Received: 20 October 2023 | Revised: 18 December 2023 | Accepted: 10 January 2024

DOI: 10.1002/ardp.202300612

FULL PAPER

ARCH PHARM DPhG
Archiv der Pharmazie

Extracting binding energies and binding modes from biomolecular simulations of fragment binding to endothiapepsin

Birte Schmitz[1] | Benedikt Frieg[1,2] | Nadine Homeyer[1] | Gisela Jessen[1†] | Holger Gohlke[1,2,3]

This study is dedicated to Prof. Dr. Gerhard Klebe on the occasion of his 70th birthday. For Holger Gohlke, this study with data from the Klebe lab signifies the closure of a circle, which was opened more than 25 years ago when Holger Gohlke became a PhD student in the Klebe lab.

[1]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany

[3]Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich, Jülich, Germany

Correspondence
Holger Gohlke, Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, Düsseldorf 40225, Germany.
Email: gohlke@uni-duesseldorf.de

Present address
Nadine Homeyer, Charité Research Organisation GmbH, Charitéplatz 110117 Berlin, Germany.

Funding information
Deutsche Forschungsgemeinschaft

Abstract

Fragment-based drug discovery (FBDD) aims to discover a set of small binding fragments that may be subsequently linked together. Therefore, in-depth knowledge of the individual fragments' structural and energetic binding properties is essential. In addition to experimental techniques, the direct simulation of fragment binding by molecular dynamics (MD) simulations became popular to characterize fragment binding. However, former studies showed that long simulation times and high computational demands per fragment are needed, which limits applicability in FBDD. Here, we performed short, unbiased MD simulations of direct fragment binding to endothiapepsin, a well-characterized model system of pepsin-like aspartic proteases. To evaluate the strengths and limitations of short MD simulations for the structural and energetic characterization of fragment binding, we predicted the fragments' absolute free energies and binding poses based on the direct simulations of fragment binding and compared the predictions to experimental data. The predicted absolute free energies are in fair agreement with the experiment. Combining the MD data with binding mode predictions from molecular docking approaches helped to correctly identify the most promising fragments for further chemical optimization. Importantly, all computations and predictions were done within 5 days, suggesting that MD simulations may become a viable tool in FBDD projects.

KEYWORDS

aspartic protease, binding-free energy, endothiapepsin, fragment-based drug discovery, molecular dynamics simulations

†During this work, Gisela Jessen deceased.

1 | INTRODUCTION

Fragment-based drug discovery (FBDD) is a lead generation strategy based on screening small compounds that typically exhibit low affinity toward protein targets.[1,2] These compounds, also known as fragments, are small in size and low in molecular weight (usually less than 300 Da) and can be used to explore the binding sites of proteins with high efficiency and specificity.[3–5] FBDD relies on biophysical techniques, such as nuclear magnetic resonance,[6] surface plasmon resonance,[7] isothermal titration calorimetry (ITC),[8] X-ray crystallography, and cryogenic electron microscopy[9] to detect and quantify the interactions between fragments and proteins. Once a fragment is identified as binding, it can be optimized by growing or linking it to other fragments to increase its potency and selectivity. FBDD has several advantages over conventional high-throughput screening, such as requiring less compound synthesis, covering more chemical space, generating more ligand-efficient leads, and facilitating structure-based drug design (SBDD).[10] FBDD has been successfully applied to various protein families and therapeutic areas, discovering several clinical candidates and approved drugs.[11–13]

Endothiapepsin is an aspartic protease first isolated from the plant pathogenic fungus *Endothia parasitica*.[14] This enzyme belongs to the family of pepsin-like aspartic proteases, which are relevant for the pathogenesis of diseases such as malaria, Alzheimer's disease, fungal infections, and hypertension.[15] Endothiapepsin has been used as a model enzyme for identifying and optimizing inhibitors of other aspartic proteases, such as renin[16] and β-secretase.[17] It is also a well-characterized model system for the elucidation of the catalytic mechanism of pepsin-like aspartic proteases.[18-20] For example, pepsin-like aspartic proteases consist of two homologous domains forming a cleft-shaped active site, which harbors the catalytic dyad formed by aspartic acid residues (D35 and D219 in endothiapepsin).[21]

Work in FBDD for endothiapepsin started with the construction of a fragment library with 364 entries without strict adherence to the rule-of-three. Initial screening for endothiapepsin binding yielded 55 hits inhibiting the enzyme by at least 40%. Subsequent crystallography studies revealed diverse binding modes of the fragments.[22] Additionally, de novo SBDD was combined with dynamic combinatorial chemistry (DCC) to generate a library of potential acylhydrazone inhibitors, directed by the target endothiapepsin. NMR and crystallography validated the binding modes of the most potent inhibitors.[23] In a follow-up study, SBDD and DCC were used to optimize a hit by focusing on an amide–π interaction.[21] Finally, by combining fragment linking and DCC, a library of bis-acylhydrazones was designed based on X-ray structures and subjected to hit identification. The most potent inhibitor displayed substantial improvement in potency compared to parent hits.[24]

In recent years, computational approaches have gained an important role in FBDD, especially in SBDD. Molecular dynamics (MD) simulations are a powerful tool to investigate the conformational dynamics and flexibility of proteins and ligands and the thermodynamics and kinetics of their binding processes.[25–27] MD simulations can provide atomic-level insights into the binding modes, interactions, and mechanisms of binding of fragments and their derivatives and identify potential sites for fragment optimization. Moreover, MD simulations can be combined with rigorous free-energy calculations to estimate the binding affinity and selectivity of fragments and guide the design of new compounds by evaluating the effect of structural modifications on the binding-free energy.[25,28] Applying MD simulations to FBDD, Martinez-Rosell et al. screened 129 fragments to identify mM-affinity fragments against the CXCL12 monomer by using a Markov state model based on a total average of 54 μs MD simulations data per fragment.[29] Linker et al. combined MD simulations and Markov state models to predict binding sites and binding modes of fragment-like small molecules, and for each protein–fragment pair, 50 μs of accumulated simulation time was computed.[30] Pan et al. performed long unbiased MD simulations of 20–39 μs length for six fragments to record spontaneous binding and unbinding events, allowing them to derive the binding affinities and kinetics, as well as binding poses directly from the MD data.[31] However, judging by the reported simulation times, MD simulations-based approaches are computationally highly demanding, and multimicrosecond long simulations require weeks or even months of computation, which limits the applicability of MD simulations in FBDD.

To investigate the current strengths and limitations of MD simulations in an FBDD project, we performed short, unbiased MD simulations of direct fragment binding to endothiapepsin. We predicted the fragments' absolute binding-free energies, as well as the fragments' binding sites and poses directly from the MD trajectories and compared our predictions with experiments. All computations were finished and predictions were made within 5 business days, in a time frame suitable to be applied within the scope of an FBDD project. Interestingly, while the absolute free energy predictions are in reasonable agreement with experiments, the MD simulations struggled with the prediction of endothiapepsin–fragment complex structures.

2 | RESULTS AND DISCUSSION

2.1 | A small fragment library

The fundamental concept of an FBDD project is to use small, simple, and easily accessible molecules, allowing for the sampling of a more extensive chemical space compared to molecules with advanced complexity.[32] As a guideline, fragments follow a rule-of-three, although this guideline has been challenged.[22] Accordingly, a drug fragment has a molecular weight <300 Da, not more than three hydrogen-bond donor and acceptor atoms, respectively, and a cLogP < 3.[3,4] Additionally, less than three rotatable bonds and a polar surface area <60 $Å^2$ have been suggested as beneficial.[3,4] In the present study, we considered fragments that mostly agree with the rule-of-three, such that all molecules have fragment-like properties (Supporting Information S1: Table S2).

To study the potential of MD simulations to predict binding affinities and binding poses of fragments in FBDD, the library was initially split into a training (Table 1) and prediction set (Table 2). To avoid any bias, the assignment was done by the experimentalists (A. Metz, G. Klebe, personal communication; the assignment relates to different series of fragments that were measured at consecutive times) and remained semi-blind for the computational chemists, that is, experimental binding affinities and crystallographic binding poses were initially provided only for the training set and only after computational results had been generated for the prediction set.

2.2 | Fragments spontaneously unbind from holo complexes during short, unbiased MD simulations

First, to investigate the structural behavior of endothiapepsin in complex with fragments, we performed unbiased MD simulations starting from the crystallographically resolved complexes of the three fragments showing the highest binding affinity toward endothiapepsin, namely **02**, **47**, and **60** (PDB ID: 3PBD, 3PM4, 3PI0[22]). Furthermore, to make binding-free energy and pose predictions for fragments binding to endothiapepsin, we also established a measure to estimate when a fragment is bound to or unbound from the enzyme. The catalytic site of endothiapepsin can be divided into three subpockets S1, S2, and S3. Initially, fragment **02** is bound to the S1 pocket, fragment **47** to the S2 pocket, and fragment **60** to the S3 pocket (Supporting Information S1: Figure S1). We considered both protonation states endothiapepsin^{Asp219} and endothiapepsin^{AspH219}, as both may be relevant under crystallization conditions.[22] To evaluate the structural variability of the binding pose, the distance d_{bind} between the fragments' centers of mass (COM; only considering nonhydrogen atoms) and the COM of the fragments' binding pockets was computed along the trajectories.

For all fragments, d_{bind} increased during short MD simulations of 100 ns length, indicating that the fragments move away from the crystallographic poses (Figure 1). The major peaks of the distance distributions are shifted toward smaller distances if the fragments are initially bound to the deprotonated/ionic endothiapepsin^{Asp219} structure, suggesting that they are more stable there (Figure 1a,c). A plausible explanation is that the cationic fragments form ionic interactions with the anionic amino acid in the starting complex. Nevertheless, independent of the Asp219 protonation state, even during MD simulations of 100 ns, the fragments are released from the catalytic site to the bulk solvent (d_{bind} > 10 Å, Figure 1), indicating that they are weakly bound in the X-ray structures, except for **60** bound to Asp219, which is not fully released to the bulk solvent after that time.

Interestingly, we observed spontaneous (re)binding of the fragments to endothiapepsin within MD simulations of 100 ns length. With a higher number of spontaneous binding and unbinding events occurred during longer simulation times, it shall be possible to estimate the binding affinities directly from the simulation data,[31,34] based on the law of mass action for reversible chemical reactions. To do so, it is critical to define when a fragment is bound to or unbound from endothiapepsin, respectively.[31,34] The question arises, however, how to define a structurally bound state for fragments weakly bound to endothiapepsin? Given the high mobility in the binding site, using the standard measure root-mean-square deviation (RMSD) seemed inappropriate.

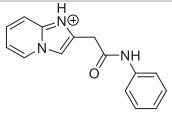
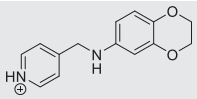
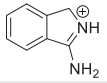
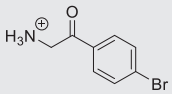
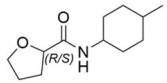
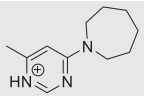
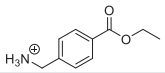
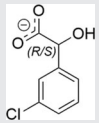
To visualize the displacement of the fragments from the crystallographic pose during 100 ns of MD simulations, we extracted multiple fragment poses from MD trajectories for increasing d_{bind} values and compared the MD poses with the crystallographic poses (Figure 1b,d, Supporting Information S1: Figure S2; Table S3). Binding poses with d_{bind} < 3 Å are still in good agreement with the crystallographic poses. Binding poses with distances 3 Å < d_{bind} ≤ 6 Å qualitatively agree with the crystallographic pose, such that the fragments bind to the correct sub-pockets and still partially overlap with the crystallographic pose. For poses with d_{bind} > 6 Å, the orientations of the fragments do not agree with the crystallographic poses but the fragments are still in the cavity of the binding site. If d_{bind} > 10 Å, the fragment can be considered fully solvated by bulk water (Supporting Information S1: Figure S3). Hence, we defined a fragment as bound for d_{bind} ≤ 6 Å and unbound for d_{bind} > 10 Å.

