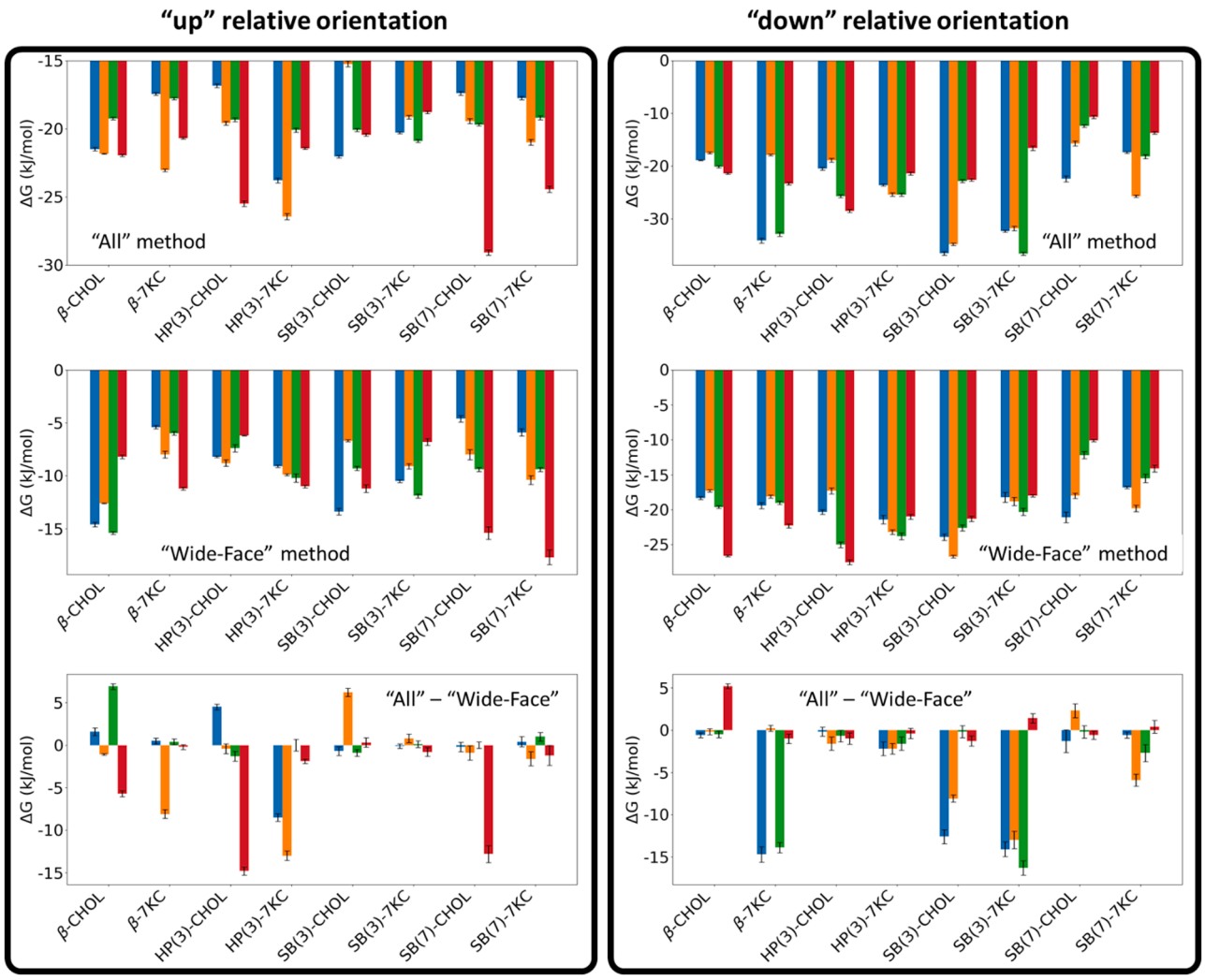

Fig. 5. ΔG^0 values obtained from all the PMF profiles calculated in the present work. For each system and integration method, four independent calculations were conducted, represented by the colors blue, orange, green, and red. The top plot displays results utilizing the entire PMF profile. The middle plot shows the results obtained from the integration from the bound region near the global minimum to the unbound region towards the wide face direction. The bottom plots highlight the differences between the two integration methods for identical simulations. The left column contains the results of simulations where the ligand is in the up orientation, while the results for the down orientation are in the right column.

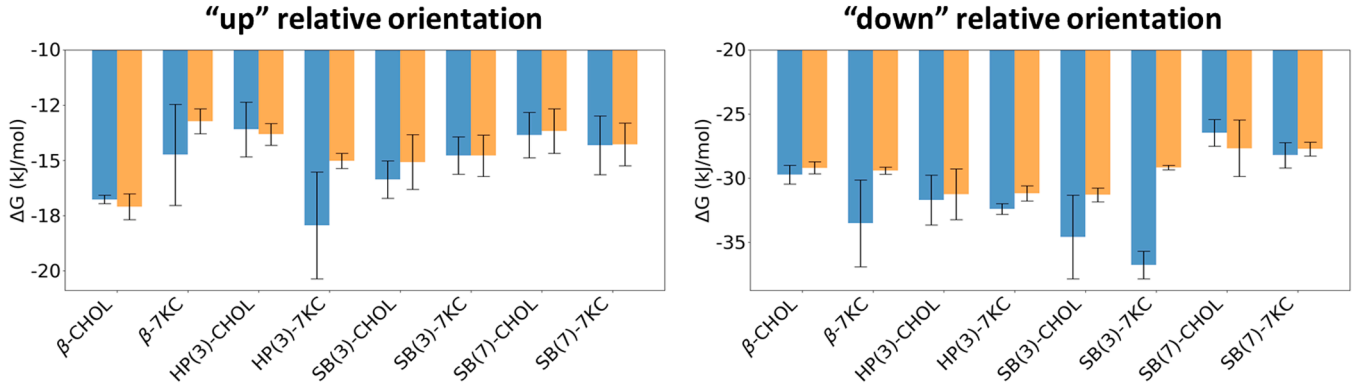

Fig. 6. Average ΔG^0 values for each group of experiments (n = 4) together with the corresponding standard deviation (error bars). The results using the 'All' integration method are in blue and those using the 'Wide-Face' method are in orange. Values with z-score > 1 (a total of 11 out of 128 values) were excluded from the calculations.

