Open Access

Article

Cosolvent Molecular Dynamics Applied to DPP4, DPP8 and DPP9: Reproduction of Important Binding Features and Use in Inhibitor Design

Olivier Beyens, Sam Corthaut, Sarah Peeters, Pieter Van Der Veken, Ingrid De Meester, and Hans De Winter*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: We present our efforts in computational drug design against dipeptidyl peptidase 4 (DPP4), DPP8 and DPP9. We applied cosolvent molecular dynamics (MD) simulations to these three protein targets of interest. Our primary motivation is the growing interest in DPP8 and DPP9 as emerging drug targets. Due to the high similarity between DPP4, DPP8 and DPP9, DPP4 was also included in these analyses. The cosolvent molecular dynamics simulations reproduce key ligand binding features and known binding pockets, while also highlighting interesting fragment positions for future ligand optimization. The resulting fragment maps from the cosolvent molecular dynamics are freely available for use in future research (https://github.com/UAMC-Olivier/DPP489_cosolvent_MD/). Detailed instructions for easy visualization of the fragment maps are provided, ensuring that the results are usable by both computational and medicinal chemists. Additionally, we used the fragment maps to search for the binding pockets with significant potential using an algorithmic approach combining top fragment locations. To discover novel binding scaffolds, a limited pharmacophore screening was performed, where the pharmacophores were based on the analyses of the cosolvent simulations. Unfortunately, inhibitory potencies were in the higher micromolar range, but we optimized the resulting scaffolds *in silico* using relative binding free energy calculations for future inhibitor design and synthesis.
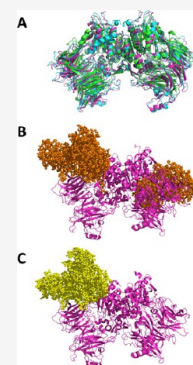
INTRODUCTION

We have applied a structural drug design methodology, namely cosolvent molecular dynamics (MD), to the drug targets dipeptidyl peptidase 4 (DPP4), 8 and 9. These three peptidases are structurally very similar serine proteases and belong to the DASH (dipeptidyl peptidase 4 activity/structure homologues) family. The three enzymes occur mainly as homodimers. Their structural similarity is demonstrated in Figure 1 (Panel A). DPP8 and DPP9 are emerging drug targets due to their putative role in cell death. Cryo-EM structures have shown that DPP9 sequesters the inflammasome sensors NLRP1 (nucleotide-binding oligomerization domain, leucine-rich repeat and pyrin domain containing protein 1) and CARD8 (caspase recruitment domain-containing protein 8) by forming protein−protein complexes.[1−3] These protein−protein complexes are visualized in Figure 1 (Panels B,C). DPP9 inhibition activates the inflammasome pathways, leading to pyroptotic cell death. As nonselective DPP8/9 inhibition leads to pyroptosis in acute myeloid leukemia (AML) cell lines, DPP9 is a potentially interesting drug target for AML.[4,5] CARD-8 mediated pyroptosis induced by DPP9 inhibition has also been observed in HIV-1 infected cells, further enhancing the interest in DPP9 as an effective drug target.[6,7] Whether DPP8 also sequesters the NLRP1 and CARD8 inflammasome sensors in a similar manner is currently unknown as no such structures are available, but similar interactions are not inconceivable due to their high similarity. Additionally, DPP8 inhibition is linked to apoptotic

cell death in multiple myeloma cell lines.[8,9] In our opinion, these discoveries warrant further investigation into DPP8 and DPP9 as drug targets. Further information with regards to the natural substrates of DPP8 and DPP9 can be found in specialized reviews.[10,11]

Contrary to DPP8 and DPP9, DPP4 is a clinically validated drug target. Four FDA approved DPP4 inhibitors (saxagliptin, alogliptin, linagliptin and sitagliptin) are available for treatment of Type 2 diabetes. The EMA has approved a fifth DPP4 inhibitor, namely vildagliptin, for clinical use in Type 2 diabetes. DPP4 inhibitors control blood glucose levels by elongating the half-life of the incretin hormones GLP-1 and GIP.[14−17] Given that DPP4 cleaves the incretin hormones, DPP4 inhibition will elongate the half-life of GLP-1 and GIP. Because DPP4 is a clinically validated drug target, novel insights into the important ligand binding features of DPP8 and DPP9 are most interesting. We did include DPP4 in our analyses due to the previously highlighted high similarity with DPP8 and DPP9. DPP4
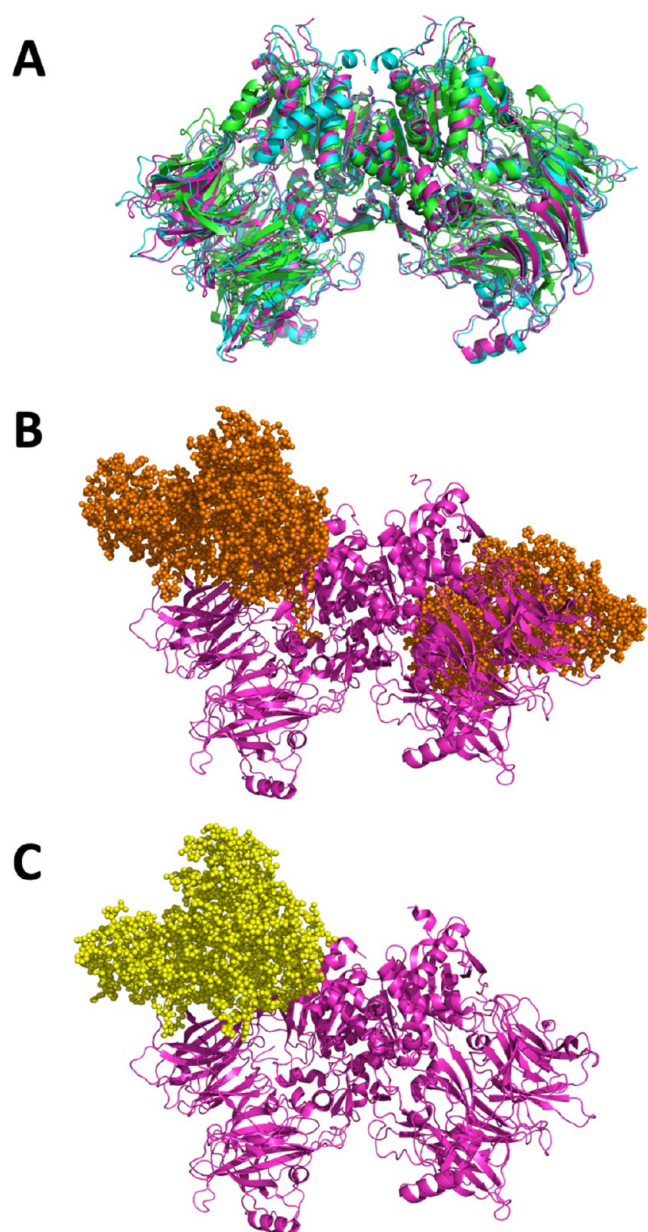
Figure 1. Panel (A) Protein structure overlap between homodimers of DPP4 (green cartoon, PDB code: 6B1E[12]), DPP8 (cyan cartoon, PDB code: 6EOP[13]) and DPP9 (magenta cartoon, PDB code: 6EOR[13]). High structural similarity is observed between DPP4, DPP8 and DPP9. Panel (B) Cryo-EM structure of the DPP9-NLRP1 complex (PDB code: 6X6A[1]). NLRP1 is represented as orange spheres. Panel (C) DPP9-CARD8 complex (PDB code: 7JKQ[3]). CARD8 is represented as yellow spheres. Note that an NLRP1 tail inserts in the DPP9 side opening to the active site, but that such interactions do not form in the DPP9-CARD8 structure.

selectivity remains a challenge in the design of DPP8/9 inhibitors, and henceforth selectivity descriptors are of interest.

Currently described DPP8/9 inhibitors that show no significant affinity toward DPP4 mainly comprise isoindoline-based scaffolds. An overview of the inhibitor structures is given in Figure 2. The binding poses of the inhibitors in Figure 2 are shown in Figure S1 (Supporting Information), overlapped with the pocket identifiers of DPP8 and DPP9. The first DPP8/9 inhibitors with selectivity against DPP4 were allo-*ile*-isoindoline (**1**) and 1G244 (**2**).[18−20] Further inhibitor design started from

these lead structures, where 1G244 modifications led to small improvements in the DPP8 over DPP9 selectivity.[21] A third DPP8/9 inhibitor scaffold was discovered by our research group by changing the cyanopyrrolidine moiety of the DPP4 inhibitor vildagliptin to isoindoline.[22] Follow-up research further explored the structure activity relationship of this line of compounds.[23] Compound **3** (UAMC-4918) was discovered this way. Compound **3** has a previously unseen DPP9 over DPP8 selectivity with a selectivity index of ∼180 (DPP9/8), but unfortunately this molecule has poor pharmacokinetic properties. Researchers from Merck discovered IceD-2 (**4**), another isoindoline containing compound.[6] Iced-2 is a highly potent DPP9 inhibitor with significant but less pronounced DPP9-to-8 selectivity. The last important molecule is a 4-oxo-β-lactam inhibitor (**5**), which has the highest DPP8 over DPP9 selectivity (SI DPP8/9 = 21) known to date.[24] We are of the opinion that this inhibitor overview highlights the limited scaffold diversity in DPP8 and DPP9 inhibitors, and therefore we believe that more scaffold diversity for DPP8/9 inhibitors is desirable.

The research presented in this work aims to gain further knowledge in the important binding features of the known ligands, to explore possible allosteric binding pockets and to discover novel ligand scaffolds. To achieve these three purposes, cosolvent MD is a suitable technology. To discover novel scaffolds, the results from the cosolvent MD are combined into suitable queries for subsequent pharmacophore screening.

The idea behind cosolvent MD is to calculate the preferred binding locations of small organic fragments, also called probes, on a target of interest. These preferred locations can be calculated in an *a posteriori* analysis of MD simulations in which the solvent is a water-ions-probes mixture instead of the standard water-ions mixture, hence the term cosolvent simulations. Cosolvent MD has been used for (allosteric) binding pocket identification and to explain ligand binding modes in earlier research.[25−29] An elaborate review of available literature in cosolvent MD has been written by Ghanakota and Carlson.[30]

Significant attention has also been focused on the development of pharmacophore hypotheses based on the favorite fragment locations in cosolvent MD.[31−33] In 'SILCS-Pharm',[33] the cosolvent-derived pharmacophore hypotheses generally outperformed other docking screening approaches. A variety of analysis methodologies for cosolvent MD simulations is freely available, such as the Cosolvent Analysis Toolkit (CAT),[34] DrugGUI[35] or Probeview.[36] This work builds further upon an analysis methodology by Bakan et al.[35] This analysis methodology makes use of the concept of interaction spots. Interaction spots can be understood as the central voxels within probe density areas. The interactions spots are then combined by the following algorithmic approach:

(1) Identification of the voxel with the lowest Grid Free Energy (GFE), where the GFE is a probe density scoring metric. This voxel will act as seed.