2.3 | Absolute binding-free energies derived from multiple short MD trajectories are mostly within 1 kcal mol^{-1} of experimental binding affinities

We performed unbiased MD simulations of fragment binding to endothiapepsin of 600 ns length to estimate binding affinities directly from the MD ensemble. Initially, 16 fragments were randomly placed around endothiapepsin to reach a concentration of approximately 50 mM in the simulation box. To consider the experimentally observed protonation–deprotonation equilibrium of Asp219 in our MD simulations, we performed five replica simulations with the deprotonated and charged amino acid (Asp219) and five with the protonated and noncharged amino acid (AspH219). In nearly all cases, we observed multiple binding and unbinding events of the fragments using the above $d_{bind.}$ criteria (Supporting Information S1: Figure S4, Supporting Information S2: Movie S1). Mean absolute standard-free energies of binding ($\Delta G°_{bind, sim}$) were computed according to Equations (1)–(4) and compared to free energies of binding from ITC experiments ($\Delta G°_{bind, exp}$) (Table 1; Table S4).

Analyzing the Asp219 and AspH219 systems independently revealed that the computed binding-free energies are marginally more accurate compared to the experimental ones in the case of AspH219 (Table 1, Supporting Information S1: Figure S5), presumably due to a better balance of electrostatic interactions and solvation contributions between the neutral side chain and the cationic fragments.[21,35] However, neither system leads to significantly more accurate predictions than the other. The best predictions were obtained with the average over the 10 independent replicas, five with Asp219 and five with AspH219 (Figure 2). We, thus, posit that the experimental protonation–deprotonation equilibrium

TABLE 1 Fragments in the training set ordered by descending $\Delta G°_{bind.,exp.}$

Fragment structure	ID	PDB ID/Reference/DOI	K_D^{exp}[a]	$\Delta G°_{bind.,\ exp.}$[b]	ΔG_{Asp219} ± SEM[b]	$\Delta G_{AspH219}$ ± SEM[b]	$\Delta G°_{bind.,sim.}$ ± SEM[b]	AD[b]	$LE_{exp.}$	$LE_{sim.}$
	**47**	3PM4[22]	0.2	−5.08	−3.86 ± 0.30	−4.53 ± 0.62	−4.19 ± 0.34	0.89	0.27	0.22
	**60**	3PIO[22]	1.4	−3.92	−4.37 ± 0.35	−3.85 ± 0.18	−4.11 ± 0.21	0.20	0.21	0.23
	**02**	3PBD[22]	1.7	−3.80	−4.34 ± 0.12	−3.35 ± 0.29	−3.84 ± 0.22	0.04	0.38	0.38
	**10**	4Y51 10.2210/pdb4Y51/pdb	3.3	−3.41	−3.62 ± 0.34	−3.74 ± 0.26	−3.68 ± 0.20	0.27	0.31	0.34
	**52**	4Y5E 10.2210/pdb4Y5E/pdb	3.7 (bound in *R*-configu-ration)	−3.34	−3.07 ± 0.33	−4.50 ± 0.20	−3.78 ± 0.30	0.44	0.22	0.23
	**04**	4Y3P 10.2210/pdb4Y3P/pdb	>10	N.A.[c]	−4.22 ± 0.25	−4.24 ± 0.38	−4.23 ± 0.21	-	-	0.30
	**18**	4YCK 10.2210/pdb4YCK/pdb	>10	N.A.[c]	−4.61 ± 0.25	−3.84 ± 0.11	−4.22 ± 0.18	-	-	0.33
	**12R**	4Y54 10.2210/pdb4Y54/pdb	N.N.	N.N.[d]	0.03 ± 0.24	0.04 ± 0.18	0.03 ± 0.14	-	-	<0.02
	12S				0.29 ± 0.28	0.12 ± 0.12	0.21 ± 0.15	-		

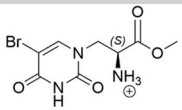
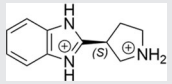
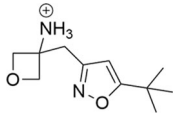
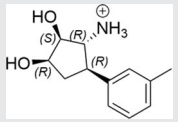
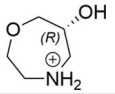
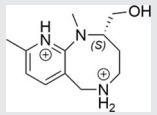
[a]In mM. K_D^{exp} values were measured as indicated in the publication or DOI given in the previous column; see also Köster.[33]

[b]In kcal mol^{-1}.

[c]Measured but too weakly binding in isothermal titration calorimetry (ITC).

[d]No binding measured in ITC.

TABLE 2 Fragments in the prediction set.

Fragment structure	ID	PDB-ID/DOI	K_D^{exp}[a]	$\Delta G°_{bind.,exp.}$[b]	ΔG_{Asp219} ± SEM[b]	$\Delta G_{AspH219}$ ± SEM[b]	$\Delta G°_{bind.,sim.}$ ± SEM[b]	AD[b]	$LE_{exp.}$	$LE_{sim.}$
	**K1A**	5QB6 10.2210/pdb5QB6/pdb	>10	N.A.[c]	−3.35 ± 0.38	−3.16 ± 0.35	−3.26 ± 0.24	-	-	0.20
	**K1B**	5QBA 10.2210/pdb5QBA/pdb	4.6	−3.21	−4.60 ± 0.39	−2.15 ± 0.14	−3.37 ± 0.45	0.17	0.23	0.24
	**K1C**	5QBB 10.2210/pdb5QBB/pdb	18.4	−2.38	−4.75 ± 0.58	−3.84 ± 0.15	−4.29 ± 0.32	1.91	0.16	0.29
	**K1D**	5QBC 10.2210/pdb5QBC/pdb	89.1	−1.44	−4.03 ± 0.40	−3.70 ± 0.27	−3.87 ± 0.23	2.42	0.10	026
	**K1E**	5QBG 10.2210/pdb5QBG/pdb	>10	N.A.[c]	−3.40 ± 0.52	−2.56 ± 0.26	−2.98 ± 0.31	-	-	0.37
	**K1F**	5QBI 10.2210/pdb5QBI/pdb	4.7	−3.20	−2.88 ± 0.38	−2.30 ± 0.42	−2.59 ± 0.28	0.60	0.20	0.16

[a]In mM. K_D^{exp} values were measured as indicated in the publication or DOI given in the previous column; see also Köster.[33]

[b]In kcal mol^{-1}.

[c]Measured but too weakly binding in isothermal titration calorimetry (ITC).

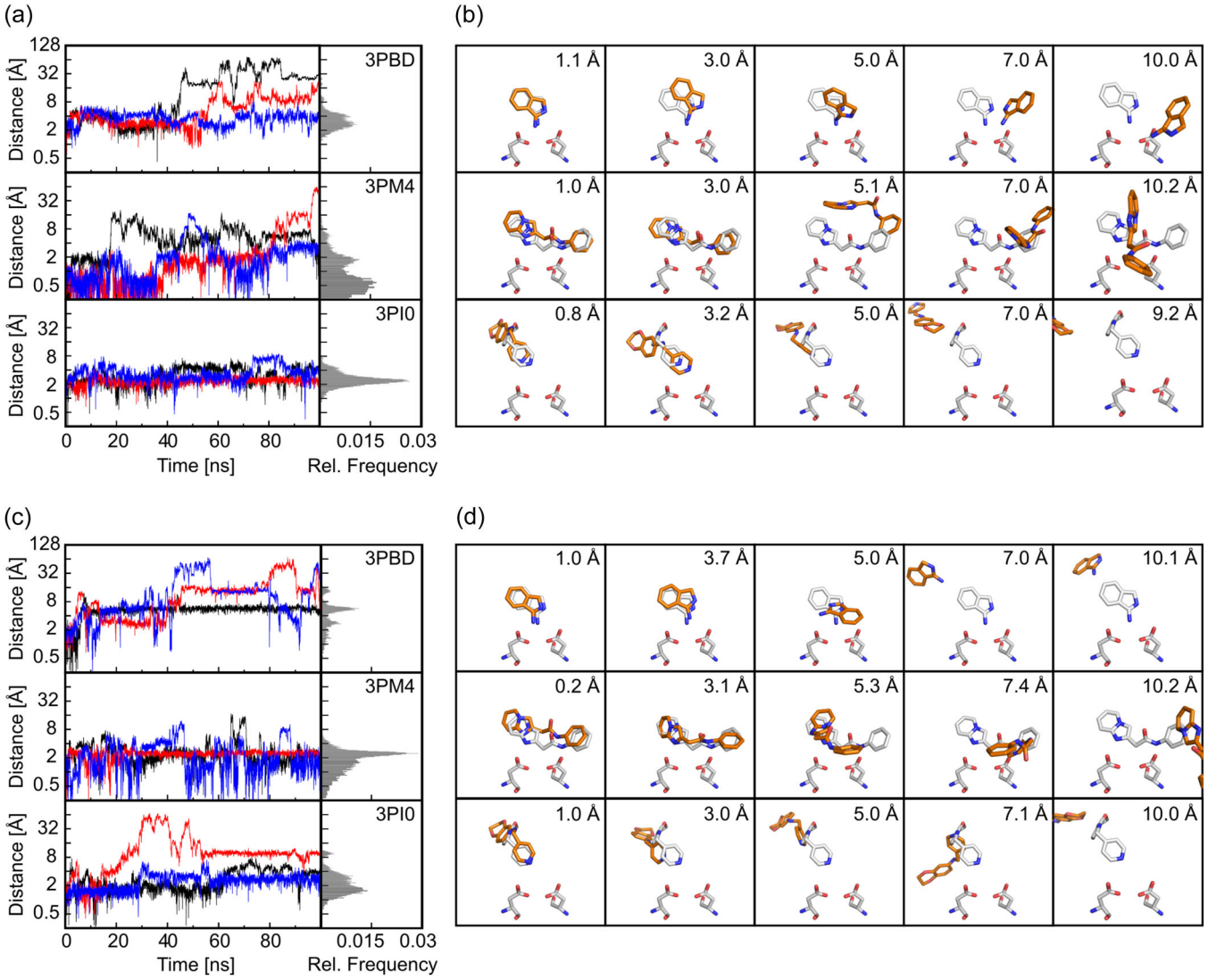


FIGURE 1 Results from molecular dynamics (MD) simulations of the holo structures. (a, c) Time series of the distances between the COM of the nonhydrogen atoms of **02** (top), **47** (middle), and **60** (bottom) and the corresponding fragment's binding site in endothiapepsin^{Asp219} (a) and endothiapepsin^{AspH219} (c) during three replica MD simulations (colored differently) of 100 ns length. All distances are summarized in a histogram (normalized to the sum over all bins; bin size = 0.04 Å) including all MD simulations. (b, d) Overlay of representative conformations of **02** (top), **47** (middle), and **60** (bottom) extracted from MD trajectories of endothiapepsin^{Asp219} (b) and endothiapepsin^{AspH219} (d) and the corresponding crystal structures PDB ID 3PBD, 3PM4, and 3PI0.[22] The fragments and amino acids Asp35 and Asp219 are shown as sticks with the crystal structure conformation colored in gray and the MD simulation conformation in orange, respectively. For clarity, all hydrogen atoms are omitted.

of Asp219 was best mimicked in our MD simulation setup, as protonated and deprotonated Asp219 species might coexist at the experimental pH of 4.6, as suggested by our p*K*a computations. We consider this global average of $\Delta G°_{bind,\,sim.}$ for comparison with $\Delta G°_{bind,\,exp}$.