$(-\Delta G^0/RT)$), meaning even small differences in ΔG^0 can lead to substantial variations in the association constants. Therefore, in all cases, the contribution of the up orientation to the quantification of binding energies is expected to be low when compared to the down orientation. The results corresponding to the down orientation do not exhibit clear specificity for any of the sterol molecules.

Table 1
Average ΔG^0 values for each group of experiments (n = 4) together with the corresponding standard deviation for the All and the Wide-Face integration methods. Values with z-score > 1 (a total of 11 out of 128 values) were excluded from the calculations.

	up		down	
	all	Wide-Face	all	Wide-Face
β-CHOL	−13.53 ± 0.39	−14.2 ± 1.2	−19.5 ± 1.4	−18.40 ± 0.94
β-7KC	−9.5 ± 4.6	−6.5 ± 1.1	−27.1 ± 6.7	−18.83 ± 0.54
HP(3)-CHOL	−7.2 ± 2.5	−7.7 ± 1.0	−23.4 ± 3.9	−22.5 ± 4.0
HP(3)−7KC	−15.9 ± 4.8	−10.05 ± 0.68	−24.80 ± 0.85	−22.4 ± 1.2
SB(3)-CHOL	−11.7 ± 1.7	−10.2 ± 2.5	−29.2 ± 6.5	−22.6 ± 1.1
SB(3)−7KC	−9.6 ± 1.7	−9.6 ± 1.9	−33.6 ± 2.2	−18.33 ± 0.34
SB(7)-CHOL	−7.7 ± 2.1	−7.3 ± 2.0	−12.9 ± 2.1	−15.4 ± 4.4
SB(7)−7KC	−8.6 ± 2.7	−8.6 ± 1.9	−16.4 ± 2.0	−15.5 ± 1.1

A more robust support for this discussion could be provided by applying appropriate statistical tests.

3.2. Statistical results

3.2.1. All vs. Wide-Face methods

To explore the differences in ΔG^0 values, uncertainties, and reproducibility between the two integration methods (All and Wide-Face) across both up and down orientations, the average values shown in Fig. 6 and in Table 1 were employed. The standard deviation was taken as a measurement of the reproducibility of the ΔG^0 values in each independent simulation.

The Bland-Altman plots suggest a small bias between the two methods (represented by the red line in Fig. 7). In addition to this visual comparison, a quantitative analysis aimed to assess whether or not both methods are equivalent to each other was performed. The analysis was performed in different ways: (*i*) ignoring the uncertainties; and (*ii*) considering the uncertainties to weight the differences in the results obtained from both methods. Note that outliers were excluded from the calculation of the average values. To assess the differences in ΔG^0 values derived from the two integration methods, both the paired t-test and the Wilcoxon test were utilized. Initial comparisons, which did not account for uncertainties, suggested significant differences between the methods. However, when incorporating the standard deviation of the average values into the analysis, there are no significant differences between the two integration methods (p-value > 0.05) (Table 2).

The t-test for independent samples and the Mann-Whitney test were employed to check for differences in the uncertainties of both methods. Both tests suggest that the uncertainties are significantly different with p-values of 0.030 and 0.018, respectively. The mean difference of the standard deviations is 1.83 ± 0.49 (larger for the All method). This means that the data obtained using the Wide-Face method is significantly more reproducible than those using the All method, even though the two methods provide equivalent mean values.

3.2.2. Up vs. down orientation

This analysis was performed using the mean values to determine if significant differences exist between the up and down orientations, independent of the reproducibility of repeated experiments for the same system. It was conducted separately for both the All and Wide-Face integration methods.

The Bland-Altman plots show a clear systematic bias between the results provided by both orientations, with a negative excess in the ΔG^0 value of more than 10 kJ/mol, as shown in Fig. 8. Table 3.

No statistical differences were observed between the uncertainties of the two orientations, the t-test for independent samples, and the Mann-Whitney test providing values of 0.51 and 0.80 for the All method and 0.79 and 0.51 for the Wide-Face method.

Table 2
Comparison of the average difference in ΔG^0 values ($\Delta G^0_{all} - \Delta G^0_{wide}$) obtained from all the PMF plots with different considerations.

Analysis Method	Mean Difference (kJ/mol)	p-value t-test	p-value Wilcoxon
Results using the mean values for each system			
Without Uncertainties	−2.67 ± 1.10	0.028	0.0125
With Uncertainties	−1.96 ± 1.04	0.080	