(2) The next lowest interactions spot within 6.2 Å is added to the "pocket".

(3) Step 2 is repeated until the pocket is comprised of a desired number of spots (usually seven or eight). An addition of an interaction spot to that pocket is rejected if the effective charge would become lower than − 2*e*.

The main modification added in this work is the implementation of a backtracking search to identify the optimal combinations instead of the greedy search performed by Bakan

Allo-ile-isoindoline
1

1G244
2

UAMC-4918
3

ICeD-2
4

4-oxo-β-lactam inhibitor 12
5
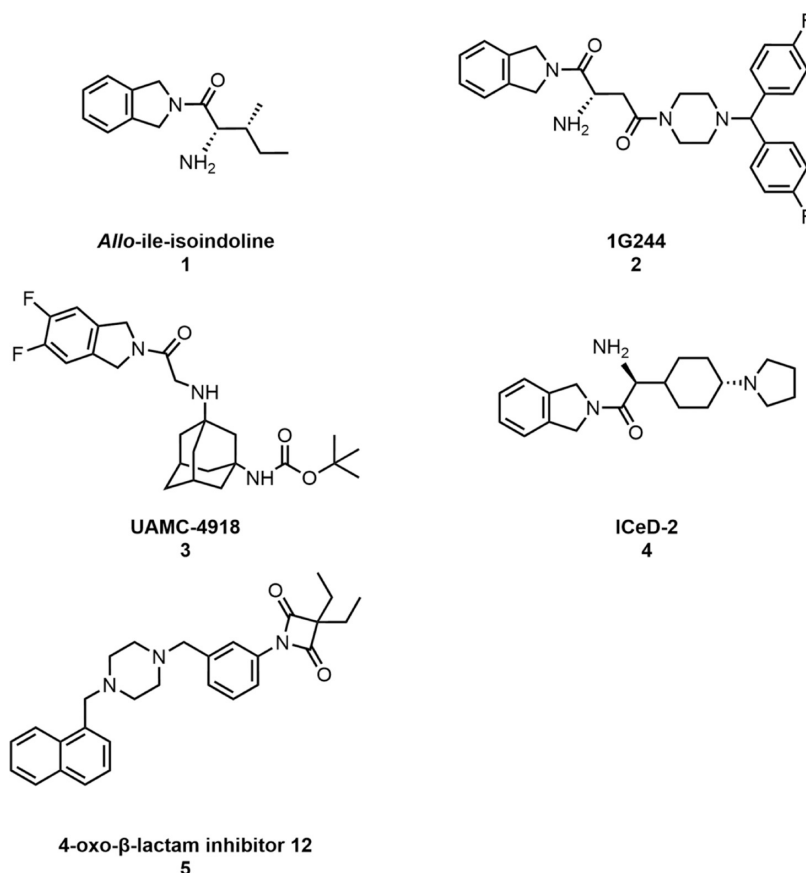
**Figure 2.** Structures of selected DPP8/9 inhibitors.[6,18−20,23,24]

et al.[35] In addition, hydrophobic probes were also added to the analyses. The full implementation of our methodology for combining the probe densities can be found in the methodology section.

The application of the cosolvent MD simulations serves the purpose of investigating three research objectives: (1) gain insight into the important binding features, (2) exploration of alternative binding pockets, and (3) novel scaffold discovery. The main targets are DPP8 and DPP9, but due to their high similarity we are also interested in DPP4, as we believe insights into DPP4 could be useful for the design of selective inhibitors for DPP8/9.

METHODOLOGY

Cosolvent Molecular Dynamics. Cosolvent MD simulations were performed on the three targets of interest (DPP4, DPP8 and DPP9) with three different cosolvent mixtures. The cosolvent mixtures are summarized in Figure 3. The first cosolvent mixture is identical to the mixture used by Bakan et al.,[35] and is composed of the water-soluble organic fragments isopropanol, isopropyl ammonium, acetate and acetamide. The second cosolvent mixture is a mixture between water and isobutane, while the third mixture is composed of water and benzene. The Plumed Automatic Restraining Tool (PART)[37] is used to prevent hydrophobic aggregation in the second and third cosolvent mixture.

The protein structures of DPP4 (PDB: 6B1E[12]), DPP8 (PDB: 6EOP[13]) and DPP9 (PDB: 6EOR[13]) were prepared in an elaborate procedure. The full homodimer structure was used and all ligands and crystal waters present were removed. The
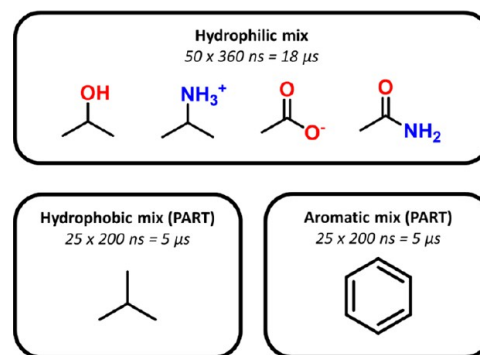
Figure 3. Overview of the three different cosolvent mixtures applied on all three targets. The cosolvent probes in the first mixture are isopropanol, isopropyl ammonium, acetate and acetamide. The cosolvent molecule in the hydrophobic mixture is isobutane, while benzene is the probe in the aromatic mixture. In the hydrophobic and aromatic mixes restraints from the Plumed Automatic Restraining Tool (PART)[37] were used to prevent hydrophobic aggregation. The listed timings are the lengths of the MD simulations.

mutation in amino acid position 36 in the DPP4 structure was restored to the original amino acid. Missing loops in the proteins were modeled using the ModLoop webserver.[38,39] Optimal flips of histidine, asparagine and glutamine were estimated using Molprobity and then checked manually.[40] Amino acid protonation states were calculated using PROPKA and H++ and then analyzed manually.[41−45] If two flips or protonation states were both likely to occur, the first state was assigned in the first monomer and the second state in the second monomer. The target protein was placed in a cubic box with periodic boundary

Table 1. Overview of the System Compositions (Number of Molecules *n* and Concentration *C*) for the Hydrophilic Mixture, Hydrophobic Mixture and Aromatic Mixture[a]

	DPP4		DPP8		DPP9	
	n	*C* (M)	*n*	*C* (M)	*n*	*C* (M)
hydrophilic mix						
isopropanol	3374	1.52	4338	1.53	3856	1.52
isopropyl ammonium	476	0.21	612	0.22	544	0.21
acetate	476	0.21	612	0.22	544	0.21
acetamide	476	0.21	612	0.22	544	0.21
sodium ions	22		44		25	
water	96,040		123,480		109,760	
hydrophobic mix						
isobutane	515	0.23	657	0.23	600	0.23
sodium ions	22		44		25	
water	111,755		142,569		130,200	
aromatic mix						
benzene	515	0.23	657	0.23	600	0.23
sodium ions	22		44		25	
water	111,755		142,569		130,200	

[a]Concentrations were determined from the box volumes after the NPT equilibrations (averaged over the replicates). Note that we opted for a constant number of water molecules per probe molecule similar to the reference publication,[35] rather than a constant concentration relative to the box size. The number of water molecules per probe molecule was set to 20 for the hydrophilic simulation and to 217 for the hydrophobic and aromatic mix. In the hydrophilic simulation, 70% of the total probe count is composed of isopropanol probes, while the isopropyl ammonium, acetate and acetamide probe types each represent 10% of the probe count.

conditions (PBC), where the principal axis of the protein was placed parallel with an axis of the box. The box size was chosen in such a way that any protein element was at least 12 Å away from the edge of the box. Initial box sizes were then slightly enlarged further, if required, to equalize probe concentrations across the DPP4, DPP8 and DPP9 systems.

The hydrophilic mixture (isopropanol, isopropyl ammonium, acetate and acetamide) was simulated using GROMACS 2021.3[46,47] using a time step of 2 fs. Fifty different replicas were simulated for the hydrophilic mixture. The chosen force field was CHARMM36m (July 2021 version), and the probes were parametrized using CGenFF version 4.6.[48−52] The force field cutoff was set to 12 Å with a smooth switching to zero between 10 and 12 Å. The Particle Mesh Ewald (PME)[53] methodology was used to treat the electrostatics, where the short-range cutoff for electrostatics was set to 12 Å. The probes were assigned different random starting locations in each of the replicas, after which the remainder of the box was filled with neutralizing sodium ions and original TIP3P waters.[54] Composition of the systems are shown in Table 1. These probe concentrations are in accordance with the reference publication of the hydrophilic mixture.[35]

After assigning the probe locations, each of the 50 different replicas was minimized in two phases. In the first phase a maximum of 50,000 steps of steepest descent were performed where the waters were modeled as flexible waters. The steepest descent step size was set to 0.01 Å. In the second minimization phase, LINCS constraints[55] were applied to bonds involving hydrogen, water flexibility was removed and the step size was increased to 0.1 Å. LINCS bond constraints to bonds involving hydrogen were also used in all subsequent simulations. After minimization, an equilibration was performed in three phases, applied separately to each replica: an NVT equilibration with protein heavy atom restraints, an NPT equilibration with protein heavy atom restraints and an NPT equilibration without restraints. The NVT equilibration of each replica had a length of 1 ns. Temperature coupling was performed using the V-

rescale thermostat[56] with a reference temperature of 300 K and a time constant of 0.1 ps. Protein and nonprotein elements were coupled separately. Replicas were equilibrated in the NPT ensemble for a total of 2 ns, split evenly between the restrained and unrestrained simulation. Pressure control was performed using the C-rescale barostat[57] with a reference pressure of 1.0 bar and a time constant of 2.0 ps. Actual production runs lasted 360 ns per replica, leading to a total simulation time of 18 μs for each target.

The methodologies for the hydrophobic and aromatic mixtures were very similar to that of the hydrophilic mixture. System compositions can be found in Table 1. The first and foremost difference with the hydrophilic mixture is the use of intermolecular repulsion potentials to prevent lipophilic aggregation. The methodology to apply intermolecular repulsion potentials was based on PART[37] in combination with Gromacs 2021.3 patched with PLUMED 2.7.2.[58−60] The default PART repulsion potentials for the production runs were used, but for the NVT equilibrations a softer potential was selected with the following parameters: $k \approx 0.02$ kcal mol^{-1} Å^{-2}, $a = 8.0$ Å, $s = 1.0$, $e = 2.0$. The softer potentials were used to avoid large forces due to probes inserted close to each other in the starting structure. The distance calculations were based on the central carbon of isobutane and on the center of mass of all benzene atoms. Additionally, only 25 replicas per target for the hydrophobic and aromatic mixtures were run and the length of each production run was 200 ns. Consequently, the total simulation time for each target-hydrophobic and target-aromatic mixture was 5 μs. The total production simulation time across all targets and all mixtures was 84 μs.