Initially, we aimed to reproduce the binding affinities for the training set of experimentally characterized fragments, namely, **47**, **60**, **02**, **10**, **52R**, **18**, **04**, and **12**. In general, $\Delta G°_{bind,\,sim.}$ values agree well with $\Delta G°_{bind,\,exp}$ values from ITC experiments (summarized in Figure 2 and Table 1). For fragments **47**, **60**, **02**, **10**, **52R**, the AD (see Equation 5) between $\Delta G°_{bind,\,sim.}$ and $\Delta G°_{bind,\,exp.}$ is <1 kcal mol^{-1}, and hence within chemical accuracy.[36] $\Delta G°_{bind,\,sim.}$ is not significantly different from $\Delta G°_{bind,\,exp.}$ for **60**, **02**, **10**, and **52R**. The binding affinity for **47** was underestimated. As **47** shows the strongest binding affinity to endothiapepsin in the present data set, longer simulation times might be needed to obtain a better equilibrium between unbound and bound states. In addition, we correctly identified fragment **12** as nonbinding fragment; **12** makes repulsive interactions between the negatively charged fragment and the catalytic Asp35 and Asp219. For fragments **04** and **18**, the computed $\Delta G°_{bind,\,sim.}$ of almost 4 kcal mol^{-1} indicates them as binding, but experimental binding affinities could not be obtained. However, **04** and **18** form a complex with endothiapepsin according to crystal structure analysis (Table 1; see also below).

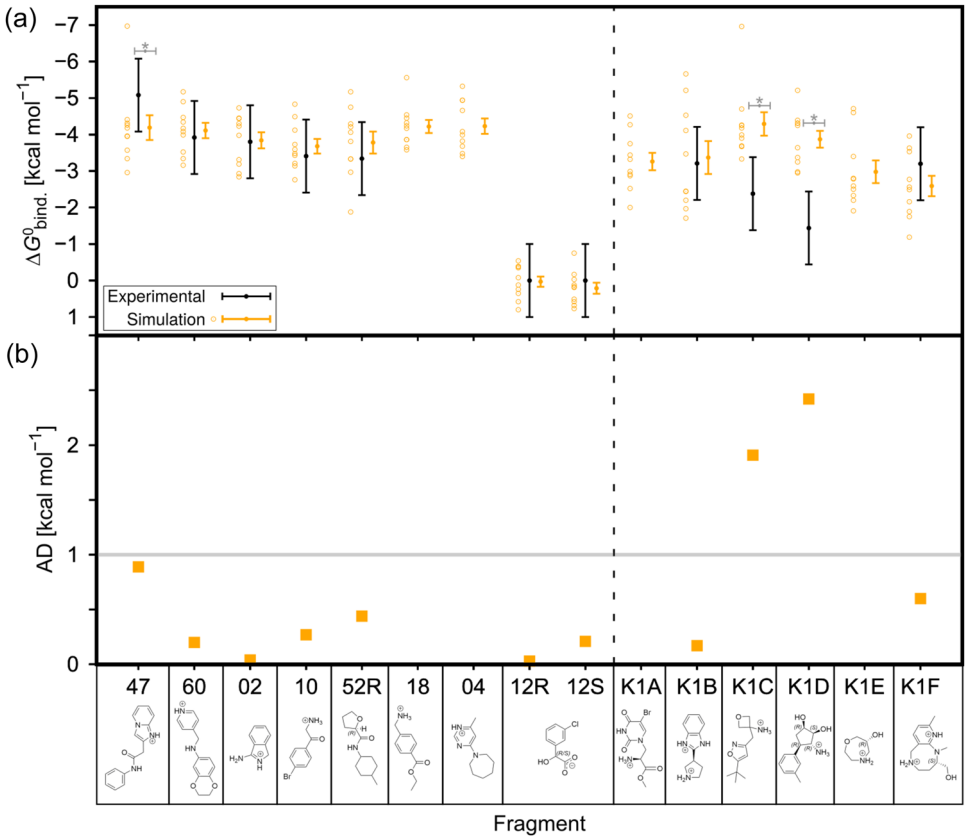


FIGURE 2 Absolute standard binding-free energies calculated from unbiased molecular dynamics (MD) simulations of fragment binding and comparison to experiment. (a) Absolute standard binding-free energies ($\Delta G°_{bind.,\,sim.}$) for each of the 10 trajectories (Asp219 and Asp**H**219 variant) are shown as open orange circles and the average as filled orange circles with error bars denoting the standard error of the mean (SEM; for $n = 10$). The experimentally determined $\Delta G°_{bind.\,exp.}$ is plotted as black circle with error bars denoting the ±1 kcal mol^{-1} interval. *Denotes $p < 0.05$ (one sample Student t-test). (b) The absolute deviation (AD) between $\Delta G°_{bind.,\,sim.}$ and $\Delta G°_{bind.\,exp.}$ (Equation 5) is shown as orange squares. The gray line at AD = 1 kcal mol^{-1} indicates the ±1 kcal mol^{-1} tolerance interval. The dashed line separates training from prediction set.

To investigate whether binding affinities for fragments can be reliably predicted, we applied the simulation procedure above to six additional fragments of the prediction set (summarized in Figure 2 and Table 2). Note that the information on the experimental binding affinities was provided to the computational chemists after they predicted $\Delta G°_{bind,\,sim}$. The comparison to $\Delta G°_{bind,\,exp.}$ revealed nonsignificant deviations only for fragments **K1B** and **K1F**, with an AD < 1 kcal mol^{-1}. In contrast, for fragments **K1C** and **K1D**, the AD between $\Delta G°_{bind,\,sim.}$ and $\Delta G°_{bind,\,exp.}$ is > 1 kcal mol^{1} and $\Delta G°_{bind,\,sim.}$ is significantly different to $\Delta G°_{bind,\,exp}$. Considering only the protonated variant of Asp219 (AspH219), $\Delta G°_{bind,\,sim.}$ improves in respect to $\Delta G°_{bind,\,exp.}$ for fragment **K1C** (Asp219: $\Delta G°_{bind,\,sim.} = -4.75 \pm 0.58$ kcal mol^{-1}; $p = 0.07$ vs. AspH219: $\Delta G°_{bind,\,sim.} = -3.84 \pm 0.15$ kcal mol^{-1}, Table 2), although residue Asp219/AspH219 does not directly interact with **K1C** (Supporting Information S1: Figure S6). Still, a neutral sidechain of Asp219 may be more favorable, considering the nearby *t*-butyl group of **K1C**. Regarding **K1D**, our simulations overestimate the binding affinities in both variants Asp219 and AspH219. A possible reason might be that **K1D** binds to a different endothiapepsin subpocket than those found in PDB IDs 3PBD, 3PM4, and 3PI0. **K1A** and **K1E** were categorized as weakly binding fragments with $\Delta G°_{bind,\,exp.}$ assumed to range from −2.7 to −1.4 kcal mol^{-1} (see also Table 2), but $\Delta G°_{bind,\,sim.}$ overestimates their affinity.

Overall, absolute binding-free energies computed from multiple short MD simulations reasonably agree with experimental binding affinities for some but not all fragments. Considering that modern rigorous-free energy calculations reproduce experimental results within ~2 kcal mol^{-1}[37–43] and that the accuracy of the calculations depends on the biophysical complexity of the investigated systems, the accuracy of the procedure reported herein is comparable. In this regard, we predicted seven fragments (**47**, **60**, **02**, **10**, **52**, **K1B**, **K1F**) with an AD < 1 kcal mol^{-1} and five fragments (**18**, **04**, **K1A**, **K1C**, **K1D**, **K1E**) with 1 < AD < 2.5 kcal mol^{-1}.

While rigorous-free energy calculations are often linked to extensive and long computing times, in the present study, we kept the total computing time on purpose low. All production simulations were concluded within 3 days, using a multi-CPU-multi-GPU computing environment to initiate all replica simulations simultaneously, which produced MD trajectories of at least 280 ns length within 24 h for each GPU (Supporting Information S1: Figure S7). Furthermore, extending the time interval for the integration of motion from 2 to 4 fs[44] halves the necessary computing time without a significant loss in the accuracy

of absolute binding-free energies predicted for the fragments **47**, **60**, and **02** (Supporting Information S1: Figure S8).

Although the prediction of absolute binding-free energies was within chemical accuracy for most cases, a solely energetic analysis might not be sufficient to correctly classify a fragment as a potential hit molecule. Hence, we extended our analysis by structurally characterizing fragment–protein interactions.

2.4 | Binding poses derived from MD trajectories partially agree with crystallographically resolved binding poses

Next, we intended to reproduce (for the training set) and predict (for the prediction set) the most likely binding poses for the fragments when bound to endothiapepsin. Therefore, all fragment poses identified above as bound to endothiapepsin^{Asp219} were further subjected to hierarchical clustering. After a coarse clustering followed by a refined clustering, the representatives from the largest cluster were compared to the available X-ray structures. As a measure for similarity, we calculated the RMSD of all nonhydrogen atoms between the pose obtained from MD simulations and the X-ray pose after superimposing the protein structures (termed $RMSD_{sim./X\text{-}ray}$ hereafter). Interestingly, although the binding-free energies could be well reproduced within our MD simulations, we recognized some discrepancies between the simulated and experimental binding poses (summarized in Figures 3 and 4).