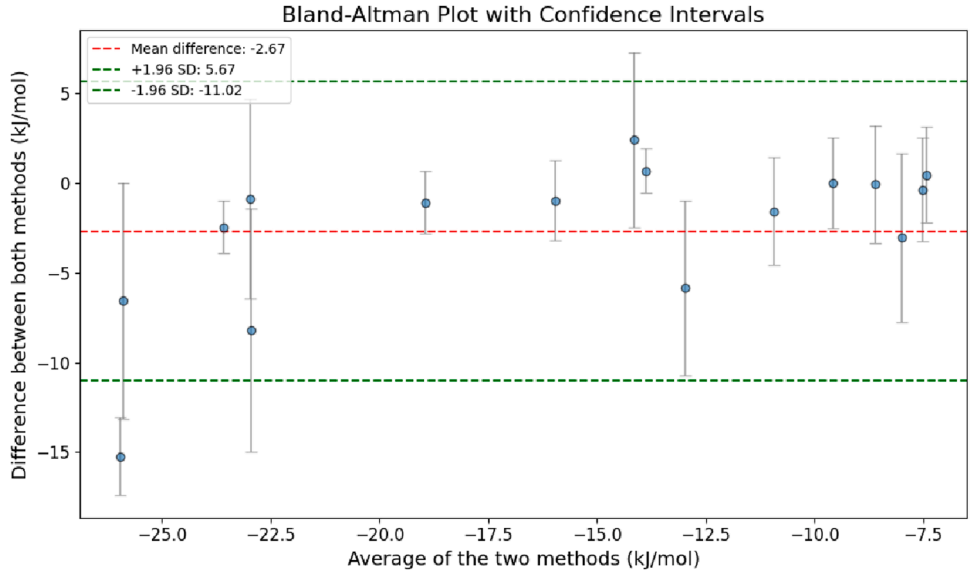

Fig. 7. Differences in ΔG^0 values obtained from the integration of the whole PMF profile and using just the wide face ($\Delta G^0_{All} - \Delta G^0_{Wide\text{-}Face}$) as a function of the average of both values ($0.5(\Delta G^0_{All} + \Delta G^0_{Wide\text{-}Face})$). The mean of the difference together with the confidence interval (mean difference ± 1.96 times the standard deviation of the differences) are shown as a dashed red line and dashed green lines, respectively.

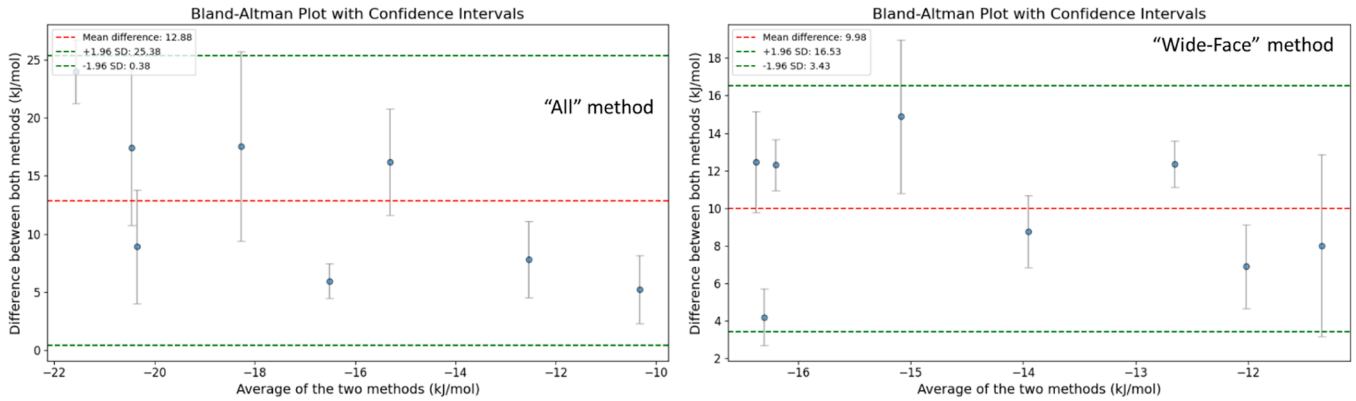

Fig. 8. Differences in mean ΔG^0 values obtained for the up and down orientations of each system for the All (left) and Wide-Face (right) integration methods. The mean of the difference together with the confidence interval (mean difference ± 1.96 times the standard deviation of the differences) are shown as a dashed red line and dashed green lines, respectively.

Table 3
Comparison of the average difference in ΔG^0 values ($\Delta G^0_{up} - \Delta G^0_{down}$) obtained from the mean values of the four PMF replicas obtained for each system and orientation. Both integration methods were employed, providing similar results, as indicated in the Table. The p-values obtained from both the paired t-test and the Wilcoxon tests are lower than 0.05 in all cases, indicating that the difference between ΔG^0_{up} and ΔG^0_{down} is significant, with lower (more negative) values for ΔG^0_{down}.

Analysis Method	Mean Difference (kJ/mol)	p-value t-test	p-value Wilcoxon
Results using the All method with mean values of each system			
Without Uncertainties	12.88 ± 2.41	0.0011	0.00781
With Uncertainties	9.79 ± 2.35	0.004	
Results using Wide Face method with mean values of each system			
Without Uncertainties	9.98 ± 1.26	0.0001	0.00781
With Uncertainties	9.92 ± 118	0.0000	

3.2.3. *Cholesterol vs. 7KC targets*

Again, the same analysis as in the previous section was performed, *i.e.* using the mean values of the four replicas obtained for each system and orientation. Again, the analysis was performed using the two integration methods separately.

The Bland-Altman plots (Fig. 9) suggest systematic bias of different sign, where the All method is positive and the Wide Face method is negative, depending on the employed method to integrate the PMF profiles. The application of statistical tests also provides contradictory results for both integration methods, but with no significant systematic differences between both sterols (Table 4).

Table 4
Comparison of the average difference in ΔG^0 values ($\Delta G_{cholesterol} - \Delta G_{7KC}$) obtained from the mean values of the four PMF replicas obtained for each system and orientation.