Calculation of Probe Densities and Binding Pockets. To calculate the probe densities in DPP4, DPP8 and DPP9, all trajectories were aligned onto a single reference frame. Alignment to the reference template was performed using MDAnalysis.[61,62] Subsequently, the MDAnalysis DensityAnalysis class was used to calculate probe densities. Each probe was binned on a grid composed of cubic voxels with a width of 1 Å,

and probe densities were then converted to grid free energies using eq 1

$$\mathrm{FE} = -RT \ln\left(\frac{n}{n_{\mathrm{expected}}}\right) \qquad (1)$$

In eq 1, R represents the ideal gas constant, T is the temperature, n is the calculated density for the grid voxel under study and n_{expected} is the expected density, calculated by dividing the total number of atoms of the atom group in the box by the average box volume throughout the simulation. Resulting maps were visualized using PyMol (Version 2.4.1, Schrödinger LLC) and Vida (Version 5.0.4, OpenEye Scientific[63]).

Probe densities were combined into predicted pockets using the concept of combining interaction spots. For all probe types, except benzene, the position of the central carbon atom was taken as the center for the density calculations, while for benzene this was the center of mass. The interaction spots were determined in an algorithmic way, based on prior research.[35] Each probe type was treated separately, and a list of all voxels was collected with their corresponding GFE values. Interaction spots were calculated as follows:

(1) Make two copies of the list of all voxels. The first copy will be immutable and will contain all original voxels ("original list"), the other copy will be dynamic ("dynamic list") and subject to changes in the next steps.

(2) Pick the voxel with the best (most negative) GFE score in the dynamic list of voxels;

(3) Check whether this voxel has a lower GFE than its neighboring voxels (note that this criterium was checked against the original list):

    a. If not, then remove the voxel from the dynamic voxel list;

    b. If yes, then add this voxel to the interaction spots and remove all voxels within the interaction radius

Table 2. Interaction Radii Used in the Calculation of the Interaction Spots[a]

probe type	used interaction radius (Å)
isopropanol	2.60
isopropyl ammonium	2.68
acetate	2.35
acetamide	2.27
isobutane	2.60
benzene	2.50

[a]Interaction radii of isopropanol, isopropyl ammonium, acetate and acetamide are based on the values used by Bakan et al.[35]

(see Table 2) from the dynamic list of voxels to be considered;

(4) Return to step 2, unless the next voxel has a GFE > −1.5 kcal/mol

Note that this procedure is similar but not identical to that of Bakan et al.[35] For example, multiple interaction spots from different probe types for the same voxel were allowed in our work if different probe types meet all criteria.

Searching for clusters of probe densities, which is of interest for studying the binding pockets, is a difficult task to perform manually. When less stringent density cut-offs are used, many density regions can appear that make manual analysis difficult.

An algorithmic approach is henceforth more suited for this problem. The algorithmic approach used was based on prior work,[35] but hydrophobic probe locations and a backtracking algorithm to ensure optimal pocket compositions were added to the earlier methodology. The algorithmic approach to combine the interaction spots into pockets was as follows:

(1) Take as input the list of interaction spots;

(2) Pick the interaction spot with the lowest GFE score as pocket seed;

(3) Keep adding interaction spots within 6.2 Å of the interaction spots currently in the pocket. Backtrack over all options for combining interaction spots. Apply the following rules:

    a. A pocket can contain no more than 7 interaction spots;

    b. Two interaction spots need to be at least 2 Å apart;

    c. The sum of the number of interaction spots from the hydrophobic or aromatic type can be at most 4;

    d. A pocket can contain maximum one interaction spot of acetate or isopropyl ammonium.

(4) Pick the pocket with lowest (most negative) sum of GFE scores and store this pocket for this seed. Remove the interaction spots forming the pocket. Start again with step two to look for the next pocket using the next seed spot;

(5) Rank all pockets according to the summed GFE score.

Compared to the original implementation of this algorithm,[35] a number of improvements were made. First, information about hydrophobic probes was added. Second, a recursive search instead of a greedy search was used when combining interaction spots into pockets. Third, using the revised concept of interaction spots ensures that only the strongest probe binding locations are considered. Finally, the stringent charge requirements mimic the protonation status of known DPP inhibitors.

In order to build a pharmacophore model, the locations of the central carbon atoms alone do not provide sufficient information. While considering the center of the fragments works well to identify clusters of fragment densities, pharmacophore models require not only the locations of the specific functional groups, but also the directionality of pharmacophore features for hydrogen bonding groups. For this purpose, individual atom densities of the functional groups were also taken into account. An overview of the atoms considered for the density calculations is provided in Table 3.

Pharmacophore Screening. Selected fragment locations were converted to PDB format and imported into Maestro (Schrödinger LLC). The Phase[64] pharmacophore screening tool

Table 3. Overview of the Atoms Used in Density Calculations[a]

probe type	atoms considered for density analysis
isopropanol	central carbon, hydroxyl oxygen, hydroxyl hydrogen
isopropyl ammonium	central carbon, nitrogen atom, polar hydrogens
acetate	central carbon, oxygen atoms
acetamide	central carbon, oxygen atom, nitrogen atom, polar hydrogens
isobutane	central carbon, noncentral carbon atoms
benzene	center of mass, individual carbon atoms (ring orientation)

[a]Calculating the densities of additional atoms besides the central carbon enables the study of the preferred orientation of the probe.
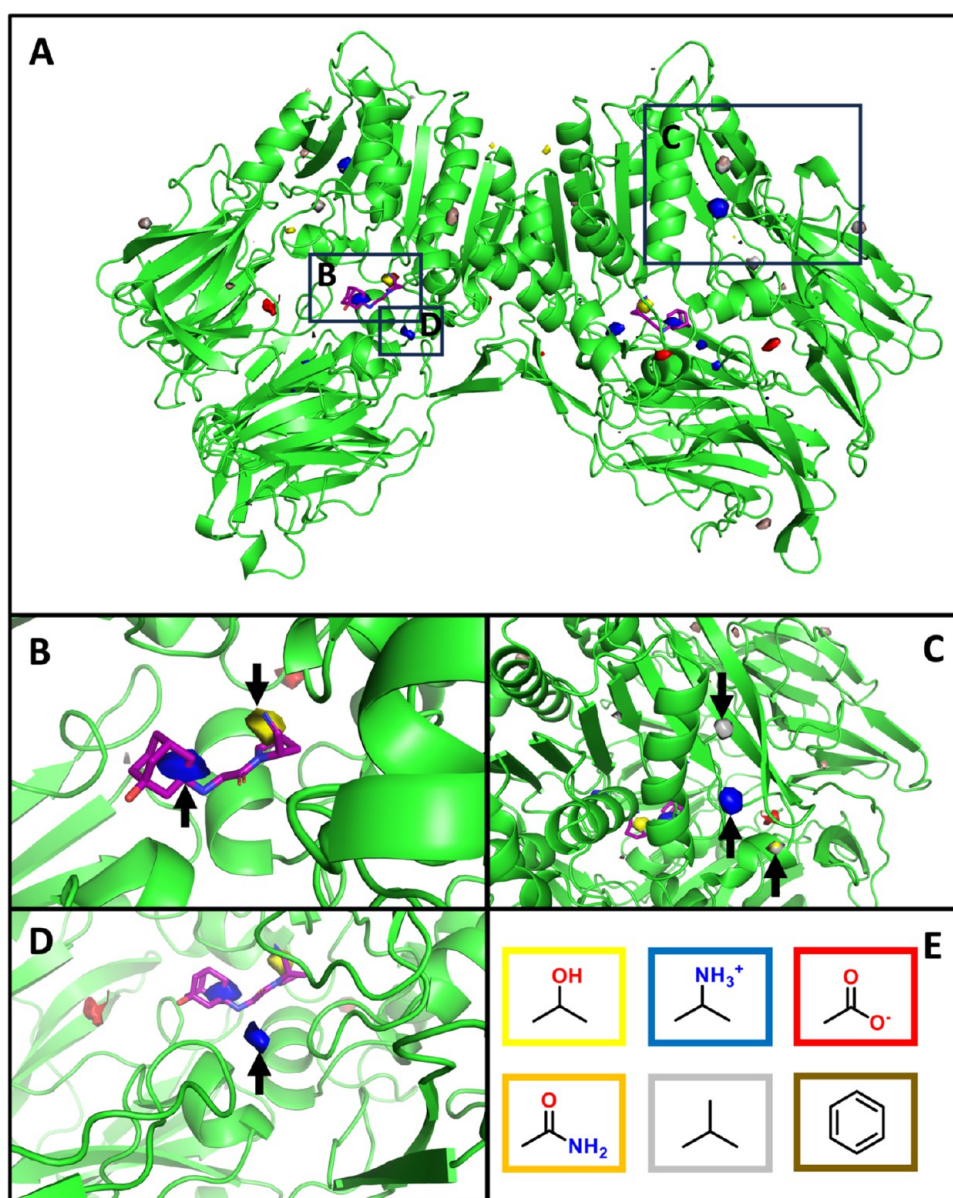
Figure 4. Density analysis for DPP4. The following cut-offs were used: 0.03 Å^{-3} for isopropanol, isobutane and benzene; 0.02 Å^{-3} for isopropyl ammonium acetate and acetamide. Panel (A) shows the density locations in the full DPP4 structure (green), for visualization purposes overlaid with the cocrystallized inhibitor vildagliptin (purple). Boxes are used to highlight the approximate locations of Panels (B–D). Panel (B) highlights the densities in the active site. A high-scoring isopropanol density close to the catalytic serine can be observed. Additionally, there is a strong isopropyl ammonium density near the secondary amine of vildagliptin. Panel (C) shows a second cluster of high scoring densities near a known glycosylation site (Asn 520). Panel (D) zooms into the side entry channel of DPP4, where a strong isopropyl ammonium density can be observed. The color codes (isopropanol: yellow, isopropyl ammonium: dark blue, acetate: red, acetamide: orange, isobutane: gray and benzene: brown) are shown graphically in Panel (E). In all panels the arrows indicate the discussed densities.

was used for virtual screening of the Chembridge, Enamine Hit Locator and Asinex databases (approximately 1.8 million compounds total). Database compounds were minimized, and their protonation states were calculated using Ligprep. Following minimization, Phase databases were created and a maximum number of 100 conformations per compound were generated, where the energy of each conformation was not allowed to be more than 3.82 kcal/mol higher than the minimized structure.

Screening results for the active site models were filtered to contain a primary or secondary amine, as the current active site inhibitors for DPP4, DPP8 or DPP9 contain this moiety. Compound **18** (see Supporting Information) does not satisfy

this criterion, as **18** was a manual addition to the selected compound pool based on a promising binding posing observed in early test runs of the protocol. After performing the pharmacophore screen, the thirty top scoring compounds according to the PhaseScreenScore were evaluated manually to determine whether the binding pose was reasonable. This ranking procedure was repeated for every pharmacophore model.