The best matches of poses from our MD simulations and X-ray were obtained for fragments **02** ($RMSD_{sim./X\text{-}ray}$ = 0.7 Å; Figure 3a,h) and **10** ($RMSD_{sim./X\text{-}ray}$ = 1.9 Å; Figure 3b,h). As to fragments **04** and **60**, the poses from MD simulations at least partially overlap with the X-ray poses and the respective $RMSD_{sim./X\text{-}ray}$ is <5 Å (Figure 3c,d,h). As to the remaining fragments **18**, **47**, and **52R**, the MD poses differ from the X-ray structures with $RMSD_{sim./X\text{-}ray}$ up to 10 Å (Figure 3e–h). Interestingly, poses or epitopes can be correctly identified in the cases where also $\Delta G^0_{bind.}$ was satisfactorily estimated, particularly for fragments **02** and **10**. By contrast, only partially overlapping or no good binding poses were identified for **04** and **18**, for which binding affinity predictions failed. However, the reverse does not hold: although binding affinity predictions were satisfactory for **47**, **60**, and **52R**, the binding poses were only partially overlapping or differed markedly. Although **47** has the strongest affinity toward endothiapepsin in vitro, the predicted binding pose deviates from the crystallographic one by

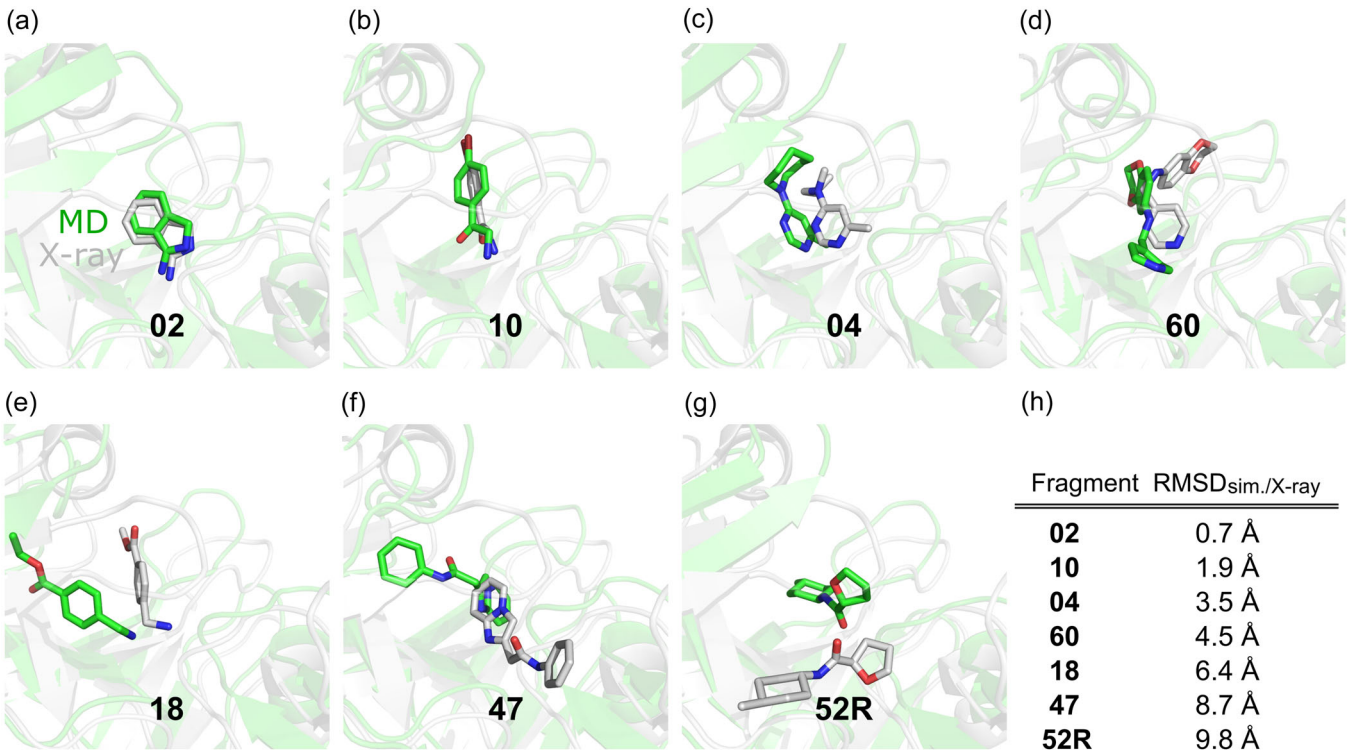


FIGURE 3 Predicted versus crystallographically resolved binding poses. Comparison of representative fragment conformations (green) extracted from the highest populated cluster of the molecular dynamics (MD) ensemble and crystal structures (white) of fragments (**02** (a): PDB ID 3PBD,[22] **10** (b): PDB ID 4Y51, **04** (c): PDB ID 4Y3P, **60** (d): PDB ID 3PI0,[22] **18** (e): PDB ID 4YCK, **47** (f): PDB ID 3PM4,[22] **52R** (g): PDB ID 4Y5E). The fragment is always shown in stick representation. The secondary structure is displayed as cartoon representation. Please note that in PDB ID 4Y3P (**04**) the 7-membered ring and in PDB ID 4YCK (**18**) the terminal carbon atom are not resolved completely. For clarity, all hydrogen atoms are omitted. (h) Fragment-wise root mean square deviation ($RMSD_{sim./X\text{-}ray}$) between the MD simulation pose and the X-ray pose (including all nonhydrogen atoms), after superimposing the protein backbone atoms.

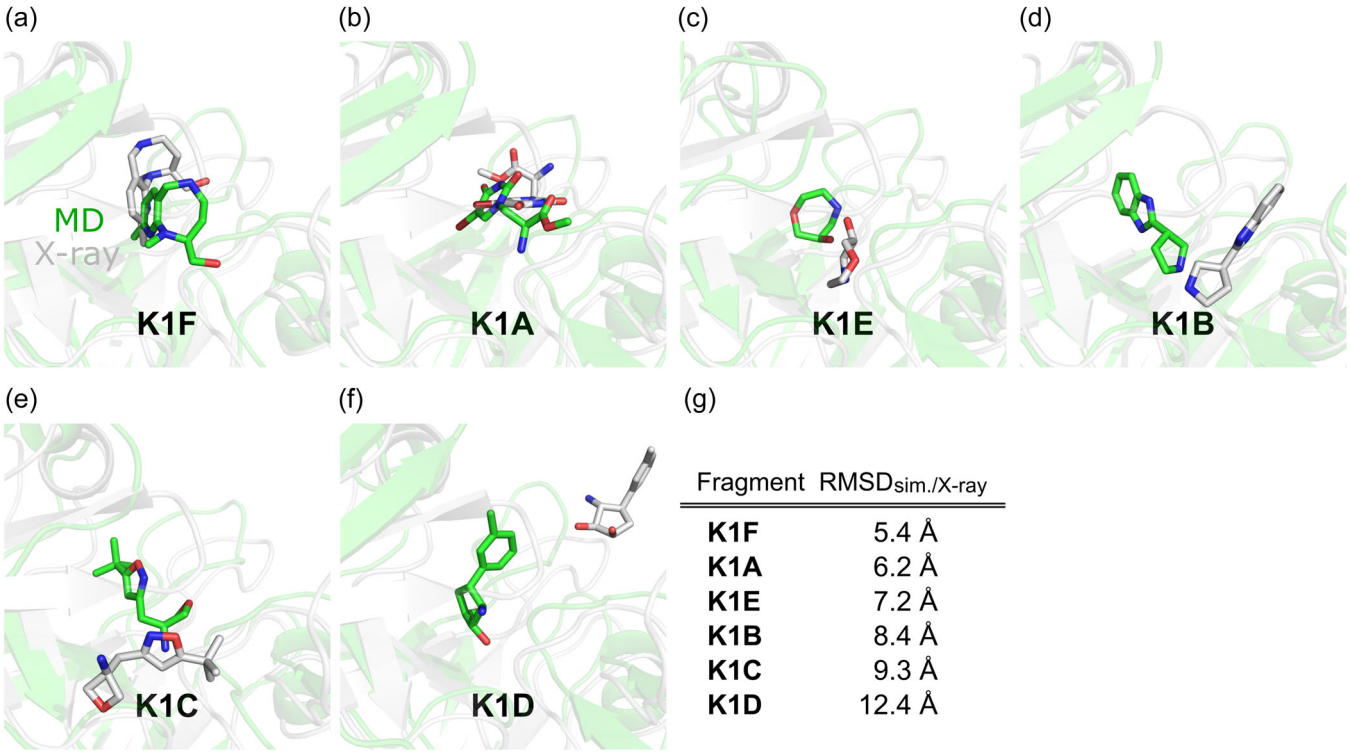


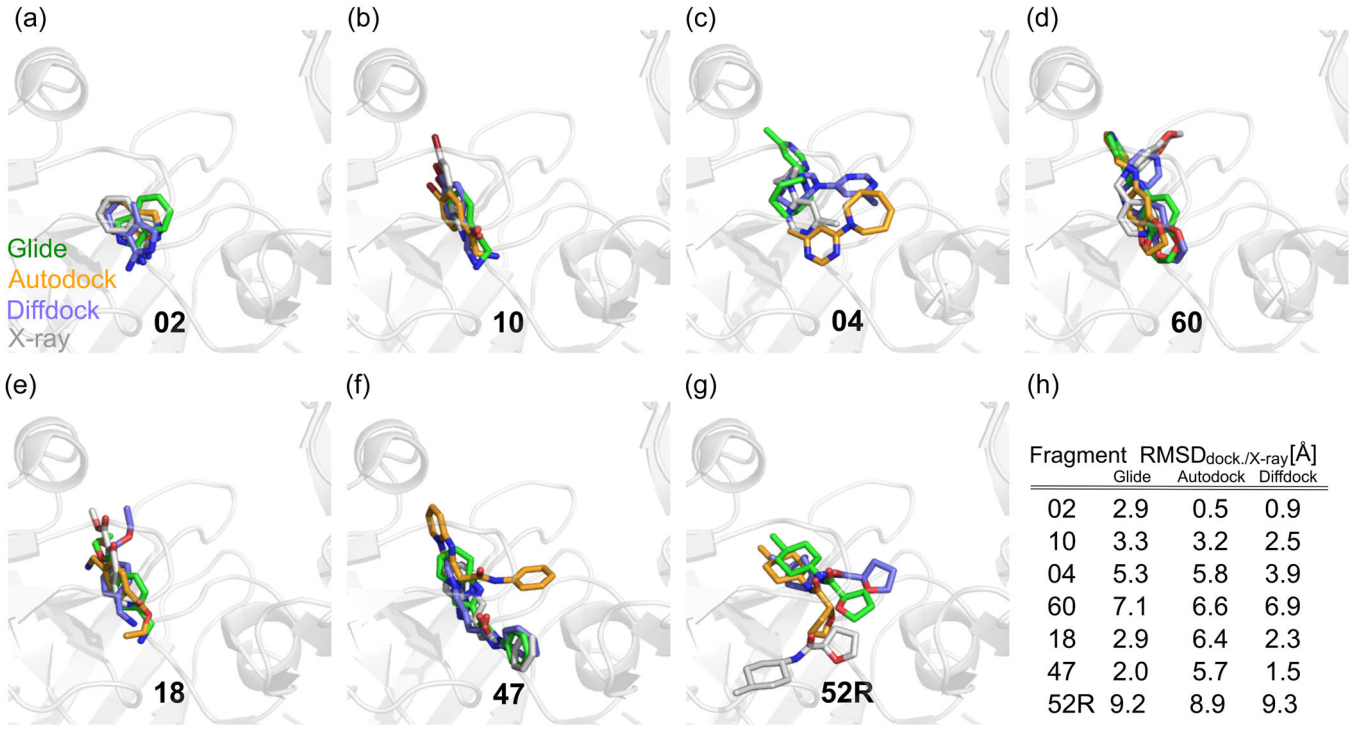
FIGURE 4 Predicted versus crystallographically resolved binding poses. Comparison of representative fragment conformations (green) extracted from the highest populated cluster of the molecular dynamics (MD) ensemble and crystal structures (white) of fragments (**K1F** (a): PDB ID 5QBI, **K1A** (b): PDB ID 5QB6, **K1E** (c): PDB ID 5QBG, **K1B** (d): PDB ID 5QBA, **K1C** (e): PDB ID 5QBB, **K1D** (f): PDB ID 5QBC; for DOIs associated with the PDB IDs see Table 2). The fragment is always shown in stick representation. The secondary structure is displayed as cartoon representation. For clarity reasons, all hydrogen atoms are omitted. (g) Fragment-wise root mean square deviation ($RMSD_{sim./X\text{-}ray}$) between the MD simulation pose and the X-ray pose (including all nonhydrogen atoms), after superimposing the protein backbone atoms.