Analysis Method	Mean Difference (kJ/mol)	p-value t-test	p-value Wilcoxon
Results using the All method with mean values of each system			
Without Uncertainties	2.54 ± 1.56	0.148	0.195
With Uncertainties	1.02 ± 1.23	0.433	
Results using Wide Face method with mean values of each system			
Without Uncertainties	-1.07 ± 1.17	0.931	0.742
With Uncertainties	-1.37 ± 1.21	0.295	

Additionally, no statistical differences were observed between the uncertainties of the mean values for both sterols, the t-test for independent samples, and the Mann-Whitney test providing values of 0.53 and 0.51 for the All method and 0.073 and 0.16 for the Wide-Face method.

4. Discussion

The present study provides a comprehensive investigation into the inclusion complexes formed between different βCDs (βCD, HPβCD with DS3, and SBβCD with DS3 and DS7) and sterol molecules, specifically CHOL and 7KC, through guided metadynamics simulations followed by PMF calculations. The research aims to establish robust and efficient methods to characterize the interaction between CDs and small ligands.

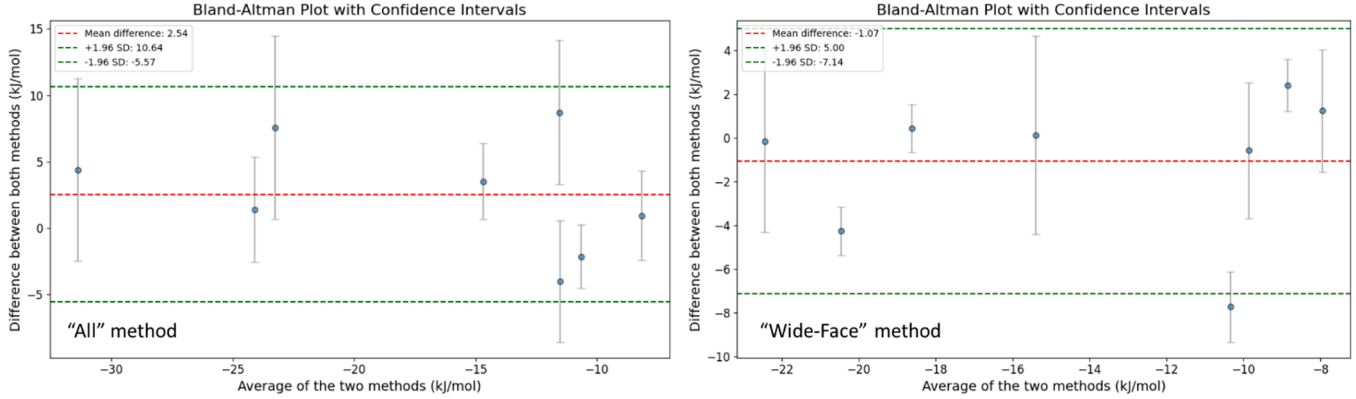

Fig. 9. Differences in mean ΔG^0 values obtained for cholesterol and 7KC with the different CDs for the All (left) and Wide-Face (right) integration methods. The mean of the difference together with the confidence interval (mean difference ± 1.96 times the standard deviation of the differences) are shown as a dashed red line and dashed green lines, respectively.

For this, a relatively large set of trajectories is used to assess the reproducibility in the determination and accuracy of ΔG^0 values by different approaches. Additionally, the ability of the different CDs to distinguish the two sterols and their relative orientation into the cavity are assessed. Complementarily, the obtained results shed light on the interactions between the molecules employed in our calculations.

All employed CDs display a strong preference for the down orientation, indicated by more negative ΔG^0 values by ~10 kJ/mol, compared to the opposite orientation of the sterol in the CD cavity. The positions of all minima (displacement between the CD and the sterol) vary for different CDs, but they are consistent for both sterols within each CD and orientation. SBβCD with DS= 7 is the only exception to this behavior due to the presence of two local minima in the up orientation at similar locations but with different depths for both sterols. This preference could be due to the fact that the hydrophobic headgroup of the sterol is in better contact with the hydrophobic CD cavity, while the hydrophilic tail of the sterol is in better contact with water in this orientation. Subtle differences in sterol specificity among CDs was observed. This could be due to the significant differences in ΔG^0 values in the up orientation, and the impact of this contribution to the overall association constant is limited due to the much higher significance of the down orientation. Note that the total association constant for the interaction between a given CD-ligand pair results from the sum of the association constants in the two orientations and the transformation from the ΔG^0 to the K_A space is performed by the exponential function. Thus, relatively small differences in ΔG^0 are magnified when they are transformed to K_A, as it happens for all the studied systems when combining ΔG^0 values obtained from both orientations of the sterol in the CD cavity.

The roughness of the energy profile is evident around the narrow face of the CD, in contrast to the wide face, making it harder to sample. The results obtained from the two integration methods (All and Wide-Face) were consistent, with no significant differences between them. However, the Wide-Face method was significantly more reproducible and it is cheaper, from the computational point of view, than the All method, since the former requires the sampling of a much smaller range of the CV used in the metadynamics simulations. This method is expected to be sufficient for randomly substituted or unsubstituted CDs interacting with steroidal target molecules.

The study successfully reveals energy profiles, optimized complex structures, and quantitative estimations of binding affinities for the inclusion complexes between βCDs and sterols. The computed PMF profiles offer a detailed understanding of the energy landscapes governing the association between the host βCDs and the guest sterols. The rationalization of these energy landscapes provides insights into the mechanisms of inclusion, shedding light on the interactions and driving forces governing the formation of inclusion complexes. The employed approach to exploring selective sequestration of specific molecules using modified βCDs opens up possibilities for designing novel molecular systems for targeted drug delivery or other applications. Our study is focused only on 1:1 complexes, but it is worth mentioning that other stoichiometries such as 2:1 (two CD molecules threaded by the same ligand), 1:2 (one CD molecule simultaneously binding 2 sterol molecules), or even aggregates of higher order could significantly contribute to the solubilization of a given ligand by CDs. [24,27,73–77] However, inclusion complexes are of special interest because the interaction mechanism leading to these structures is expected to take advantage of the particular design of the CD molecules to specifically distinguish between different ligands. The CDs employed in this study did not successfully differentiate CHOL from 7KC, but it has been demonstrated, with approaches similar to those employed in the present work, that CD dimers, with the monomers joined by a chemical linker, are indeed able to distinguish these two sterol molecules.