This procedure resulted in 24 compounds that were acquired and submitted for biochemical evaluation. Prior to this, LC-MS measurements were used to examine compound purities (data not shown).
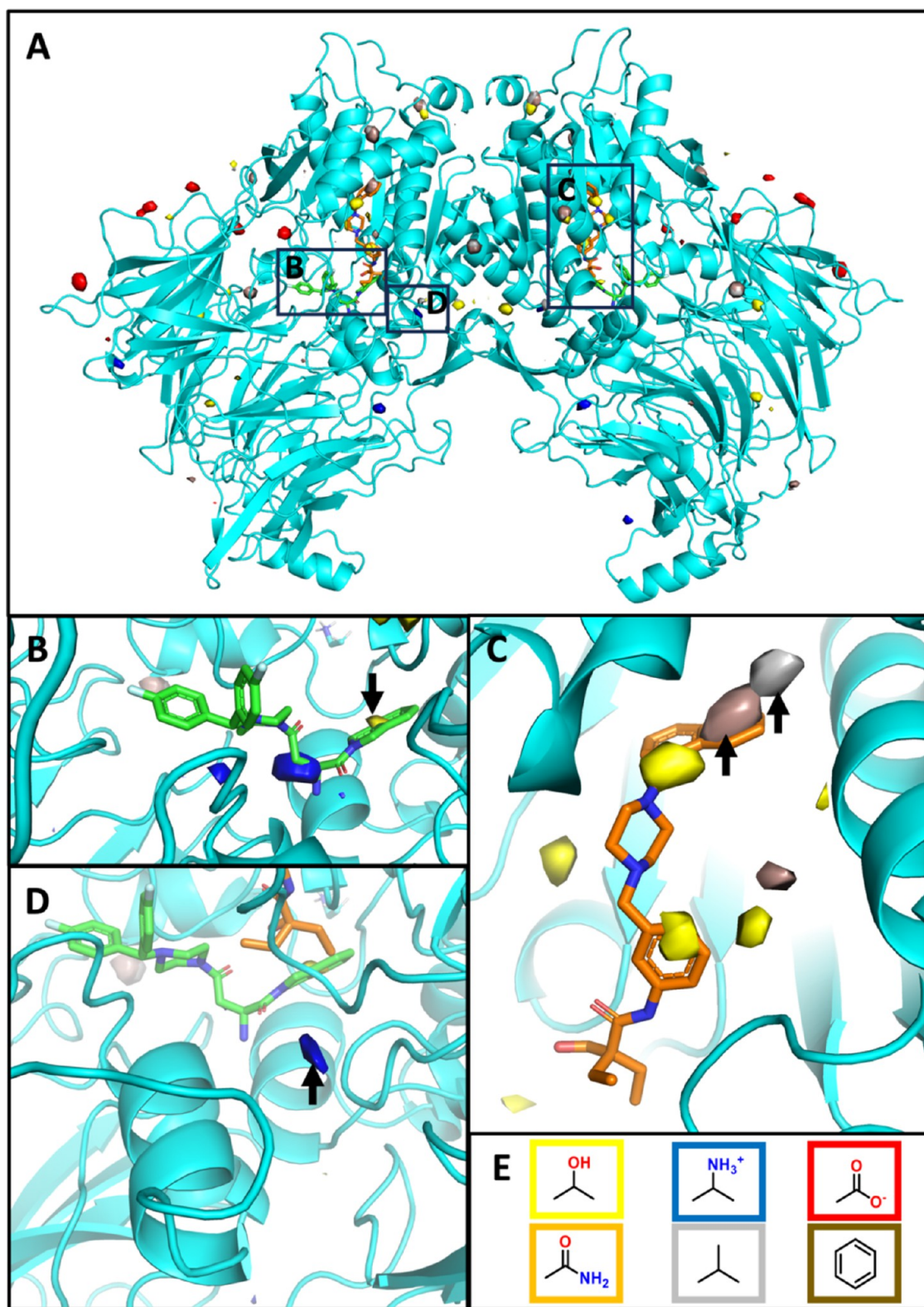
Figure 5. Overview of the density analysis for DPP8 (cyan). The following cut-offs were used: 0.03 Å^{-3} for isopropanol, isobutane and benzene; 0.02 Å^{-3} for isopropyl ammonium acetate and acetamide. Panel (A) shows the density locations in the full DPP8 structure in complex (solely for visualization purposes) with the cocrystallized inhibitor 1G244 (green) and the 4-oxo-β-lactam inhibitor 12. Panel (B) shows a close-up of the active site with 1G244. A high scoring isopropanol density is observed close to the catalytic serine and a high scoring isopropyl ammonium density is located near the primary amine of the ligand. Note that the density cutoff for isopropyl ammonium was lowered to 0.01 Å^{-3} in Panel (B). Panel (C) highlights the location of a cluster of high-scoring densities in the binding pocket of the 4-oxo-β-lactam inhibitor 12. Panel (D) shows a positively charged probe density located in the side entry channel between the α/β-hydrolase domain and the β-propellor domain. Panel (E) shows the color legend (isopropanol: yellow, isopropyl ammonium: dark blue, acetate: red, acetamide: orange, isobutane: gray and benzene: brown) graphically. In all panels the arrows indicate the discussed densities.
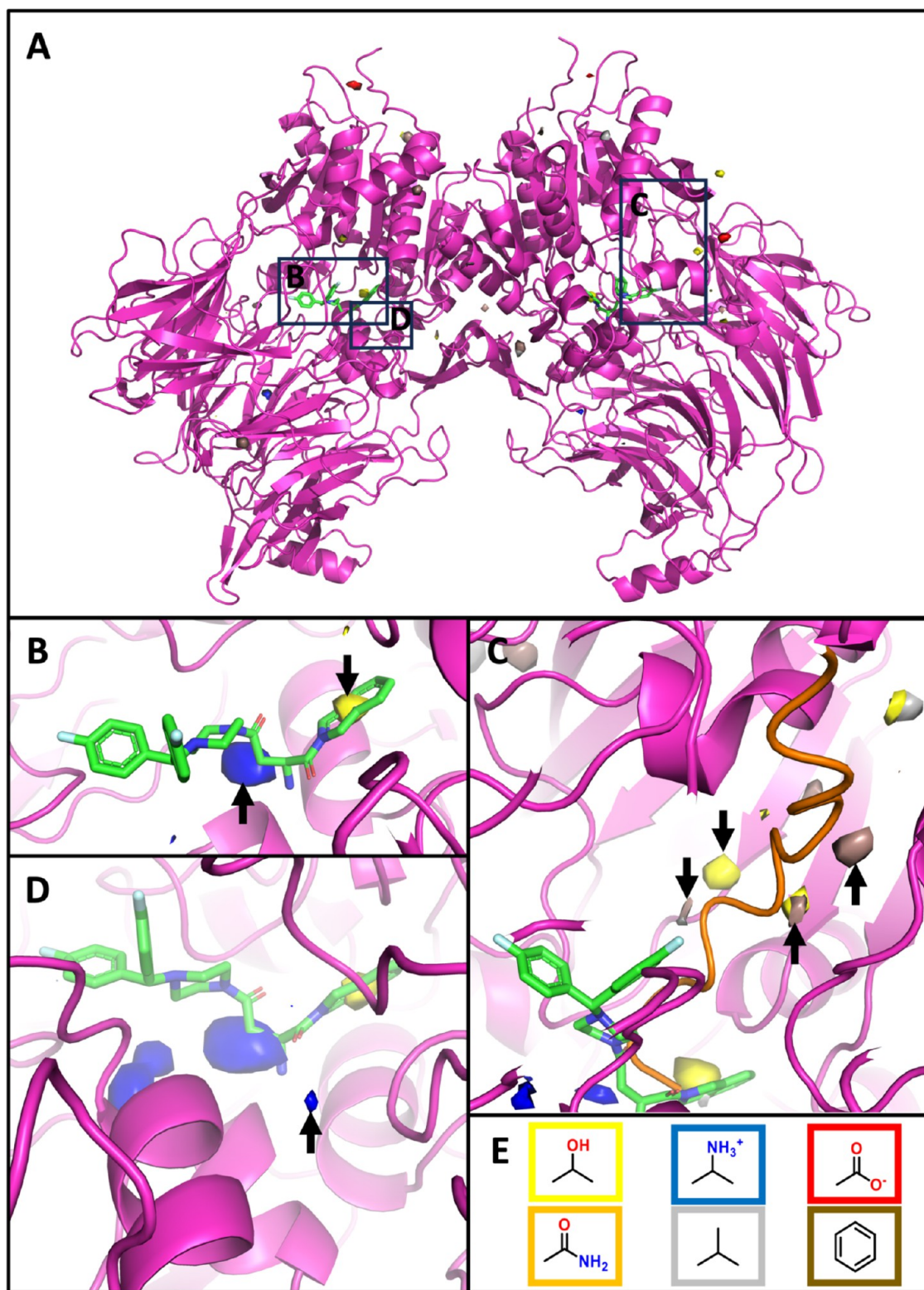
Figure 6. Overview of the density analysis for DPP9 (magenta). The following cut-offs were used: 0.03 Å^{-3} for isopropanol, isobutane and benzene; 0.02 Å^{-3} for isopropyl ammonium, acetate and acetamide. Panel (A) shows the density locations in the full DPP9 structure in complex with the cocrystallized inhibitor 1G244 (green). Panel (B) zooms to the cocrystallize inhibitor (1G244) and highlights that a high scoring isopropanol density is observed in the active site, and that a high-scoring isopropyl ammonium is located near the primary amine of 1G244. Note that in Panel (B) the isopropyl ammonium cutoff was reduced to 0.01 Å^{-3}. Panel (C) zooms to the side entry channel in DPP9 and highlights high scoring densities overlapping with the tail of NLRP1 inserting toward the active site. Note that in Panel (C) all cut-offs were reduced to 0.01 Å^{-3} except for isopropanol, for which the cutoff was reduced to 0.02 Å^{-3}. Panel (D) shows a positively charged probe density located in the side entry channel of DPP9. Note that in Panel (D) the isopropyl ammonium cutoff was reduced to 0.005 Å^{-3}. Panel (E) shows the probe color legend (isopropanol: yellow, isopropyl ammonium: dark blue, acetate: red, acetamide: orange, isobutane: gray and benzene: brown) graphically. In all panels the arrows indicate the discussed densities.