$RMSD_{sim./X\text{-}ray}$ = 8.7 Å. In the case of the prediction set, the poses from MD simulations deviate even stronger from the X-ray poses, with $RMSD_{sim./X\text{-}ray}$ ≥ 5.4 Å (Figure 4). Note, however, that X-ray poses were obtained for fragments **04**, **18**, **K1A**, and **K1D** even though they have been categorized as non- or weakly binding fragments by ITC experiments.

In search of an explanation for the discrepancy between the outcome from absolute binding-free energy estimation versus binding mode prediction from the MD trajectories, we analyzed the MD trajectories toward structural changes within the subpockets S1, S2, and S3. Throughout the MD simulations, we observed an opening of the β-hairpin loop (residues 78–83) (Supporting Information S1: Figures S9–S11). This observation aligns with reported X-ray structures of the endothiapepsin homolog aspartic protease plasmepsin II, for which closed and open β-hairpin loop conformations are known.[45,46] Except for fragment **12**, the fragments interact with the β-hairpin loop, which might stabilize the closed conformation in the X-ray structure (Supporting Information S1: Figure S12). In turn, X-ray structure-like endothiapepsin conformations are only sporadically observed in our MD simulations, but fragment-bound configurations occur preferentially when coupled to a closed-loop configuration (Supporting Information S1: Figures S13 and S14). Much longer MD simulations or biased simulation schemes would be required to increase the occurrence frequency of these rare events, which might lower the structural differences between poses from MD simulations versus X-ray.

Finally, to test whether molecular docking programs can correctly predict the fragment binding pose, we docked the fragments into one X-ray conformation of endothiapepsin that contained the β-hairpin loop in the closed conformation. In contrast to predicting the binding poses of fragments by MD simulations, molecular docking methods are standard tools in modern drug discovery to determine protein-bound ligand configurations.[47,48] Using a static, nonflexible X-ray conformation in molecular docking may help to find the correct fragment pose if the conformation resembles the bound state. Again, we used the RMSD of all nonhydrogen atoms between the docking pose and the X-ray pose after superimposing the protein structures (termed $RMSD_{dock./X\text{-}ray}$ hereafter) to measure structural similarity.

The docking poses only partially agree with the X-ray poses, and the overall performance is similar to the prediction by MD simulations (Figures 5 and 6, Supporting Information S1: Figure S15). While a few docking poses closely mimic the experimental poses, for example, **02**, **47**, **10**, **18**, or **K1F** ($RMSD_{dock./X\text{-}ray}$ < 2.5 Å), most of the docking poses reveal $RMSD_{dock./X\text{-}ray}$ > 4 Å (Figure 7). This result is surprising given that three docking algorithms were used, increasing



Fragment	$RMSD_{dock./X\text{-}ray}$[Å] Glide	Autodock	Diffdock
02	2.9	0.5	0.9
10	3.3	3.2	2.5
04	5.3	5.8	3.9
60	7.1	6.6	6.9
18	2.9	6.4	2.3
47	2.0	5.7	1.5
52R	9.2	8.9	9.3

FIGURE 5 Docked versus crystallographically resolved binding poses. Docked fragments with the highest GlideXPScore (green), most favorable binding energy from the largest cluster (orange), and highest confidence (slate blue) in comparison to crystal structures (white) of fragments (**02** (a): PDB ID 3PBD,[22] **10** (b): PDB ID 4Y51, **04** (c): PDB ID 4Y3P, **60** (d): PDB ID 3PI0,[22] **18** (e): PDB ID 4YCK, **47** (f): PDB ID 3PM4,[22] **52R** (g): PDB ID 4Y5E). The fragment is always shown in stick representation. The secondary structure is displayed as cartoon representation. Please note that in PDB ID 4Y3P (**04**) the 7-membered ring and in PDB ID 4YCK (**18**) the terminal carbon are not resolved completely. For clarity, all hydrogen atoms are omitted. (h) Fragment-wise root mean square deviation ($RMSD_{dock./X\text{-}ray}$) between the docked pose and the X-ray pose (including all nonhydrogen atoms), after superimposing the protein backbone atoms.

the likelihood of finding a good pose, for docking into a bound endothiapepsin conformation. One might speculate that fragment binding to endothiapepsin involves multiple coupled steps, which may include an initial opening of the β-hairpin loop for access to the binding site followed by the closing of the β-hairpin loop upon fragment binding that cannot be captured when using static protein structures.

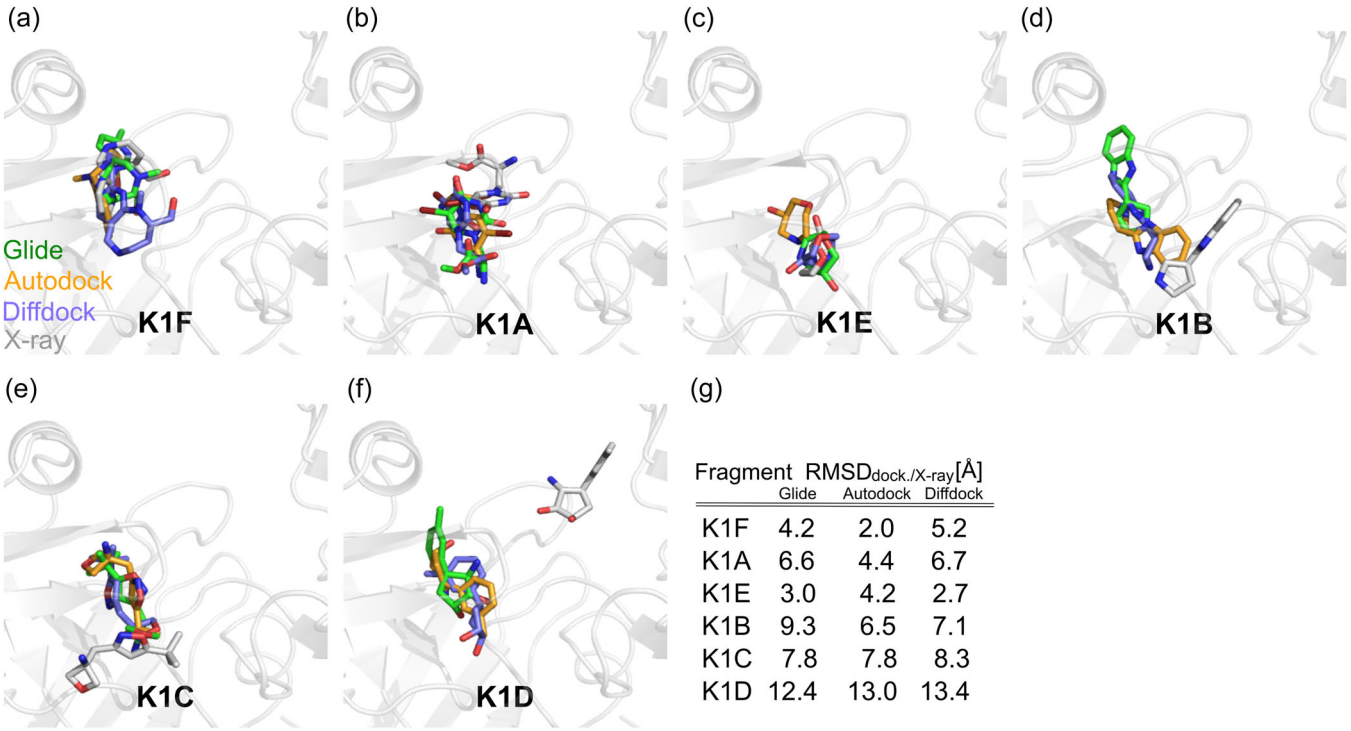
2.5 | MD simulations mimic the experimental binding pose of fragments with a high ligand efficiency

In FBDD, a promising hit fragment is not necessarily identified by its absolute binding-free energy but by its binding-free energy relative to its number of nonhydrogen atoms (i.e., ligand efficiency [LE][49,50]). Fragments with LE > 0.3 kcal mol^{-1} per nonhydrogen atom are considered promising hit fragments for further optimization toward a nanomolar inhibitor.[51] In our training and prediction sets, only two fragments reveal an experimental LE ($LE_{exp.}$) > 0.3, **02** ($LE_{exp.}$ = 0.38) and **10** ($LE_{exp.}$ = 0.31). For both fragments, the simulated $LE_{sim.}$ are also >0.3 (**02**: $LE_{sim.}$ = 0.38; **10**: $LE_{sim.}$ = 0.34) (Table 1). Notably, for both fragments, the MD-predicted poses align well with the X-ray poses (Figure 3a,b,h). Even though fragments **18**, **04**, and **K1E** also reveal $LE_{sim.}$ > 0.3, their $LE_{exp.}$ are <0.3 (Tables 1 and 2) and their simulated binding poses disagree with the experimental binding poses (Figures 3c,e,h and 4c,g).

Finally, it is interesting to note that for ligands with $LE_{sim.}$ > 0.3 (**02**, **04**, **10**, **18**, **K1E**), docking poses and the MD-predicted pose are similar to each other only in the cases of **02** and **10** (Supporting Information S1: Figure S15). For **04**, **18**, and **K1E**, on the other hand, the docking programs and MD simulations do not yield similar poses (Supporting Information S1: Figure S15). This observation may be relevant if the experimental binding mode of a fragment remained unknown in an FBDD project. Indeed, the observation that different techniques yield similar poses could be used for decision-making in a computational FBDD project.

2.6 | Assessment and scope of the approach

So far, multimicrosecond MD simulations per fragment have been used to characterize the fragment–protein interaction.[29–31] By contrast, to keep the computational demands on purpose low, we performed short MD simulations of unbiased fragment binding to endothiapepsin, derived absolute binding-free energies and binding modes from the simulation



Fragment	$RMSD_{dock./X\text{-}ray}$[Å] Glide	Autodock	Diffdock
K1F	4.2	2.0	5.2
K1A	6.6	4.4	6.7
K1E	3.0	4.2	2.7
K1B	9.3	6.5	7.1
K1C	7.8	7.8	8.3
K1D	12.4	13.0	13.4

FIGURE 6 Docked versus crystallographically resolved binding poses. Docked fragments with highest GlideXPScore (green), most favorable binding energy from the largest cluster (orange), and highest confidence (slate blue) in comparison to crystal structures (white) of fragments (**K1F** (a): PDB ID 5QBI, **K1A** (b): PDB ID 5QB6, **K1E** (c): PDB ID 5QBG, **K1B** (d): PDB ID 5QBA, **K1C** (e): PDB ID 5QBB, **K1D** (f): PDB ID 5QBC; for DOIs associated with the PDB IDs see Table 2). The fragment is always shown in stick representation. The secondary structure is displayed as cartoon representation. For clarity, all hydrogen atoms are omitted. (g) Fragment-wise root mean square deviation ($RMSD_{dock./X\text{-}ray}$) between the docked pose and the X-ray pose (including all nonhydrogen atoms), after superimposing the protein backbone atoms.