Overall, in addition to the new insight that our calculations provide on the specific interactions between different CDs and the two studied sterol molecules, we propose a protocol that could be employed in massive PMF calculations of CD-ligand systems. This protocol is based on metadynamics simulations that only require the sampling in one dimension (the displacement of the ligand along the symmetry axis of the CD) in the direction of the wide face of the cavity with some restraints (the relative orientation of the ligand with respect to the CD and the distance from the ligand to the symmetry axis of the CD). The ΔG^0 and association constant values obtained from these calculations were highly reproducible and allowed us to save significant computational resources compared to other methods. We intend to implement this method into an automated workflow for designing modified CDs tailored for specific applications.

5. Conclusion

This study focuses on the computational study of inclusion complexes between multiple βCDs and sterols, using metadynamics simulations and PMF calculations to understand their interactions. The methods reveal that CDs-sterol inclusion complexes are generally preferred in the "down" orientation, with notable differences in binding affinities and energy profiles among different CDs. These differences highlight the impact of number and type of substitution on the characteristics of CD inclusion complexes. The research demonstrates the consistency and efficiency of the Wide-Face integration method in these calculations, offering insights for designing faster, specialized methods to analyze CD-sterol inclusion complexes *in silico*. These results are significant for various applications, including targeted drug delivery and novel drug design. Finally, the automated techniques presented here showcase the potential of accelerating CD development by creating and using large amounts of data for training AI and ML algorithms. These advancements are poised to not only improve our understanding of CD inclusion complexes, but also open new horizons for a world where tailored, CD-based solutions can be developed more efficiently and effectively than ever before.

CRediT authorship contribution statement

Amelia M. Anderson: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Ángel Piñeiro:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Rebeca García-Fandiño:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Matthew S. O'Connor**: Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of generative AI and AI-assisted technologies in the writing process During the preparation of this work the author(s) used OpenAI 's ChatGPT in order to improve the readability and flow of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: AMA, RGF, AP, and MSO report financial support was provided by Cyclarity Therapeutics. Inc. AMA, RGF, AP, and MSO report a relationship with Cyclarity Therapeutics Inc that. includes: employment, consultation or advisory, equity or stocks, and travel reimbursement. RGF and AP report a relationship with the company MD.USE that includes: employment. and equity or stocks.

Acknowledgments

This work was supported by Cyclarity Therapeutics, the European Union's Horizon Europe Research and Innovation Programme (Marie Sklodowska-Curie grant agreement Bicyclos N° 101130235), the Spanish Agencia Estatal de Investigación (AEI) and the ERDF (PID2019-111327GB-I00 and PDC2022-133402-I00), by Xunta de Galicia and the ERDF (ED431B 2022/36 and Centro Singular de Investigación de Galicia, 2019–2022, Grant ED431G2019/03). All calculations were carried out at CESGA Supercomputing Center.

During the preparation of this work the author(s) used ChatGPT 4 in order to improve readability and flow of the language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2024.02.011.