Biochemical Evaluation. DPP4 was purified from human seminal plasma, as described previously.[65] Gateway-entry clones for human DPP8 and DPP9 were purchased from Dharmacon (Accession numbers DQ891733 and DQ892325 respectively). Recombinant human DPP8 and DPP9 was expressed using the N-terminal BaculoDirect insect cell expression system (Invitrogen) in Sf9 cells. Both proteins were purified as described by De Decker et al.[66]

Screening of the compounds was performed by kinetic measurement of the hydrolysis of the chromogenic substrate Ala-Pro-para-nitroanilide at 405 nm. Substrate concentrations were chosen to approximate the K_m value of the respective enzyme: 25 μM (DPP4), 300 μM (DPP8) or 150 μM (DPP9). The buffer for the enzymatic reaction was 0.05 M HEPES-NaOH at pH 7.4 with 0.1% Tween-20, 0.1 mg/mL BSA and 150 mM NaCl. Compounds were tested in duplicate wells at 10 and 50 μM with a final reaction volume of 100 μL. Each screening was measured in duplicate. Inhibitors were preincubated with the enzyme for 15 min at 37 °C. Subsequently, the substrate was added, and the enzymatic reaction was measured in an Infinite F200 Pro (Tecan Benelux) for at least 10 min at 37 °C. Magellan software (Tecan Benelux) was used to reduce the kinetic data into initial velocities (v_0, initial velocity in the absence of inhibitor, v_i, initial velocity in the presence of inhibitor). Since some compounds showed some degree of color, inhibitor blanks (inhibitor with buffer and substrate) were measured and taken into account. For every assay measurement, a positive control was tested. The positive control was sitagliptin (1 and 0.1 μM) for DPP4 and 1G244 (1 and 0.1 μM) for DPP8 and DPP9.

Alchemical Binding Free Energy Calculations. Based on the results of the biochemical evaluation (see below), two compounds were selected for further optimization using relative binding free energy (RBFE) calculations. New inhibitors based on the two selected compounds were designed by overlap with the cosolvent MD maps and by a qualitative molecular sculpting procedure based on the target structure. RBFE calculations to estimate the affinity of the newly designed compounds were performed using nonequilibrium work calculations,[67] where the hybrid topologies were generated using PMX.[68] One ligand per homodimer was simulated. The protein−ligand systems were placed in the center of a dodecahedral box, with the size of the dodecahedron chosen in such a way that any protein element was at least 12 Å away from the box edge. Similarly, the ligand systems were also placed in a dodecahedral box of which the size was chosen in such a way that that any ligand element was at least 12 Å away from the box edge. Each box was solvated with TIP3P water[54] and sodium and chloride ions were added to achieve neutrality at an ion concentration of 0.15 M. The Amber99SB-ILDN force field modified by PMX[68,69] and Joung and Cheatham ion parameters[70] was used. Simulations were run using the GROMACS v2023.3 engine.[46,47] Ligand topologies were prepared using ACPYPE v. 2022.6.6,[71] where ligands were parametrized via GAFF2.[72] The force field cutoff was set to 11 Å, with a smooth switching to zero starting at 10 Å. Electrostatics were handled using the PME methodology[53] with a short-range cutoff of 11 Å.

For each ligand, four states were simulated (ligand in water $\lambda = 0$ and $\lambda = 1$, ligand in protein: $\lambda = 0$ and $\lambda = 1$), and each state was simulated in four separate replicas. The systems were minimized using the same procedure as the two cosolvent MD minimization phases. In the MD simulations, LINCS constraints[55] were applied to all bonds to ensure constraints to dummy hydrogens. Prior to the production MD runs, separate

unrestrained NVT (0.5 ns) and NPT (1 ns) equilibrations were run. In all MD simulations, the V-rescale thermostat[56] controlled system temperature with a reference temperature of 298 K and a time constant of 0.1 ps. In all MD simulations except the NVT equilibration, pressure control was performed using the C-rescale barostat[57] with a reference pressure of 1 bar and a time constant of 2.0 ps. Following equilibration, conventional MD simulations for each replica were run. These conventional MD simulations had a length of 50 ns, and 100 frames were extracted from each simulation. As four replicates were performed for each conventional MD run, a total of 400 snapshots were gathered in each state (ligand in water, modified ligand in water, ligand in protein and modified ligand in protein). The short transitions were started from these snapshots and had a length of 50 ps in which statistics about $\frac{dH}{d\lambda}$ (with H the Hamiltonian) were gathered.

In the postprocessing stage, work values on the system were computed from the $\frac{dH}{d\lambda}$ curves using the PMX analysis tools.[68] Forward and the reverse transition data were combined in the calculation of a ΔG via the Bennett Acceptance Ratio (BAR).[73,74] The $\Delta\Delta G$ value was then calculated by subtracting the free energy contribution of mutating the ligand in water from the free energy contribution of changing the ligand in the protein−ligand complex ($\Delta\Delta G = \Delta G_{PL \to PL'} - \Delta G_{L \to L'}$). Error estimates were calculated assuming that the $\Delta G_{PL \to PL'}$ and $\Delta G_{L \to L'}$ estimations are uncorrelated. In some cases, a thermodynamic pathway was constructed over more than one mutation to connect two compounds. In such cases the $\Delta\Delta G$ values were summed, and the standard deviation was approximated by summing the variances of the $\Delta\Delta G$ estimates.

■ RESULTS AND DISCUSSION

Analyzing the Fragment Densities. *DPP4.* We started the analyses by inspecting the fragment densities. The high-scoring
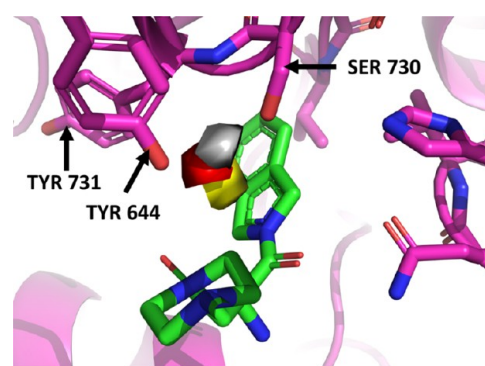
Figure 7. Overview of the orientation of the hydrogen bond donor/acceptor interactions of isopropanol near the catalytic center in DPP9 (magenta cartoon and magenta sticks). The cocrystallized inhibitor 1G244 (green sticks) is shown as a reference. The isopropanol central carbon density is shown in yellow (cutoff: 0.03 Å^{-3}), the isopropanol oxygen atom density is shown in red (cutoff: 0.02 Å^{-3}) and the hydroxyl hydrogen atom is shown in silver (cutoff: 0.02 Å^{-3}).

fragment densities in DPP4 are shown in Figure 4. High scoring fragment densities are observed in the active site (panel B), validating the hypothesis that strong fragment densities should be observed in relevant binding areas. Isopropanol densities are observed close to the catalytic center. It is found that, in the case of the DPP4 simulations, isopropyl ammonium molecules often
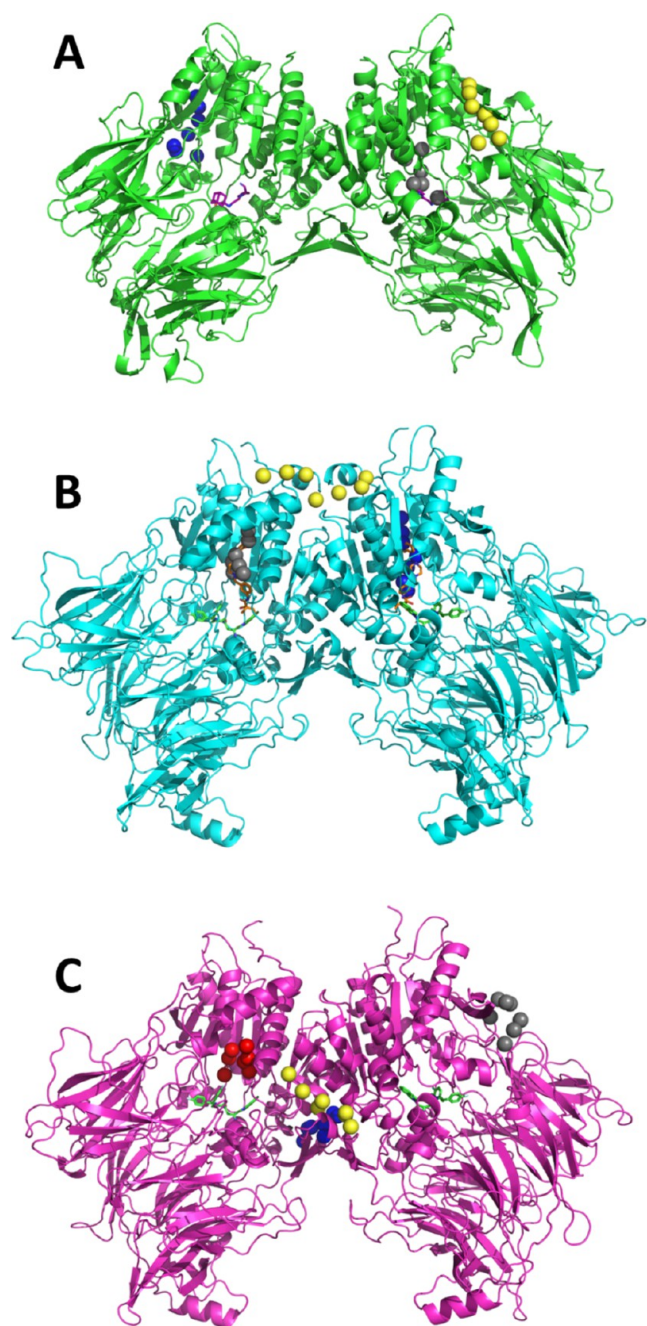
Figure 8. Highest ranking pockets in DPP4 (panel A, green), DPP8 (panel B, cyan) and DPP9 (panel C, magenta). In all panels, the highest ranking pocket is shown in gray, the second highest in dark blue and the third highest in yellow. Panel (A) shows that the top-ranked DPP4 pocket (gray spheres) is located in the active site. The second (blue) and third (yellow) highest ranked pockets are located near a known glycosylation site (Asn 520). Panel (B) shows that the top-two pockets (gray and blue) in DPP8 overlap with the binding location of compound **5**. Panel (C) highlights that the three highest-scoring pockets in DPP9 are located on the surface of the protein or near the homodimer interface. In Panel (C) the fourth highest-scoring pocket (red) is also included, as this is located in the area where the NLRP1 tail enters in the DPP9 structure (see Figure 1).

reside near the location where the positively charged secondary amine of vildagliptin is also located. Furthermore, a high-scoring cluster of fragment densities in a known glycosylation site in DPP4 (Asn 520) is also observed (Figure 4C). We however

speculate that the therapeutic potential of this pocket is debatable, as no biological function of the DPP4 glycosylation sites is known.[75] Furthermore, the glycosylation site is located in a surface pocket, which may prove to be difficult for drugs to bind due to its large solvent exposed area.

In Figure 4D, a high scoring isopropyl ammonium density in the side entry channel toward the active site is shown. This density location is of interest as all FDA and EMA approved inhibitors of DPP4 (inhibitors vildagliptin, saxagliptin, alogliptin, linagliptin and sitagliptin) have a positively charged moiety. A possible hypothesis could be that this affinity for positively charged moieties in the entry channel could lower the energy barrier of the binding process. Additionally, as DPP4 cleaves off a dipeptide from the N-terminus of the substrate, a similar mechanism could be conceivable for enzyme−substrate binding processes.