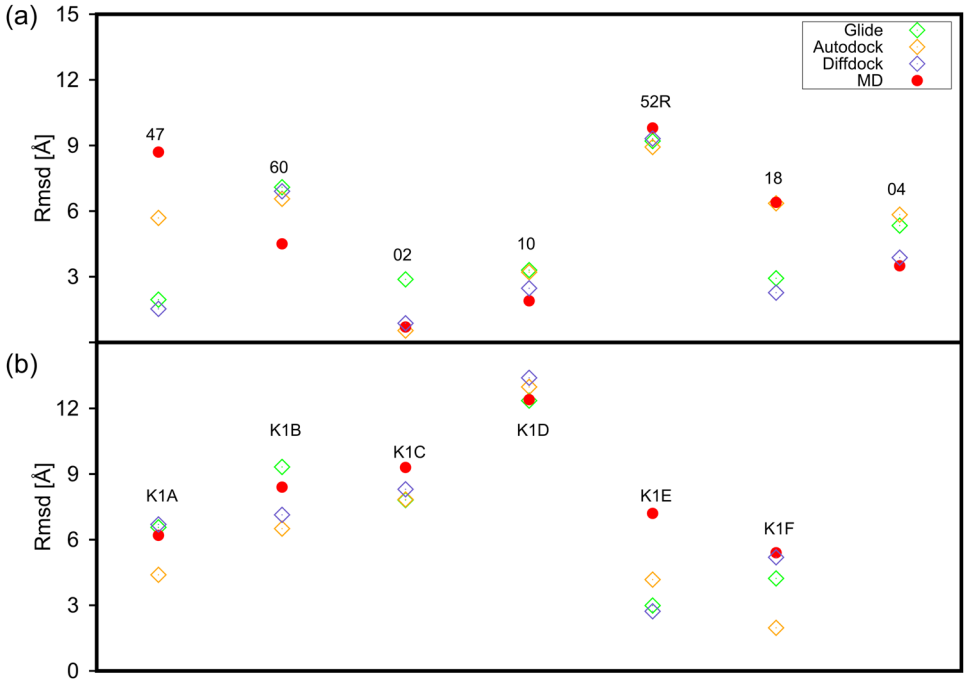


FIGURE 7 Comparison of predicted binding poses from molecular dynamics (MD) and standard docking procedures. RMSD to the crystal pose (for PDB IDs see Figures 3 and 4) of docked fragments in the training (a) and prediction set (b) with highest GlideXPScore (green), highest binding energy from AutoDock docking (orange), highest confidence from DiffDock (slate blue), as well as the highest populated cluster of the MD ensemble (red).

data and compared the simulation results with experimental data. Our primary motivation was to accomplish all MD simulations-related tasks, from system setup to the final prediction of binding energies and modes, within one business week and obtain accurate predictions.

We performed five MD simulation replicas of 600 ns length for each fragment. Within the computational workflow, the most time-consuming step is the production simulations to generate the MD ensembles. Hence, this step should be as short as possible. On the other hand, insufficient sampling would limit the applicability,[52] as our analysis strategy also requires that we record multiple binding and unbinding events of the fragments to and from the protein, which requires sufficiently long simulation times. In the present study, all production simulations were performed on a mixed-type multi-CPU-multi-GPU computing cluster and finished in less than 72 h, leaving enough time for the automated trajectory analysis using bash and Python scripts. Within 600 ns of simulation time, we observed multiple binding (between 20 and 75) and unbinding (between 1 and 16) events, indicating, in general, adequate sampling. Furthermore, we used a relatively high (50 mM) fragment concentration in our setup to increase the likelihood of fragment binding. Still, evaluating what is the best compromise between computational efficiency and adequate sampling is likely needed for every new FBDD project.

We obtained a fair agreement of simulated and experimentally measured absolute binding affinities. Note in this context that the experimental uncertainty of independent measurements of pK_i values for ligands in ChEMBL yielded a standard deviation of 0.54 log units.[53] Accurately characterizing fragment–protein interactions is considered an even more challenging task.[54,55] Due to the small affinity range of fragments, this considerably limits the maximum attainable agreement per se. Overall, the accuracy of absolute binding-free energies from multiple short MD simulations, as reported herein, is comparable with that of rigorous (relative) free energy calculations,[37–43] which are considered computationally demanding. Previous studies on the energetic characterization of fragments suggested that multi-microsecond length MD trajectories are needed to accurately describe the binding affinities,[29–31] which would limit the application of MD simulations in a drug discovery project. However, the complexity of the investigated biological target might also influence the outcome, and biological targets with well-defined and conformationally stable binding pockets should be easier to characterize by MD-based methods than highly dynamic targets.

In the case of endothiapepsin and other pepsin-like aspartic proteases, the binding kinetics depend on the conformational dynamics of the β-hairpin motif that covers the active site, the flap region.[56,57] The flap, in particular a conserved tyrosine residue,[58] covers the enzyme's active site but can adopt different conformations, including open, intermediate, or closed, depending on the presence or absence of a substrate or ligand.[59,60] We observed a highly flexible and dynamic flap region that can undergo large-scale motions, which is also known from previous studies on human cathepsin-D[61] or BACE-1,[62,63] structural homologs to endothiapepsin. In the X-ray structures of the endothiapepsin-fragment complexes, the fragments are bound to the closed endothiapepsin conformation. Note, however, that the crystals were obtained by soaking, which may bias the protein conformation. The closed endothiapepsin conformation is only sporadically observed during our simulations. Fragment-bound configurations occur preferentially in our simulations when coupled to a closed-loop conformation. The low probability of finding a closed conformation during MD simulations provides a plausible explanation for why MD simulations struggle to reproduce the X-ray conformation of the bound fragment. As we used a closed enzyme conformation for docking, this may explain why the docking performs similarly in reproducing the X-ray-like conformation.

The fragments investigated here bind weakly to endothiapepsin, as is typical for fragments. Previous studies showed that the LE of a particular compound is a key factor in determining docking performance and that poses in better agreement with the experiment are often found for fragments with higher LE.[64] Fragments with LE > 0.3 are considered promising candidates for chemical optimization toward a potential drug.[49] In our datasets, only fragments **02** and **10** have an LE > 0.3, and for both, we predicted binding poses by MD simulations and docking that are almost identical to the X-ray configurations, in agreement with the above observation.

What do our results imply for decision-making during an FBDD project, considering that the project aim is to identify promising but usually weakly binding fragments for the subsequent chemical optimization toward a potential highly affine drug? Using unbiased, short MD simulations to estimate binding affinities from the MD ensembles, $LE_{sim.} > 0.3$ was found for fragments **02**, **10**, **18**, **04**, and **K1E**, which would qualify these fragments for a binding mode prediction. The MD predicted binding poses and the poses from three different docking approaches were similar only for **02** and **10**, which we, thus, would suggest as the most promising fragments for further optimization. Remarkably, our predictions are in agreement with the ITC-based categorization of the fragments, which also identified **02** and **10** with $LE_{exp.} > 0.3$ as potential optimization candidates.

3 | CONCLUSION

Taken together, we used short and unbiased MD simulations and molecular docking to characterize fragment binding to endothiapepsin and compared the results to ITC and X-ray experiments. The fragments' absolute free energies directly predicted from the MD trajectories are in fair agreement with the experiment. Combining the MD data with binding mode predictions from molecular docking approaches helped to correctly identify the most promising fragments for further chemical optimization. Importantly, all computations and predictions were done within 5 days, suggesting that MD simulations may become a viable tool in FBDD projects. Future studies on larger fragment libraries and other targets shall be conducted to confirm the observed trends. Finally, in view of

the importance of protein conformational ensembles in protein function,[65] explicitly considering protein dynamics upon fragment binding may also be helpful to discover cryptic sites[66] or exploit binding to allosteric sites.[67,68]

4 | EXPERIMENTAL

4.1 | System preparation

To investigate the energetic and structural features of fragment binding to endothiapepsin, we performed all-atom MD simulations of free fragment diffusion around endothiapepsin.[69,70] As to endothiapepsin, we extracted Cartesian coordinates from the crystal structure PDB ID 3PCW.[22] pK_a values were calculated for all titratable amino acids using the H^{++} server[71] and PROPKA3.0 program,[72] and protonation states were assigned according to the pH of 4.6 used in experiments.[22] As to the catalytically important amino acid Asp219,[22] the calculated pK_a ranged from 3.6 to 4.9 for the H^{++} Server and from 4.5 to 9.5 for the PROPKA program, such that both species (deprotonated/ionic [net charge −1] and protonated/neutral [net charge 0]) are relevant at pH 4.6. This is also in line with previous observations that Asp219 but not Asp35 is supposed to carry a proton.[20,73] We thus prepared two protein structures, one carrying the deprotonated/ionic Asp219 (referred to as endothiapepsin^{Asp219}) and one carrying the protonated/neutral AspH219 (referred to as endothiapepsin^{AspH219}). ACE caps were connected to the N-terminal amino acids to avoid artificially charged termini. Hydrogen atoms not resolved in the crystal structure were added according to the ff99SB library.[74,75] As to the fragments, initial 3D structures were generated according to the structural formulas provided in Tables 1 and 2 using Sybyl.[76] Protonation states for a pH of 4.6 were determined using Epik.[77] The initial structures were subjected to quantum mechanical (QM) geometry optimization using Gaussian 09[78] at the HF/6-31 G* level of theory.

We aimed for a fragment concentration of 50 mM, and thus, randomly placed 16 QM-optimized fragments and the protein structure into the simulation box using packmol.[79] The systems were solvated with TIP3P water,[80] and 0.1 M NaCl was added to mimic experimental buffer concentrations. Following this procedure, we prepared 16 individual simulation systems, including a small fragment library of 14 fragments of which two existed as enantiomers, for each of the two protein systems (endothiapepsin^{Asp219} and endothiapepsin^{AspH219}). The 14 fragments originate from an in-house library and have been selected by the experimentalists, who knew the availability of experimental data, for example, X-ray complex structures, for all the fragments at that time (more details are provided in the results section). Note that we investigated only one type of fragment during one simulation and not a mixture.

As to fragments 02, 47, and 60, we analogously prepared systems for the crystallographically resolved binding poses (PDB IDs: 3PBD, 3PM4, 3PI0[22]; Supporting Information S1: Table S1). In the following, these systems are referred to as the holo structures.

4.2 | MD simulations

All MD simulations were performed with the AMBER 14 suite of programs[81] using the GPU-accelerated code of *pmemd*.[82] The ff99SB force field[74,75] was used to describe the protein dynamics, while we used the GAFF force field[83] to describe the fragments. As to the fragments, atomic partial charges were derived following the RESP procedure[84] by fitting HF/6-31 G* electrostatic potentials. The particle mesh Ewald (PME) method[85] was used to treat long-range electrostatic interactions using a direct-space nonbonded cutoff of 8 Å. Bond lengths involving bonds to hydrogen atoms were constrained using SHAKE.[86] If not reported differently, the time step for integrating Newton's equation of motion during MD simulations was set to 2 fs.

The systems were initially minimized by applying harmonic restraints with force constants of at least 5 kcal mol^{-1} Å^{-2} to all solute atoms. Applying harmonic restraints with the force constants of 5 kcal mol^{-1} Å^{-2} to all solute atoms, NVT simulations were carried out for 50 ps, during which the systems were heated to 300 K. Subsequently, we adjusted the solvent density to 1.00 g/cm^3 during NPT simulations, applying the Berendsen thermostat for temperature control (using a time constant of 0.5 ps for heat bath coupling) and the Berendsen barostat for pressure control (using a pressure relaxation time of 0.5 ps). Finally, the positional restraints on solute atom positions were gradually reduced to zero during NVT simulations, and subsequently, the systems were subjected to unrestrained NVT production simulations. Temperature control was done using the Berendsen thermostat with a time constant of 10 ps.