References

[1] Del Valle EMM. Cyclodextrins and their uses: a review. Process Biochem 2004;39(9):1033–46. https://doi.org/10.1016/S0032-9592(03)00258-9.
[2] Loftsson T, Brewster ME. Pharmaceutical applications of cyclodextrins. 1. Drug solubilization and stabilization. J Pharm Sci 1996;85(10):1017–25. https://doi.org/10.1021/js950534b.
[3] Manor PC, Saenger W. Topography of cyclodextrin inclusion complexes. III. Crystal and molecular structure of cyclohexaamylose hexahydrate, the water dimer inclusion complex. J Am Chem Soc 1974;96(11):3630–9. https://doi.org/10.1021/ja00818a042.
[4] Chatjigakis AK, Donze C, Coleman AW, Cardot P. Solubility behavior of.beta.-cyclodextrin in water/cosolvent mixtures. Anal Chem 1992;64(14):1632–4. https://doi.org/10.1021/ac00038a022.
[5] Frijlink HW, Visser J, Hefting NR, Oosting R, Meijer DKF, Lerk CF. The pharmacokinetics of β-cyclodextrin and hydroxypropyl-β-cyclodextrin in the rat. Pharm Res 1990;7(12):1248–52. https://doi.org/10.1023/A:1015929720063.
[6] Irie T, Fukunaga K, Pitha J. Hydroxypropylcyclodextrins in parenteral use. i: lipid dissolution and effects on lipid transfers in vitro. J Pharm Sci 1992;81(6):521–3. https://doi.org/10.1002/jps.2600810609.
[7] Kovac AL. Sugammadex: the first selective binding reversal agent for neuromuscular block. J Clin Anesth 2009;21(6):444–53. https://doi.org/10.1016/j.jclinane.2009.05.002.
[8] Moser JG, Rose I, Wagner B, Wieneke T, Vervoorts A. Taxol inclusion complexes with a cyclodextrin dimer: possibilities to detoxify chemotherapeutics and to target drugs specifically to tumors? J Incl Phenom Macrocycl Chem 2001;39(1):13–8. https://doi.org/10.1023/A:1008114826524.
[9] Skanchy DJ, Xie G-H, Tait RJ, Luna E, Demarest C, Stobaugh JF. Application of sulfobutylether-β-cyclodextrin with specific degrees of substitution for the enantioseparation of pharmaceutical mixtures by capillary electrophoresis. Electrophoresis 1999;20(13):2638–49. https://doi.org/10.1002/(SICI)1522-2683(19990901)20:13<2638::AID-ELPS2638>3.0.CO;2-Z.
[10] Arima H, Motoyama K, Higashi T. Potential use of cyclodextrins as drug carriers and active pharmaceutical ingredients. Chem Pharm Bull (Tokyo) 2017;65(4):341–8. https://doi.org/10.1248/cpb.c16-00779.
[11] Cai W, Sun T, Liu P, Chipot C, Shao X. Inclusion mechanism of steroid drugs into β-cyclodextrins. Insights from free energy calculations. J Phys Chem B 2009;113(22):7836–43. https://doi.org/10.1021/jp901825w.
[12] Loftsson T, Moya-Ortega MD, Alvarez-Lorenzo C, Concheiro A. Pharmacokinetics of cyclodextrins and drugs after oral and parenteral administration of drug/cyclodextrin complexes. J Pharm Pharmacol 2016;68(5):544–55. https://doi.org/10.1111/jphp.12427.
[13] Barbosa PFP, Cumba LR, Andrade RDA, do Carmo DR. Chemical modifications of cyclodextrin and chitosan for biological and environmental applications: metals and organic pollutants adsorption and removal. J Polym Environ 2019;27(6):1352–66. https://doi.org/10.1007/s10924-019-01434-x.
[14] Challa R, Ahuja A, Ali J, Khar RK. Cyclodextrins in drug delivery: an updated review. AAPS PharmSciTech 2005;6(2):E329–57. https://doi.org/10.1208/pt060243.
[15] de Boer T, de Zeeuw RA, de Jong GJ, Ensing K. Recent innovations in the use of charged cyclodextrins in capillary electrophoresis for chiral separations in pharmaceutical analysis. Electrophoresis 2000;21(15):3220–39. https://doi.org/10.1002/1522-2683(20000901)21:15<3220::AID-ELPS3220>3.0.CO;2-X.
[16] Hedges, A.R.; Shieh, W.J.; Sikorski, C.T. Use of Cyclodextrins for Encapsulation in the Use and Treatment of Food Products. In Encapsulation and Controlled Release of Food Ingredients; Risch, S. J., Reineccius, G. A., Eds. ACS Symposium Series; American Chemical Society: Washington, DC; 1995; Vol. 590, pp 60–71. https://doi.org/10.1021/bk-1995–0590.ch006.
[17] Marques HMC. A review on cyclodextrin encapsulation of essential oils and volatiles. Flavour Fragr J 2010;25(5):313–26. https://doi.org/10.1002/ffj.2019.
[18] Trotta F, Cavalli R. Characterization and applications of new hyper-cross-linked cyclodextrins. Compos Interfaces 2009;16(1):39–48. https://doi.org/10.1163/156855408X379388.
[19] Anderson AM, Kirtadze T, Malanga M, Dinh D, Barnes C, Campo A, Clemens DM, Garcia-Fandiño R, Piñeiro Á, O'Connor MS. Cyclodextrin dimers: a versatile approach to optimizing encapsulation and their application to therapeutic extraction of toxic oxysterols. Int J Pharm 2021;606:120522. https://doi.org/10.1016/j.ijpharm.2021.120522.
[20] Final Progress Reports: CycloPure, Inc.: Remediation of Perfluorinated Chemicals in Water Using Novel High-Affinity Polymer Adsorbents (Superfund Research Program). National Institute of Environmental Health Sciences. https://tools.niehs.nih.gov/srp/programs/progress_report.cfm?Project_ID=R43ES029401 (Accessed 2023–09-06).
[21] Alcalde MA, Antelo A, Jover A, Meijide F, Tato JV. Solubilization of cholesterol in aqueous solution by two β-cyclodextrin dimers and a negatively charged β-cyclodextrin derivative. J Incl Phenom Macrocycl Chem 2009;63(3–4):309–17. https://doi.org/10.1007/s10847-008-9524-3.
[22] Atger VM, de la Llera Moya M, Stoudt GW, Rodrigueza WV, Phillips MC, Rothblat GH. Cyclodextrins as catalysts for the removal of cholesterol from macrophage foam cells. J Clin Invest 1997;99(4):773–80. https://doi.org/10.1172/JCI119223.
[23] Caira M, Bourne S, Samsodien H, Smith V. Inclusion Complexes of 2-methoxyestradiol with dimethylated and permethylated β-cyclodextrins: models for cyclodextrin-steroid interaction. Beilstein J Org Chem 2015:2616–30. https://doi.org/10.3762/bjoc.11.281.
[24] Christoforides E, Papaioannou A, Bethanis K. Crystal structure of the inclusion complex of cholesterol in β-cyclodextrin and molecular dynamics studies. Beilstein J Org Chem 2018;14:838–48. https://doi.org/10.3762/bjoc.14.69.
[25] de Jong MR, Engbersen JF, Huskens J, Reinhoudt DN. Cyclodextrin dimers as receptor molecules for steroid sensors. Chem Weinh Bergstr Ger 2000;6(21):4034–40. https://doi.org/10.1002/1521-3765(20001103)6:21<4034::aid-chem4034>3.0.co;2-3.
[26] Flood KG, Reynolds ER, Snow NH. Determination of apparent association constants of steroid–cyclodextrin inclusion complexes using a modification of the hummel–dreyer method. J Chromatogr A 2001;913(1–2):261–8. https://doi.org/10.1016/S0021-9673(01)00567-2.
[27] Kučáková K, Dolenský B. Molecular structure study of a heptakis(2,3,6-Tri-O-Methyl)-β-cyclodextrin complex of cholesterol. Steroids 2020;155:108555. https://doi.org/10.1016/j.steroids.2019.108555.
[28] Sursyakova VV, Levdansky VA, Rubaylo AI. Evaluation of the effect of background electrolyte composition and independence of parameters in determining binding constants of betulin derivatives to B- and dimethyl-β-cyclodextrins by affinity capillary electrophoresis. J Sep Sci 2022;45(19):3745–53. https://doi.org/10.1002/jssc.202200453.
[29] Šolínová V, Mikysková H, Kaiser MM, Janeba Z, Holý A, Kašička V. Estimation of apparent binding constant of complexes of selected acyclic nucleoside phosphonates with β-cyclodextrin by affinity capillary electrophoresis. Electrophoresis 2016;37(2):239–47. https://doi.org/10.1002/elps.201500337.
[30] Schönbeck C, Kari J, Westh P. ITC analysis of polydisperse systems: unravelling the impact of sample heterogeneity. Anal Biochem 2024;687:115446. https://doi.org/10.1016/j.ab.2023.115446.
[31] Welliver M. Discovery, development, and clinical application of sugammadex sodium, a selective relaxant binding agent. Drug Des Devel Ther 2008:49. https://doi.org/10.2147/DDDT.S2757.
[32] Anderson AM, Garcia-Fandiño R, Piñeiro Á, O'Connor MS. Unraveling the molecular dynamics of sugammadex-rocuronium complexation: a blueprint for cyclodextrin drug design. Carbohydr Polym 2024:122018. https://doi.org/10.1016/j.carbpol.2024.122018.
[33] Tahıl G, Delorme F, Le Berre D, Monflier É, Sayede A, Tilloy S. Curated dataset of association constants between a cyclodextrin and a guest for machine learning. Chem Data Collect 2023;45:101022. https://doi.org/10.1016/j.cdc.2023.101022.
[34] Anderson AM, Manet I, Malanga M, Clemens DM, Sadrerafi K, Piñeiro Á, García-Fandiño R, O'Connor MS. Addressing the complexities in measuring cyclodextrin-sterol binding constants: a multidimensional study. Carbohydr Polym 2024;323:121360. https://doi.org/10.1016/j.carbpol.2023.121360.
[35] Ramírez D. Computational methods applied to rational drug design. Open Med Chem J 2016;10:7–20. https://doi.org/10.2174/1874104501610010007.
[36] Arsiccio A, Rospiccio M, Shea J-E, Pisano R. Force field parameterization for the description of the interactions between hydroxypropyl-β-cyclodextrin and proteins. J Phys Chem B 2021;125(27):7397–405. https://doi.org/10.1021/acs.jpcb.1c04033.
[37] Cézard C, Trivelli X, Aubry F, Djedaïni-Pilard F, Dupradeau F-Y. Molecular dynamics studies of native and substituted cyclodextrins in different media: 1. charge derivation and force field performances. Phys Chem Chem Phys 2011;13(33):15103. https://doi.org/10.1039/c1cp20854c.
[38] Mazurek AH, Szeleszczuk Ł, Gubica T. Application of molecular dynamics simulations in the analysis of cyclodextrin complexes. Int J Mol Sci 2021;22(17):9422. https://doi.org/10.3390/ijms22179422.
[39] Ott K-H, Meyer B. Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations. J Comput Chem 1996;17(8):1068–84. https://doi.org/10.1002/(SICI)1096-987X(199606)17:8<1068::AID-JCC14>3.0.CO;2-A.