DPP8. In Figure 5, a similar analysis is given for DPP8. Similarly to DPP4, a density of isopropanol close to the catalytic serine is observed (panel B), which is in line with the strong hydrogen bond acceptor character of this catalytic residue. A strong density of the positively charged probe, namely isopropyl ammonium, can be found close to the location where the primary amine of the cocrystallized DPP8/9-inhibitor 1G244 is found (panel B). A second cluster of high-scoring probe densities is observed in the region where the 4-oxo-β-lactam inhibitor 12 (compound **5**[24]) likes to reside (panel C). A strong lipophilic density is observed in the region that corresponds to the binding pocket of the compound's naphthalene ring. The strong isobutane density next to the naphthalene ring hints toward further optimization with lipophilic substituents on the naphthalene ring. Next, a positively charged probe density in the side entry channel between the α/β-hydrolase domain and the β-propellor domain is found, similarly to DPP4 (panel D).

DPP9. The fragment density map for DPP9 is shown in Figure 6. In accordance with the results in DPP4 and DPP8, a strong isopropanol density is located close to the catalytic site (panel B). When investigating this isopropanol density further, it becomes clear that the isopropanol is acting as a hydrogen bond donor to the catalytic serine (see Figure 7 below). In DPP4 and DPP8 similar interactions are found, where in DPP4 a hydrogen donation toward the hydroxyl group of tyrosine amino acid can also be seen. In DPP9, the isopropanol density near the catalytic serine also appears to be acting as a hydrogen bond acceptor with the hydroxyl of Y644 and possibly with the nitrogen backbone atom of Y731. Another interesting density cluster is located in the side entry channel in DPP9, located on the interface between the α/β-hydrolase domain and the β-propellor domain. High-scoring isopropanol, isobutane and benzene probe densities overlap with the location, where the NLRP1 tail is located when binding (panel C). Finally, and again in full accordance with the results for the other two targets, a high scoring positively charged probe density close to the primary amine of the cocrystallized 1G244 is found (panel D).

The three main conclusions from the density analysis are as follows:

- Strong scoring densities of isopropanol and isopropyl ammonium are observed in the main binding pockets of DPP4, DPP8 and DPP9.

- A positive charge density in the side entry channels of DPP4, DPP8 and DPP9 is observed.
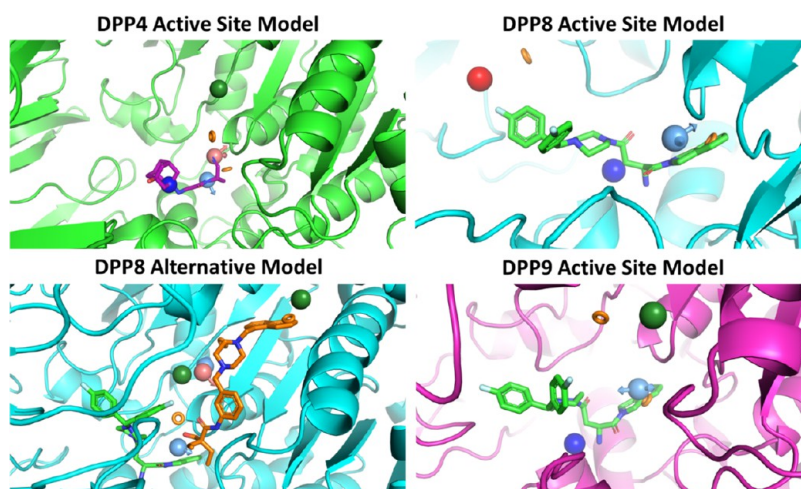
Figure 9. Overview of the four pharmacophore models used in the screening. Note that the color coding is the default Maestro color coding: salmon spheres as hydrogen bond acceptors, light blue spheres as donors, dark blue spheres as positively charged groups, red spheres as negatively charged groups, green spheres are hydrophobic features and orange rings are aromatic features. Orientations of the features was based on the orientation spots (see Methodology section).

- High scoring densities in alternative pockets are observed, such as in the 4-oxo-β-lactam inhibitor 12 pocket in DPP8 or in the NLRP1 tail binding area in DPP9.

The probe density maps are freely available at: https://github.com/UAMC-Olivier/DPP489_cosolvent_MD/.

Pocket Analysis. After visualizing the simplest scoring metric (probe densities), an analysis was performed to check whether normalizing the densities for probe concentration via the GFE value leads to different conclusions. After normalization, it was found that the interaction spots of aromatic rings in the active sites of DPP4, DPP8 and DPP9 have similar GFE values to those from the hydrogen bond donors/acceptors solvents (Figure S2, Supporting Information). This can be explained by the lower concentration of these aromatic probes, meaning that normalization to concentration will improve their relative scoring.

Analyses of the results of the algorithmic approach to combine interaction spots into binding pockets (see Methodology section) shows that the pockets that scored the best according to the algorithm correspond with known binding areas in DPP4 and DPP8 (Figure 8). DPP9 forms an exception to this, as in this case the highest-scoring pockets are all located on the protein surface or at the homodimerization interface. High scoring pockets are also observed near the glycosylation site (Asn 520) of DPP4 with strong fragment densities (Figure S3 Panels A,B, Supporting Information). In DPP8, the highest-scoring combinations of densities are located near the area where the 4-oxo-β-lactam inhibitor 12 is located in its crystal structure. The models indicate a strongly lipophilic pocket in the binding area of the ligand's naphthalene moiety (Figure S3 Panel C). The third highest-scoring pocket in DPP8 is also located on the protein surface, indicating a likely spurious pocket.

As the three highest-scoring pockets in DPP9 appear on the surface or at the homodimerization interface, the fourth highest-scoring combination of interaction spots was also included for further evaluation. The fourth pocket is located in the area of the S' pockets where the NLRP1 tail enters the side tunnel toward the active site. The fourth highest-scoring pocket in DPP9 is visualized in more detail in Figure S3D.

Pharmacophore Screening. Based on the cosolvent MD results, four pharmacophore models were generated. The first three models were built upon the respective active site regions, and the fourth model (the "alternative model") was built from the densities that make up the top-scoring pocket in DPP8. An overview of these four pharmacophore hypotheses is shown in Figure 9. In the active site pharmacophore models of DPP4 and DPP9, densities from the S' pockets were included as these scored favorably, coupled with a positively charged group near the isopropyl ammonium densities and either an aromatic ring or hydrogen bond donor near the catalytic site. In DPP8, we opted to add pharmacophores in the S2 extended (S2e) pocket, as in this region notable differences were observed between DPP8 and DPP9 in the corresponding fragment densities for a negatively charged probe. Consequently, the negatively charged probe interaction spot in the DPP8 S2e site was included in the DPP8 Active Site pharmacophore model. The pharmacophores in the S2e pocket were coupled with the positively charged density near the amine groups of the cocrystallized ligands and a donor or aromatic pharmacophore at the catalytic serine. In the DPP8 "alternative model", the strongly scoring density in the lipophilic pocket binding the naphthalene ring of 4-oxo-β-lactam inhibitor 12 was coupled to a hydrogen bond donor to the catalytic serine in the active site. Some strongly scoring connecting densities were also included.

Based on these four models, four corresponding pharmacophore screens were performed, resulting in a total of 24 predicted hits that were purchased. These compounds, together with the originating virtual screen sources, are listed in Table S1 (Supporting Information). Out of these 24 compounds, ten compounds originated from the "DPP4 Active Site" screen, eight from the "DPP8 Active Site", and four from the "DPP9 Active Site" screen. The last two compounds were identified by the "DPP8 Alternative Model" virtual screen.

The results of the *in vitro* evaluation of the selected compounds are shown in Table 4. Unfortunately, all tested compounds were weak binders at best. Compounds **8**, **13** and **23** show very modest binding and were selected for further analyses. Compounds **13** and **23** have some affinity for a target for which they were not selected, but this is not surprising given the high degree of similarity between the targets in the active sites. The lack of stronger binding hits can have several causes. First, a set of 24 compounds is a small test set, and therefore it is not

**Table 4. In Vitro Screening Results for the Selected Compounds**[a]

| | | v_i/v_0 (%) | | | | | |
| | | DPP4 | | DPP8 | | DPP9 | |
compound	concentration (μM)	run1	run2	run1	run2	run1	run2
6	50	**112**	**94**	92	98	106	107
	10	**94**	**97**	99	100	105	104
7	50	**93**	**94**	95	99	99	100
	10	**96**	**96**	99	99	103	97
8	50	**76**	**72**	89	93	92	95
	10	**92**	**91**	96	98	100	105
9	50	**93**	**94**	97	101	105	104
	10	**100**	**96**	96	99	105	104
10	50	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
	10	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
11	50	**93**	**93**	92	95	102	102
	10	**101**	**109**	97	98	106	103
12	50	**92**	**102**	97	100	102	104
	10	**87**	**99**	101	100	98	99
13	50	**95**	**99**	102	98	81	68
	10	**99**	**106**	101	98	90	82
14	50	**99**	**101**	102	95	91	90
	10	**107**	**100**	99	101	95	93
15	50	**101**	**100**	97	97	95	98
	10	**101**	**102**	105	102	73	97
16	50	93	93	**96**	**94**	99	97
	10	99	93	**101**	**99**	106	99
17	50	97	96	**97**	**94**	105	100
	10	96	95	**98**	**98**	103	98
18	50	100	97	**88**	**92**	97	89
	10	93	95	**96**	**100**	104	105
19	50	97	104	**100**	**100**	105	104
	10	96	102	**98**	**100**	106	104
20	50	106	102	**98**	**99**	92	94
	10	105	101	**101**	**101**	96	99
21	50	98	102	**98**	**101**	95	96
	10	96	98	**97**	**102**	97	98
22	50	95	100	**106**	**100**	97	101
	10	98	103	**101**	**97**	100	98
23	50	98	101	**100**	**100**	87	78
	10	101	108	**108**	**102**	98	91
24	50	103	110	**95**	**91**	97	99
	10	94	99	**98**	**99**	98	95
25	50	96	98	**95**	**98**	95	99
	10	101	98	**103**	**101**	110	93
26	50	94	100	81	87	**92**	**99**
	10	94	95	99	94	**94**	**100**
27	50	106	96	92	78	**92**	**95**
	10	102	99	100	87	**101**	**103**
28	50	100	97	103	88	**102**	**102**
	10	95	94	103	94	**102**	**104**
29	50	100	95	100	84	**92**	**92**
	10	101	92	102	97	**99**	**99**

[a]The compounds' effect on the enzymatic activity (v_i/v_0) is shown for two concentrations (10 and 50 μM). Compound 10 did not dissolve in DMSO and was not analyzed (N.A.). Compound 12 showed some degree of precipitation. Inhibition values in bold indicate the target for which the virtual screen was designed (see Table S1).

unlikely that based on the underlying hit rate for these enzymes no stronger binders were discovered. Expanding the set of selected compounds is a first option for finding stronger binders. Second, we opted for the chosen methodology for its simplicity and due to resource constraints. Including a follow-up consensus scoring docking approach or even short MD studies of the complexes could further improve the compound selection procedure.