For the *holo* structures, we performed three independent MD runs each for 100 ns and at temperatures of 300.0, 300.1, and 300.2 K, respectively. As to the fragment diffusion simulations, we performed five independent MD runs each for 600 ns at temperatures of 300.0, 300.1, 300.2, 300.3, and 300.4 K for endothiapepsin^{Asp219} and endothiapepsin^{AspH219}, respectively, if not reported differently. This resulted in a cumulative MD simulation time of 600 ns × 16 × 5 × 2 = 96 µs. Note that in all fragment diffusion simulations, the fragments were not biased by a guiding force. Coordinates were stored in a trajectory file every 100 ps for analysis. Note that we also probed the influence of varying simulation times and a 4 fs time-step[44] for integrating Newton's equation on the quality of our results. These cases are explicitly mentioned below.

4.3 | General analyses of MD simulations

Structural analyses were performed with the program cpptraj[87] of the AmberTools14 suite.[88] If not reported differently, results are expressed as mean values ± SEM and compared using a two-sided Student's *t*-test using Microsoft's Excel® program. *p*-Values < 0.05 were considered significant. Plots and figures were generated with gnuplot[89] and pymol.[90]

4.4 | Determination of binding affinities

The binding affinities for the fragments were estimated based on the law of mass action for reversible chemical reactions directly from the MD simulations data. Therefore, we described the fragment (F) binding process to endothiapepsin (E) by the below eqation

$$E + F \rightleftarrows EF, \tag{1}$$

where E and F denote the noncomplexed endothiapsin and fragment, respectively, and EF the endothiapepsin–fragment complex. Following the law of mass action, the equilibrium-binding constant K_a and dissociation constant K_d can be described by the below equation

$$K_a = \frac{[EF]}{[E][F]} = \frac{1}{K_d}, \tag{2}$$

where [E], [F], and [EF] denote the equilibrium concentration of the respective species in molar units.[25,91] To calculate the K_a from binding simulations, we expressed Equation (2) as Equation (3)

$$K_a = \frac{P_b}{P_u[F]} = \frac{1}{K_d}, \tag{3}$$

where P_u is the number of configurations in which no fragment is bound, P_b is the number of configurations in which the fragment is bound to endothiapepsin, and [F] is the initial fragment concentration of 0.05 M. The absolute standard free energy of binding $\Delta G^0_{bind.,\,sim.}$ was then derived according to the below equation[25,91]

$$\Delta G^0_{bind.,sim.} = -k_B T \ln(K_a), \tag{4}$$

where k_B is Boltzmann's constant ($k_B = 0.001987$ kcal mol^{-1} K^{-1}), and T is the temperature ($T = 300$ K).

To categorize configurations along the trajectories (i.e., coordinates saved every 100 ps) into bound and unbound states, we calculated the center of mass (COM) distance, termed $d_{bind.}$ hereafter, between the heavy atoms of the fragments and the binding pockets S1, S2, and S3.[22] We considered a fragment bound to at least one pocket if $d_{bind.} < 6$ Å and unbound if $d_{bind.} > 10$ Å, similar to former studies to study ligand binding.[31,34] From the known X-ray structures of fragments bound to endothiapepsin,[22] we anticipate that only a single fragment can bind to endothiapepsin at one time. However, in the particular case with multiple fragments bound to endothiapsin during MD simulations, which is technically possible as we have 16 fragment molecules in our system, we only considered the fragment with the smallest overall $d_{bind.}$ as the bound fragment.

$\Delta G^0_{bind,\,sim.}$ was determined for each trajectory independently. The results are then expressed as mean $\Delta G^0_{bind,\,sim.} \pm$ SEM over $n = 10$ trajectories and compared with experimentally derived binding-free energies. We performed a one-sample t-test using the experimentally derived binding-free energies as a reference. p values < 0.05 were considered significantly different.

Finally, to evaluate the agreement between $\Delta G^0_{bind.,\,exp.}$ and $\Delta G^0_{bind,\,sim.,}$ we calculated the absolute deviation (AD) according to the below equation

$$AD = \left|\Delta G^0_{bind,sim} - \Delta G^0_{bind,exp}\right|. \tag{5}$$

4.5 | Determination of binding modes

To derive the binding mode of the fragments bound to endothiapepsin, all bound conformations were extracted from the respective trajectories and subjected to a two-step hierarchical clustering approach. First, using the RMSD of the fragment after superimposing the backbone atoms of the protein as the measure, the bound conformations were clustered into five clusters. Only the largest cluster was considered for the second step, in which the final average cluster distance ε is 2.0 Å using the RMSD criterion again. The resulting representative poses of the largest cluster were subsequently compared to the X-ray poses. The clustering was performed using the cluster command in the cpptraj[87,88] program.

4.6 | Structural analyses of the MD simulations

To characterize the structural changes of the β-hairpin loop (residues 78–83), which is also part of the binding pockets S1, S2, and S3, we calculated the no-fit RMSD of the β-hairpin loop with respect to the closed X-ray conformation of endothiapepsin^{Asp219}. Therefore, we superimposed all Cα atoms except the β-hairpin residues onto the respective closed X-ray conformation and calculated the RMSD for the Cα atoms in residues 78–83 of the β-hairpin loop without fitting. As the open conformation of endothiapepsin is still unknown, we compared this data to the differences between the open (PDB ID 2IGY[45]) and closed (PDB ID 1SME[46]) conformations of plasmepsin II, a structural homolog to endothiapepsin. Additionally, to characterize the opening and closing of the β-hairpin loop, we measured the distance between the hydroxyl-oxygen in Y79 and the indol-nitrogen in W42 and compared our results with the distances between the equivalent residues Y77 and W41 in plasmepsin II. The distance of 9.3 Å in PDB ID 2IGY provided us with an orientation value about the gap in an open aspartic protease.

4.7 | Molecular docking of fragments into the binding site of endothiapepsin

The protonation states for endothiapepsin^{Asp219} were taken from the MD simulations setup, and the Gaussian-optimized fragment conformations were used for docking. For AutoDock-4[92] docking, the Lamarckian genetic algorithm (LGA) was selected for the ligand conformational search, using a 28 × 28 × 28 3D affinity grid centered around the residues in S1, S2, S3 with a 0.375 Å grid point spacing. The final binding mode model was chosen as the one with the most favorable docking energy from the largest cluster of structurally similar

binding modes; the mutual all-atom RMSD of the binding modes was used as the cluster criterion with a cutoff of 0.5 Å. For Glide docking,[93,94] fragments were prepared in Maestro[95] using LigPrep,[96] followed by an energetic minimization employing the OPLS4 force field.[97] During docking, ionization states and the fragments' chirality were kept fixed. For the protein structure, the hydrogen atoms of endothiapepsin^{Asp219} were minimized with the OPLS4 force field.[97] The final binding mode model was chosen as the one with the highest Extra-Precision docking score.[98] In the case of Diffdock,[99] we used the web interface (https://huggingface.co/spaces/simonduerr/diffdock) with standard values. The final binding mode model was chosen as the one with the highest confidence.

ACKNOWLEDGMENTS

We thank Dr. Alexander Metz and Prof. Dr. Gerhard Klebe for their support and fruitful discussions throughout the realization of this work. Holger Gohlke gratefully acknowledges the computational support provided by the "Center for Information and Media Technology" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user IDs: HKF7, VSK33). This study was funded, in part, by Deutsche Forschungsgemeinschaft (DFG) through GRK 2158/2 (project number 270650915) to Holger Gohlke. Open Access funding enabled and organized by Projekt DEAL.

CONFLICTS OF INTEREST STATEMENT

Benedikt Frieg is now an employee of AstraZeneca. The other authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in RCSB at https://www.rcsb.org/, reference number PDB IDs. These data were derived from the following resources available in the public domain: PDB IDs Tables 1 and 2, https://www.rcsb.org/.

ORCID

Holger Gohlke http://orcid.org/0000-0001-8613-1447

REFERENCES

[1] W. P. Jencks, *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046.
[2] D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discov.* **2016**, *15*, 605.
[3] H. Jhoti, G. Williams, D. C. Rees, C. W. Murray, *Nat. Rev. Drug Discov.* **2013**, *12*, 644.
[4] M. Congreve, R. Carr, C. Murray, H. Jhoti, *Drug Discov. Today* **2003**, *8*, 876.
[5] G. M. Keserű, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray, S. D. Pickett, *J. Med. Chem.* **2016**, *59*, 8189.
[6] B. Diethelm-Varela, *ChemMedChem* **2021**, *16*, 725.
[7] I. Navratilova, A. L. Hopkins, *ACS Med. Chem. Lett.* **2010**, *1*, 44.
[8] J. E. Ladbury, G. Klebe, E. Freire, *Nat. Rev. Drug Discov.* **2010**, *9*, 23.
[9] M. Saur, M. J. Hartshorn, J. Dong, J. Reeks, G. Bunkoczi, H. Jhoti, P. A. Williams, *Drug Discov. Today* **2020**, *25*, 485.
[10] P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discov.* **2007**, *6*, 211.
[11] G. Bollag, J. Tsai, J. Zhang, C. Zhang, P. Ibrahim, K. Nolop, P. Hirth, *Nat. Rev. Drug Discov.* **2012**, *11*, 873.
[12] A. J. Souers, J. D. Leverson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, H. Ding, S. H. Enschede, W. J. Fairbrother, D. C. S. Huang, S. G. Hymowitz, S. Jin, S. L. Khaw, P. J. Kovar, L. T. Lam, J. Lee, H. L. Maecker, K. C. Marsh, K. D. Mason, M. J. Mitten, P. M. Nimmer, A. Oleksijew, C. H. Park, C. M. Park, D. C. Phillips, A. W. Roberts, D. Sampath, J. F. Seymour, M. L. Smith, G. M. Sullivan, S. K. Tahir, C. Tse, M. D. Wendt, Y. Xiao, J. C. Xue, H. Zhang, R. A. Humerickhouse, S. H. Rosenberg, S. W. Elmore, *Nat. Med.* **2013**, *19*, 202.
[13] T. P. S. Perera, E. Jovcheva, L. Mevellec, J. Vialard, D. De Lange, T. Verhulst, C. Paulussen, K. Van De Ven, P. King, E. Freyne, D. C. Rees, M. Squires, G. Saxty, M. Page, C. W. Murray, R. Gilissen, G. Ward, N. T. Thompson, D. R. Newell, N. Cheng, L. Xie, J. Yang, S. J. Platero, J. D. Karkera, C. Moy, P. Angibaud, S. Laquerre, M. V. Lorenzi, *Mol. Cancer Ther.* **2017**, *16*, 1010.
[14] J. B. Cooper, *in* Handbook of Proteolytic Enzymes, Elsevier, San Diego, **2004**, pp. 104.
[15] J. Cooper, *Curr. Drug Targets* **2002**, *3*, 155.
[16] J. Cooper, W. Quail, C. Frazao, S. I. Foundling, T. L. Blundell, C. Humblet, E. A. Lunney, W. T. Lowther, B. M. Dunn, *Biochemistry* **1992**, *31*, 8142.
[17] S. Geschwindner, L. L. Olsson, J. S. Albert, J. Deinum, P. D. Edwards, T. de Beer, R. H. A. Folmer, *J. Med. Chem.* **2007**, *50*, 5903.
[18] L. Coates, P. T. Erskine, S. Mall, R. Gill, S. P. Wood, D. A. A. Myles, J. B. Cooper, *Eur. Biophys. J.* **2006**, *35*, 559.
[19] L. Coates, P. T. Erskine, S. P. Wood, D. A. A. Myles, J. B. Cooper, *Biochemistry* **2001**, *40*, 13149.
[20] L. Coates, H. F. Tuan, S. Tomanicek, A. Kovalevsky, M. Mustyakimov, P. Erskine, J. Cooper, *J. Am. Chem. Soc.* **2008**, *130*, 7235.
[21] A. Hartman, M. Mondal, N. Radeva, G. Klebe, A. Hirsch, *Int. J. Mol. Sci.* **2015**, *16*, 19184.
[22] H. Köster, T. Craan, S. Brass, C. Herhaus, M. Zentgraf, L. Neumann, A. Heine, G. Klebe, *J. Med. Chem.* **2011**, *54*, 7784.
[23] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259.
[24] M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2016**, *55*, 9422.
[25] M. De Vivo, M. Masetti, G. Bottegoni, A. Cavalli, *J. Med. Chem.* **2016**, *59*, 4035.
[26] R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, D. E. Shaw, *Annu. Rev. Biophys.* **2012**, *41*, 429.
[27] R. O. Dror, M. Ø. Jensen, D. W. Borhani, D. E. Shaw, *J. Gen. Physiol.* **2010**, *135*, 555.
[28] Z. Cournia, B. Allen, W. Sherman, *J. Chem. Inf. Model.* **2017**, *57*, 2911.
[29] G. Martinez-Rosell, M. J. Harvey, G. De Fabritiis, *J. Chem. Inf. Model.* **2018**, *58*, 683.
[30] S. M. Linker, A. Magarkar, J. Kofinger, G. Hummer, D. Seeliger, *J. Chem. Theory Comput.* **2019**, *15*, 4974.
[31] A. C. Pan, H. F. Xu, T. Palpant, D. E. Shaw, *J. Chem. Theory Comput.* **2017**, *13*, 3372.
[32] M. M. Hann, A. R. Leach, G. Harper, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856.
[33] H. Köster, *PhD Thesis*, Philipps-University Marburg, Marburg. **2012**.
[34] N. Plattner, F. Noé, *Nat. Commun.* **2015**, *6*, 7653.
[35] M. C. Ahmed, E. Papaleo, K. Lindorff-Larsen, *PeerJ* **2018**, *6*, e4967.
[36] J. A. Pople, *Angew. Chem. Int. Ed.* **1999**, *38*, 1894.
[37] D. L. Mobley, A. P. Graves, J. D. Chodera, A. C. McReynolds, B. K. Shoichet, K. A. Dill, *J. Mol. Biol.* **2007**, *371*, 1118.
[38] D. J. Cole, I. Cabeza de Vaca, W. L. Jorgensen, *MedChemComm* **2019**, *10*, 1116.