[40] Senderowitz H, Parish C, Still WC. Carbohydrates: United Atom AMBER* parameterization of pyranoses and simulations yielding anomeric free energies. J Am Chem Soc 1996;118(8):2078–86. https://doi.org/10.1021/ja9529652.

[41] Dodziuk H. Rigidity versus Flexibility. A review of experimental and theoretical studies pertaining to the cyclodextrin nonrigidity. J Mol Struct 2002;614(1):33–45. https://doi.org/10.1016/S0022-2860(02)00236-3.

[42] F. Garrido P, Calvelo M, Garcia-Fandiño R, Piñeiro Á. Rings, hexagons, petals, and dipolar moment sink-sources: the fanciful behavior of water around cyclodextrin complexes. Biomolecules 2020;10(3):431. https://doi.org/10.3390/biom10030431.

[43] Barducci A, Bonomi M, Parrinello M. Metadynamics. WIREs Comput Mol Sci 2011;1(5):826–43. https://doi.org/10.1002/wcms.31.

[44] Bussi G, Laio A. Using metadynamics to explore complex free-energy landscapes. Nat Rev Phys 2020;2(4):200–12. https://doi.org/10.1038/s42254-020-0153-0.

[45] López CA, de Vries AH, Marrink SJ. Computational microscopy of cyclodextrin mediated cholesterol extraction from lipid model membranes. Sci Rep 2013;3(1):2071. https://doi.org/10.1038/srep02071.

[46] Anderson AM, Campo A, Fulton E, Corwin A, Jerome WG, O'Connor MS. 7-ketocholesterol in disease and aging. Redox Biol 2020;29:101380. https://doi.org/10.1016/j.redox.2019.101380.

[47] Kritharides L, Kus M, Brown AJ, Jessup W, Dean RT. Hydroxypropyl-β-cyclodextrin-mediated efflux of 7-ketocholesterol from macrophage foam cells. J Biol Chem 1996;271(44):27450–5. https://doi.org/10.1074/jbc.271.44.27450.

[48] Malde AK, Zuo L, Breeze M, Stroet M, Poger D, Nair PC, Oostenbrink C, Mark AE. An automated force field topology builder (ATB) and repository: Version 1.0. J Chem Theory Comput 2011;7(12):4026–37. https://doi.org/10.1021/ct200196m.