For the three selected compounds, Phase and Glide docking poses were compared with each other to check consensus between both. This was indeed the case for compounds **8** and **13**, but for **23** such consensus was not found (Figure 10). Based
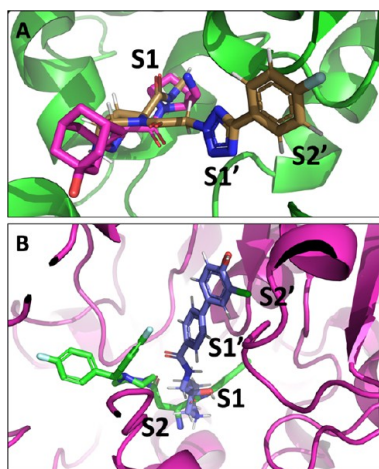
Figure 10. Panel (A) Overview of the docking pose of compound **8** (gold sticks) in DPP4 (green cartoon). Position of vildagliptin (magenta sticks) is shown as a reference. The pocket identifiers were determined via the binding mode of the Ile-Pro-Ile substrate in DPP4 (PDB: 1NU8[76]) similar to the methodology of Nabeno et al.[77] Panel (B) Overview of the predicted pose of compound **13** (purple sticks) in DPP9 (magenta cartoon). The 1G244 pose is shown in green sticks as reference. The pocket identifiers were also determined via an overlap with the binding mode of the Ile-Pro-Ile substrate in DPP4. Note that the designation of S2′ (and of S3′ in the Alchemical Free Energy Calculations section) is an approximation as Ile-Pro-Ile does not bind in the S′ pockets beyond S1′.

on this and on the affinity data, we decided to only move forward with **8** and **13** for subsequent *in silico* optimization.

Alchemical Free Energy Calculations. As the selected compounds **8** and **13** were weak binders at best, we decided to perform further *in silico* optimization of these compounds. Such *in silico* optimization has been shown to successfully turn weak binders into potent inhibitors.[24]

Compound 8 Modifications. We simulated modifications in the DPP4 S1 (Table S2), S1′ (Tables S3 and S4) and S3′ (Table S5) pockets for compound **8**. In the S1 pocket small nonaromatic rings (**30d** and **30e**; Table S2) are predicted to cause a small increase in the binding affinity, but not sufficient enough to obtain the large improvement that is required given the weak experimental binding potency of **8**. When modifying the moiety in the S1′ pocket by modifying the compound's backbone into a carbamide (Table S3) or carbamate (Table S4), no increase in binding affinities is predicted. By expanding the scaffold modifications to the S3′ pocket (Table S5), small improvements in the binding affinity are predicted for the amide moiety (**33f**), but these improvements are again not sufficient to gain the required orders of magnitude in binding affinity.

Compound 13 Modifications. With regard to the *in silico* optimization of compound **13**, we performed RBFE calculations to scan modifications in the S1′, S2′ and S3′ pockets. The first step in the molecular sculpting procedure was to remove the substituents on the terminating aromatic ring, leading to the biphenyl-containing compound **34** that has a slightly decreased binding affinity (Table S6). After removal of all aromatic substituents, a variety of aromatic substituents were tested and a nitrogen scan in the S1′ pocket was performed. As can be seen in Table S7, structures **34b** and **34d** are promising starting points with reasonable gains in binding affinity. A similar analysis was performed for the S2′ pocket, where mutations **35a** and **35h** show the strongest gains in binding affinity (−0.9 and −1.6 kcal/

mol, respectively; Table S8). When adding larger chemical moieties to the S3′ pocket of DPP9, modifications **36a**, **36e** and **36i** are most interesting (Table S9). The effect of removing the hydroxyl group from compound **13** was computed. The calculated RBFE results for the corresponding structure **37** indicate that the removal of this functional group does not improve the binding free energy (Table S10). Finally, Table S11 shows the results of the RBFE simulations when combining some of the aforementioned modifications. Compound **38a** shows strong gains in binding free energy, hypothetically turning the compound into a low micromolar or high nanomolar inhibitor. Compound **38d** also has a promising gain in $\Delta\Delta G$, and further modifications in the substituents of the terminal aromatic ring can be explored.

We conclude from the RBFE data that the tested modifications to compound **8** did not lead to the desired gains in binding free energy. For compound **13** we designed structures **38a** and **38d**. Both were predicted to lead to significant changes in binding free energy. Additional combinations (for example, combining **34d** with **36a** or **36i**) are possibly interesting as well and are part of future research.

■ CONCLUSIONS

The analyses of the cosolvent MD results showed that strong densities are observed in the active site pockets of DPP4, DPP8 and DPP9. A hydrogen bond donor/acceptor density close to the catalytic serine and a positively charged probe density near important positively charged groups of known ligands were consistently observed. A positively charged probe density was also identified in the entry channel. Furthermore, several high-scoring clusters of densities were observed. All density maps are made freely available.

When using the fragment locations and their associated GFE values in an algorithm to identify pockets, the algorithm scored the orthosteric pocket as the highest-scoring in DPP4. In DPP8, the pocket with the lowest GFE is located in the 4-oxo-β-lactam inhibitor 12 binding area. In DPP9, the strongest scoring pocket is located at the interface between the monomers, and the fourth highest-scoring cluster of densities is found in the side entry channel where the tail of NLRP1 enters the DPP9 structure.

As the aforementioned results looked promising, we initiated a pharmacophore screening from which 24 compounds were selected. *In vitro* analyses of these structures showed that the top compounds were unfortunately only weak binders at best. Subsequently, these weak binders were modified via RBFE calculations, resulting in significant gains in binding free energy with some mutations.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.4c01167.

> Overview of the pocket definitions in DPP8 (Panel A, cyan cartoon) and DPP9 (Panel B, magenta cartoon); visualization of all the interaction spots near the catalytic serine in DPP9; overview of the selected compounds from the pharmacophore screening with their respective model; overview of the modifications on compound 8 in the S1 pocket of DPP4 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Hans De Winter − *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, University of Antwerp, B-2610 Wilrijk, Belgium;* ● orcid.org/0000-0002-4450-7677; Email: hans.dewinter@uantwerpen.be

Authors

Olivier Beyens − *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, University of Antwerp, B-2610 Wilrijk, Belgium*

Sam Corthaut − *Laboratory of Medical Biochemistry, Department of Pharmaceutical Sciences, University of Antwerp, B-2610 Wilrijk, Belgium*

Sarah Peeters − *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, University of Antwerp, B-2610 Wilrijk, Belgium*

Pieter Van Der Veken − *Laboratory of Medicinal Chemistry, Department of Pharmaceutical Sciences, University of Antwerp, B-2610 Wilrijk, Belgium;* ● orcid.org/0000-0003-1208-3571

Ingrid De Meester − *Laboratory of Medical Biochemistry, Department of Pharmaceutical Sciences, University of Antwerp, B-2610 Wilrijk, Belgium;* ● orcid.org/0000-0002-3421-0124

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.4c01167

Author Contributions

O.B. performed the computational experiments. S.C. performed the biochemical assays and S.P. performed the LC-MS experiments. H.D.W. supervised the research. All authors contributed to writing the manuscript.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Hollingsworth, L. R.; Sharif, H.; Griswold, A. R.; Fontana, P.; Mintseris, J.; Dagbay, K. B.; Paulo, J. A.; Gygi, S. P.; Bachovchin, D. A.; Wu, H. DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation. *Nature* **2021**, *592* (7856), 778−783.

(2) Huang, M.; Zhang, X.; Toh, G. A.; Gong, Q.; Wang, J.; Han, Z.; Wu, B.; Zhong, F.; Chai, J. (Structural and biochemical mechanisms of NLRP1 inhibition by DPP9. *Nature* **2021**, *592* (7856), 773−777.

(3) Sharif, H.; Hollingsworth, L. R.; Griswold, A. R.; Hsiao, J. C.; Wang, Q.; Bachovchin, D. A.; Wu, H. Dipeptidyl peptidase 9 sets a threshold for CARD8 inflammasome formation by sequestering its active C-terminal fragment. *Immunity* **2021**, *54* (7), 1392−1404.

(4) Okondo, M. C.; Johnson, D. C.; Sridharan, R.; Go, E. B.; Chui, A. J.; Wang, M. S.; Poplawski, S. E.; Wu, W.; Liu, Y.; Lai, J. H.; Sanford, D. G.; Arciprete, M. O.; Golub, T. R.; Bachovchin, W. W.; Bachovchin, D. A. DPP8 and DPP9 inhibition induces pro-caspase-1-dependent monocyte and macrophage pyroptosis. *Nat. Chem. Biol.* **2017**, *13* (1), 46−53.

(5) Johnson, D. C.; Taabazuing, C. Y.; Okondo, M. C.; Chui, A. J.; Rao, S. D.; Brown, F. C.; Reed, C.; Peguero, E.; de Stanchina, E.; Kentsis, A.; Bachovchin, D. A. DPP8/DPP9 inhibitor-induced pyroptosis for treatment of acute myeloid leukemia. *Nat. Med.* **2018**, *24* (8), 1151−1156.

(6) Moore, K. P.; Schwaid, A. G.; Tudor, M.; et al. A phenotypic screen identifies potent DPP9 inhibitors capable of killing HIV-1 infected cells. *ACS Chem. Biol.* **2022**, *17* (9), 2595−2604.

(7) Clark, K. M.; Kim, J. G.; Wang, Q.; et al. Chemical inhibition of DPP9 sensitizes the CARD8 inflammasome in HIV-1-infected cells. *Nat. Chem. Biol.* **2023**, *19* (4), 431−439.

(8) Sato, T.; Tatekoshi, A.; Takada, K.; et al. DPP8 is a novel therapeutic target for multiple myeloma. *Sci. Rep.* **2019**, *9* (1), No. 18094.

(9) Kikuchi, S.; Wada, A.; Kamihara, Y.; et al. DPP8 selective inhibitor tominostat as a novel and broad-spectrum anticancer agent against hematological malignancies. *Cells* **2023**, *12* (7), No. 1100.

(10) Cui, C.; Tian, X.; Wei, L.; et al. New insights into the role of dipeptidyl peptidase 8 and dipeptidyl peptidase 9 and their inhibitors. *Front. Pharmacol.* **2022**, *13*, No. 1002871.

(11) Zolg, S.; Donzelli, L.; Geiss-Friedlander, R.N-terminal processing by dipeptidyl peptidase 9: Cut and Go! *Biochimie* **2024** .

(12) Berger, J. P.; SinhaRoy, R.; Pocai, A.; et al. A comparative study of the binding properties, dipeptidyl peptidase- 4 (DPP-4) inhibitory activity and glucose- lowering efficacy of the DPP-4 inhibitors alogliptin, linagliptin, saxagliptin, sitagliptin and vildagliptin in mice. *Endocrinol., Diabetes Metab.* **2018**, *115* (7), E1437−E1445.