[39] J. Wang, Y. Deng, B. Roux, *Biophys. J.* **2006**, *91*, 2798.
[40] M. Aldeghi, A. Heifetz, M. J. Bodkin, S. Knapp, P. C. Biggin, *J. Am. Chem. Soc.* **2017**, *139*, 946.
[41] L. Liang, H. Liu, G. Xing, C. Deng, Y. Hua, R. Gu, T. Lu, Y. Chen, Y. Zhang, *Phys. Chem. Chem. Phys.* **2022**, *24*, 9904.
[42] V. Gapsys, A. Yildirim, M. Aldeghi, Y. Khalak, D. van der Spoel, B. L. de Groot, *Comm. Chem.* **2021**, *4*, 61.
[43] V. Gapsys, L. Pérez-Benito, M. Aldeghi, D. Seeliger, H. van Vlijmen, G. Tresadern, B. L. de Groot, *Chem. Sci.* **2020**, *11*, 1140.
[44] C. W. Hopkins, S. Le Grand, R. C. Walker, A. E. Roitberg, *J. Chem. Theory Comput.* **2015**, *11*, 1864.
[45] C. Boss, O. Corminboeuf, C. Grisostomi, S. Meyer, A. F. Jones, L. Prade, C. Binkert, W. Fischli, T. Weller, D. Bur, *ChemMedChem* **2006**, *1*, 1341.
[46] A. M. Silva, A. Y. Lee, S. V. Gulnik, P. Maier, J. Collins, T. N. Bhat, P. J. Collins, R. E. Cachau, K. E. Luker, I. Y. Gluzman, S. E. Francis, A. Oksman, D. E. Goldberg, J. W. Erickson, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10034.
[47] L. Ferreira, R. Dos Santos, G. Oliva, A. Andricopulo, *Molecules* **2015**, *20*, 13384.
[48] L. Pinzi, G. Rastelli, *Int. J. Mol. Sci.* **2019**, *20*, 4331.
[49] I. J. P. de Esch, *Drug Discov. Today Technol.* **2010**, *7*, e147.
[50] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discov. Today* **2004**, *9*, 430.
[51] C. W. Murray, D. C. Rees, *Nat. Chem.* **2009**, *1*, 187.
[52] J. B. Clarage, T. Romo, B. K. Andrews, B. M. Pettitt, G. N. Phillips, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3288.
[53] C. Kramer, T. Kalliokoski, P. Gedeck, A. Vulpetti, *J. Med. Chem.* **2012**, *55*, 5165.
[54] J. Schiebel, N. Radeva, H. Köster, A. Metz, T. Krotzky, M. Kuhnert, W. E. Diederich, A. Heine, L. Neumann, C. Atmanene, D. Roecklin, V. Vivat-Hannah, J. P. Renaud, R. Meinecke, N. Schlinck, A. Sitte, F. Popp, M. Zeeb, G. Klebe, *ChemMedChem* **2015**, *10*, 1511.
[55] J. Schiebel, N. Radeva, S. G. Krimmer, X. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe, *ACS Chem. Biol.* **2016**, *11*, 1693.
[56] B. M. Dunn, *Chem. Rev.* **2002**, *102*, 4431.
[57] L. McGillewie, M. Ramesh, M. E. Soliman, *Protein J.* **2017**, *36*, 385.
[58] U. M. Nasir, F. Suzuki, T. Nagai, T. Nakagawa, Y. Nakamura, *Biosci. Biotechnol. Biochem.* **1999**, *63*, 1143.
[59] E. T. Baldwin, T. N. Bhat, S. Gulnik, M. V. Hosur, R. C. Sowder 2nd, R. E. Cachau, J. Collins, A. M. Silva, J. W. Erickson, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6796.
[60] J. Rahuel, *J. Struct. Biol.* **1991**, *107*, 227.
[61] O. A. Arodola, M. E. S. Soliman, *J. Cell. Biochem.* **2016**, *117*, 2643.
[62] H. M. Kumalo, S. Bhakat, M. E. Soliman, *J. Biomol. Struct. Dyn.* **2016**, *34*, 1008.
[63] S. Bhakat, P. Söderhjelm, *J. Chem. Inf. Model.* **2022**, *62*, 914.
[64] M. L. Verdonk, I. Giangreco, R. J. Hall, O. Korb, P. N. Mortenson, C. W. Murray, *J. Med. Chem.* **2011**, *54*, 5422.
[65] R. Nussinov, Y. Liu, W. Zhang, H. Jang, *RSC Chem. Biol.* **2023**, *4*, 850.
[66] D. Schmidt, M. Boehm, C. L. McClendon, R. Torella, H. Gohlke, *J. Chem. Theory Comput.* **2019**, *15*, 3331.
[67] G. Colombo, *Curr. Opin. Struct. Biol.* **2023**, *83*, 102702.
[68] G. La Sala, C. Pfleger, H. Käck, L. Wissler, P. Nevin, K. Böhm, J. P. Janet, M. Schimpl, C. J. Stubbs, M. De Vivo, C. Tyrchan, A. Hogner, H. Gohlke, A. I. Frolov, *Chem. Sci.* **2023**, *14*, 7057.
[69] H. Gohlke, U. Hergert, T. Meyer, D. Mulnaes, M. K. Grieshaber, S. H. J. Smits, L. Schmitt, *J. Chem. Inf. Model.* **2013**, *53*, 2493.
[70] B. Frieg, L. Gremer, H. Heise, D. Willbold, H. Gohlke, *Chem. Commun.* **2020**, *56*, 7589.
[71] J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, A. Onufriev, *Nucleic Acids Res.* **2005**, *33*, W368.
[72] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 525.
[73] L. Coates, P. T. Erskine, M. P. Crump, S. P. Wood, J. B. Cooper, *J. Mol. Biol.* **2002**, *318*, 1405.
[74] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.
[75] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins: Struct. Funct. Bioinf.* **2006**, *65*, 712.
[76] D. B. Boyd, *Rev. Comp. Chem.* **2009**, *6*, 317.
[77] J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin, M. Uchimaya, *J. Comput. Aided Mol. Des.* **2007**, *21*, 681.
[78] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Gaussian, Inc., Wallingford, CT, USA **2009**.
[79] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, *J. Comput. Chem.* **2009**, *30*, 2157.
[80] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[81] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668.
[82] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2013**,*9*, 3878.
[83] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
[84] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269.
[85] T. E. Cheatham III, J. L. Miller, T. Fox, T. A. Darden, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 4193.
[86] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[87] D. R. Roe, T. E. Cheatham III, *J. Chem. Theory Comput.* **2013**, *9*, 3084.
[88] D. A. Case, H. M. Aktulga, K. Belfon, D. S. Cerutti, G. A. Cisneros, V. W. D. Cruzeiro, N. Forouzesh, T. J. Giese, A. W. Gotz, H. Gohlke, S. Izadi, K. Kasavajhala, M. C. Kaymak, E. King, T. Kurtzman, T. S. Lee, P. Li, J. Liu, T. Luchko, R. Luo, M. Manathunga, M. R. Machado, H. M. Nguyen, K. A. O'Hearn, A. V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T. E. Cheatham, 3rd, D. R. Roe, A. Roitberg, C. Simmerling, D. M. York, M. C. Nagan, K. M. Merz Jr., *J. Chem. Inf. Model.* **2023**, *63*, 6183.
[89] T. Williams, L. Hecking, **2003**.
[90] Pymol. The Pymol molecular graphics system, V2.0. **2015**.
[91] A. C. Pan, D. W. Borhani, R. O. Dror, D. E. Shaw, *Drug Discovery Today* **2013**, *18*, 667.
[92] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785.
[93] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739.

[94] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750.
[95] Maestro, *Schrödinger, LLC*, New York, NY **2015**.
[96] LigPrep, *Schrödinger, LLC*, New York, NY **2015**.
[97] C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner, E. D. Harder, *J. Chem. Theory Comput.* **2021**, *17*, 4291.
[98] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177.
[99] G. Corso, H. Stärk, B. Jing, R. Barzilay, T. Jaakkola, *arXiv.2210.01776* **2022**.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: B. Schmitz, B. Frieg, N. Homeyer, G. Jessen, H. Gohlke, *Arch. Pharm.* **2024**, e2300612. https://doi.org/10.1002/ardp.202300612