[49] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126(1):014101. https://doi.org/10.1063/1.2408420.

[50] Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 1981;52(12):7182–90. https://doi.org/10.1063/1.328693.

[51] Darden T, York D, Pedersen L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. J Chem Phys 1993;98(12):10089–92. https://doi.org/10.1063/1.464397.

[52] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103(19):8577–93. https://doi.org/10.1063/1.470117.

[53] Van Gunsteren WF, Berendsen HJC. A leap-frog algorithm for stochastic dynamics. Mol Simul 1988;1(3):173–85. https://doi.org/10.1080/08927028808080941.

[54] Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comput Chem 1992;13(8):952–62. https://doi.org/10.1002/jcc.540130805.

[55] Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18(12):1463–72. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

[56] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[57] Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 1995;91(1–3):43–56. https://doi.org/10.1016/0010-4655(95)00042-E.

[58] Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 2004;25(13):1656–76. https://doi.org/10.1002/jcc.20090.

[59] Bonomi M, Branduardi D, Bussi G, Camilloni C, Provasi D, Raiteri P, Donadio D, Marinelli F, Pietrucci F, Broglia RA, Parrinello M. PLUMED: a portable plugin for free-energy calculations with molecular dynamics. Comput Phys Commun 2009;180(10):1961–72. https://doi.org/10.1016/j.cpc.2009.05.011.

[60] Bonomi M, Bussi G, Camilloni C, Tribello GA, Banáš P, Barducci A, Bernetti M, Bolhuis PG, Bottaro S, Branduardi D, Capelli R, Carloni P, Ceriotti M, Cesari A, Chen H, Chen W, Colizzi F, De S, De La Pierre M, Donadio D, Drobot V, Ensing B, Ferguson AL, Filizola M, Fraser JS, Fu H, Gasparotto P, Gervasio FL, Giberti F, Gil-Ley A, Giorgino T, Heller GT, Hocky GM, Iannuzzi M, Invernizzi M, Jelfs KE, Jussupow A, Kirilin E, Laio A, Limongelli V, Lindorff-Larsen K, Löhr T, Marinelli F, Martin-Samos L, Masetti M, Meyer R, Michaelides A, Molteni C, Morishita T, Nava M, Paissoni C, Papaleo E, Parrinello M, Pfaendtner J, Piaggi P, Piccini G, Pietropaolo A, Pietrucci F, Pipolo S, Provasi D, Quigley D, Raiteri P, Raniolo S, Rydzewski J, Salvalaglio M, Sosso GC, Spiwok V, Šponer J, Swenson DWH, Tiwary P, Valsson O, Vendruscolo M, Voth GA, White A, The PLUMED consortium. Promoting transparency and reproducibility in enhanced molecular simulations. Nat Methods 2019;16(8):670–3. https://doi.org/10.1038/s41592-019-0506-8.

[61] Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: New Feathers for an Old Bird. Comput Phys Commun 2014;185(2):604–13. https://doi.org/10.1016/j.cpc.2013.09.018.

[62] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

[63] team, T pandas development. Pandas-Dev/Pandas: Pandas; 2023. https://doi.org/10.5281/zenodo.7794821.

[64] McKinney W. Data structures for statistical computing in python. PROC 9th Python Sci Conf Scipy 2010;2010:56–61.

[65] Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE. Array programming with NumPy. Nature 2020;585(7825):357–62. https://doi.org/10.1038/s41586-020-2649-2.

[66] Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng 2007;9(3):90–5. https://doi.org/10.1109/MCSE.2007.55.

[67] Grønlund, A.; Larsen, K.G.; Mathiasen, A.; Nielsen, J.S.; Schneider, S.; Song, M. Fast Exact K-Means, k-Medians and Bregman Divergence Clustering in 1D. arXiv April 25; 2018. http://arxiv.org/abs/1701.07204 (Accessed 2023–04-04).

[68] Wu X. Optimal quantization by matrix searching. J Algorithms 1991;12(4):663–73. https://doi.org/10.1016/0196-6774(91)90039-2.

[69] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52(3–4):591–611. https://doi.org/10.1093/biomet/52.3-4.591.

[70] Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat 1947;18(1):50–60. https://doi.org/10.1214/aoms/1177730491.

[71] Wilcoxon F. Individual comparisons by ranking methods. Biom Bull 1945;1(6).

[72] Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician 1983;32(3):307. https://doi.org/10.2307/2987937.

[73] Castagne D, Dive G, Evrard B, Frédérich M, Piel G. Spectroscopic studies and molecular modeling for understanding the interactions between cholesterol and cyclodextrins. J Pharm Pharm Sci 2010;13(3):362–77. https://doi.org/10.18433/J3BS34.

[74] Nishijo J, Moriyama S, Shiota S, Kamigauchi M, Sugiura M. Interaction of heptakis (2,3,6-Tri-O-Methyl)-b-cyclodextrin with cholesterol in aqueous solution 2004;52(12).

[75] Ravichandran R, Divakar S. Inclusion of Ring A of cholesterol inside the -cyclodextrin cavity: evidence from oxidation reactions and structural studies. J Incl Phenom Mol Recognit Chem 1998;30:253–70.

[76] Loftsson T, Másson M, Brewster ME. Self-association of cyclodextrins and cyclodextrin complexes. J Pharm Sci 2004;93(5):1091–9. https://doi.org/10.1002/jps.20047.

[77] Messner M, Kurkov SV, Jansook P, Loftsson T. Self-assembled cyclodextrin aggregates and nanoparticles. Int J Pharm 2010;387(1–2):199–208. https://doi.org/10.1016/j.ijpharm.2009.11.035.