(13) Ross, B.; Krapp, S.; Augustin, M.; et al. Structures and mechanism of dipeptidyl peptidases 8 and 9, important players in cellular homeostasis and cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E1437−E1445, DOI: 10.1073/pnas.1717565115.

(14) Mentlein, R.; Gallwitz, B.; Schmidt, W. E. Dipeptidyl-peptidase IV hydrolyses gastric inhibitory polypeptide, glucagon-like peptide-1 (7−36) amide, peptide histidine methionine and is responsible for their degradation in human serum. *Eur. J. Biochem.* **1993**, *214* (3), 829−835.

(15) Nauck, M. A.; Kleine, N.; Ørskov, C.; et al. Normalization of fasting hyperglycaemia by exogenous glucagon-like peptide 1 (7−36 amide) in type 2 (non-insulin-dependent) diabetic patients. *Diabetologia* **1993**, *36*, 741−744.

(16) Gerich, J. E. Contributions of insulin-resistance and insulin-secretory defects to the pathogenesis of type 2 diabetes mellitus. *Mayo Clin. Proc.* **2003**, *4*, 447−456, DOI: 10.4065/78.4.447.

(17) Egan, A. G.; Blind, E.; Dunder, K.; et al. Pancreatic safety of incretin-based drugs—FDA and EMA assessment. *New Engl. J. Med.* **2014**, *370* (9), 794−797.

(18) Jiaang, W.-T.; Chen, Y. S.; Hsu, T.; et al. Novel isoindoline compounds for potent and selective inhibition of prolyl dipeptidase DPP8. *Bioorg. Med. Chem. Lett.* **2005**, *15* (3), 687−691.

(19) Wu, J.-J.; Tang, H. K.; Yeh, T. K.; et al. Biochemistry, pharmacokinetics, and toxicology of a potent and selective DPP8/9 inhibitor. *Biochem. Pharmacol.* **2009**, *78* (2), 203−210.

(20) Lankas, G. R.; Leiting, B.; Roy, R. S.; et al. Dipeptidyl peptidase IV inhibition for the treatment of type 2 diabetes: potential importance of selectivity over dipeptidyl peptidases 8 and 9. *Diabetes* **2005**, *54* (10), 2988−2994.

(21) Van Goethem, S.; Matheeussen, V.; Joossens, J.; et al. Structure−activity relationship studies on isoindoline inhibitors of dipeptidyl peptidases 8 and 9 (DPP8, DPP9): is DPP8-selectivity an attainable goal? *J. Med. Chem.* **2011**, *54* (16), 5737−5746.

(22) Benramdane, S.; De Loose, J.; Beyens, O.; et al. Vildagliptin-derived Dipeptidyl Peptidase 9 (DPP9) inhibitors: identification of a DPP8/9-specific lead. *ChemMedChem* **2022**, *17* (15), No. e202200097.

(23) Benramdane, S.; De Loose, J.; Filippi, N.; et al. Highly selective inhibitors of dipeptidyl peptidase 9 (DPP9) derived from the clinically used DPP4-inhibitor vildagliptin. *J. Med. Chem.* **2023**, *66* (18), 12717−12738.

(24) Carvalho, L. A.; Ross, B.; Fehr, L.; et al. Chemoproteomics-enabled Identification of 4-oxo-*β*-lactams as inhibitors of dipeptidyl peptidases 8 and 9. *Angew. Chem., Int. Ed.* **2022**, *61* (47), No. e202210498.

(25) Comitani, F.; Gervasio, F. L. Exploring cryptic pockets formation in targets of pharmaceutical interest with SWISH. *J. Chem. Theory Comput.* **2018**, *14* (6), 3321−3331.

(26) Kimura, S. R.; Hu, H. P.; Ruvinsky, A. M.; et al. Deciphering cryptic binding sites on proteins by mixed-solvent molecular dynamics. *J. Chem. Inf. Model.* **2017**, *57* (6), 1388−1401.

(27) Schmidt, D.; Boehm, M.; McClendon, C. L.; et al. Cosolvent-enhanced sampling and unbiased identification of cryptic pockets suitable for structure-based drug design. *J. Chem. Theory Comput.* **2019**, *15* (5), 3331−3343.

(28) Guvench, O.; Mackerell, A. D. Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput. Biol.* **2009**, *5* (7), No. e1000435.

(29) Zariquiey, F. S.; Jacoby, E.; Vos, A.; et al. Divide and conquer. Pocket-opening mixed-solvent simulations in the perspective of docking virtual screening applications for drug discovery. *J. Chem. Inf. Model.* **2022**, *62* (3), 533−543.

(30) Ghanakota, P.; Carlson, H. A. Driving structure-based drug discovery through cosolvent molecular dynamics. *J. Med. Chem.* **2016**, *59* (23), 10383−10399.

(31) Lee, J. Y.; Krieger, J. M.; Li, H.; et al. Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations. *Protein Sci.* **2020**, *29* (1), 76−86.

(32) Mejías, C.; Guirola, O. Pharmacophore model of immunocheckpoint protein PD-L1 by cosolvent molecular dynamics simulations. *J. Mol. Graphics Modell.* **2019**, *91*, 105−111.

(33) Yu, W.; Lakkaraju, S. K.; Raman, E. P.; et al. Pharmacophore modeling using site-identification by ligand competitive saturation (SILCS) with multiple probe molecules. *J. Chem. Inf. Model.* **2015**, *55* (2), 407−420.

(34) Zariquiey, F. S.; de Souza, J. V.; Bronowska, A. K. Cosolvent analysis toolkit (CAT): a robust hotspot identification platform for cosolvent simulations of proteins to expand the druggable proteome. *Sci. Rep.* **2019**, *9* (1), No. 19118.

(35) Bakan, A.; Nevins, N.; Lakdawala, A. S.; et al. Druggability assessment of allosteric proteins by dynamics simulations in the presence of probe molecules. *J. Chem. Theory Comput.* **2012**, *8* (7), 2435−2447.

(36) Graham, S. E.; Leja, N.; Carlson, H. A. MixMD probeview: robust binding site prediction from cosolvent simulations. *J. Chem. Inf. Model.* **2018**, *58* (7), 1426−1433.

(37) Beyens, O.; De Winter, H. Preventing lipophilic aggregation in cosolvent molecular dynamics simulations with hydrophobic probes using Plumed Automatic Restraining Tool (PART). *J. Cheminf.* **2024**, *16* (1), No. 23.

(38) Fiser, A.; Do, R. K. G.; Šali, A.; et al. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753−1773.

(39) Fiser, A.; Andrej, S. ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **2003**, *19* (18), 2500−2501.

(40) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **2017**, *27* (1), 293−315.

(41) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; et al. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.

(42) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; et al. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comp.* **2011**, *7*, 525−537.

(43) Gordon, J. C.; Myers, J. B.; Folta, T.; et al. H++: a server for estimating pKa's and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, *1*, W368−W372.

(44) Myers, J.; Grothaus, G.; Narayanan, S.; et al. A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins: Struct., Funct., Bioinf.* **2006**, *63* (4), 928−938.

(45) Anandakrishnan, R.; Boris, A.; Alexey, V. O. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537−W541.

(46) Van Der Spoel, D.; Lindahl, E.; Hess, B.; et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26 **2005**, *26*, 1701−1718.

(47) Abraham, M. J.; Murtola, T.; Schulz, R.; et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(48) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31* (4), 671−690.

(49) Vanommeslaeghe, K.; MacKerell, A. D., Jr Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52* (12), 3144−3154.

(50) Yu, W.; He, X.; Vanommeslaeghe, K.; et al. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **2012**, *33* (31), 2451−2468.

(51) Huang, J.; Mackerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34* (25), 2135−2145.

(52) Huang, J.; Rauscher, S.; Nawrocki, G.; et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14* (1), 71−73.

(53) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: an N· log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.

(54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; et al. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(55) Hess, B.; Bekker, H.; Berendsen, H. J. C.; et al. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463−1472.

(56) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126* (1), No. 014101.

(57) Bernetti, M.; Bussi, G. Pressure control using stochastic cell rescaling. *J. Chem. Phys.* **2020**, *153* (11), No. 114107.

(58) The PLUMED Consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16* (8), 670−673.

(59) Tribello, G. A.; Bonomi, M.; Branduardi, D.; et al. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604−613.

(60) Bonomi, M.; Branduardi, D.; Bussi, G.; et al. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180* (10), 1961−1972.

(61) Gowers, R.; Linke, M.; Barnoud, J. et al. In *MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations*, Proceedings of the 15th Python in Science Conference; SciPy, 2016.

(62) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; et al. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.

(63) *Vida 5.0.4*; OpenEye Scientific Software, Inc.: Santa Fe, NM.

(64) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; et al. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput.-Aided Mol. Des.* **2006**, *20* (10−11), 647−671.

(65) De Meester, I.; Vanhoof, G.; Lambeir, A. M.; et al. Use of immobilized adenosine deaminase (EC 3.5. 4.4) for the rapid purification of native human CD26/dipeptidyl peptidase IV (EC 3.4. 14.5). *J. Immunol. Methods* **1996**, *189* (1), 99−105.

(66) De Decker, A.; Vliegen, G.; Van Rompaey, D.; et al. Novel small molecule-derived, highly selective substrates for fibroblast activation protein (FAP). *ACS Med. Chem. Lett.* **2019**, *10* (8), 1173−1179.

(67) Gapsys, V.; Pérez-Benito, L.; Aldeghi, M.; et al. Large scale relative protein ligand binding affinities using non-equilibrium alchemy. *Chem. Sci.* **2020**, *11* (4), 1140−1152.

(68) Gapsys, V.; Michielssens, S.; Seeliger, D.; et al. pmx: Automated protein structure and topology generation for alchemical perturbations. *J. Comput. Chem.* **2015**, *36* (5), 348−354.

(69) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78* (8), 1950−1958.

(70) Joung, I. S.; Cheatham, T. E., III Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Chem. Phys. B* **2008**, *112* (30), 9020−9041.

(71) da Silva, A. W. S.; Vranken, W. F. ACPYPE-Antechamber python parser interface. *BMC Res. Notes* **2012**, *5*, No. 367.

(72) He, X.; Man, V. H.; Yang, W.; et al. A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* **2020**, *153* (11), No. 114502.

(73) Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **1999**, *60* (3), No. 2721.

(74) Shirts, M. R.; Bair, E.; Hooker, G.; et al. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett.* **2003**, *91* (14), No. 140601.

(75) Aertgeerts, K.; Ye, S.; Shi, L.; et al. N-linked glycosylation of dipeptidyl peptidase IV (CD26): Effects on enzyme activity, homodimer formation, and adenosine deaminase binding. *Protein Sci.* **2004**, *13* (1), 145−154.

(76) Thoma, R.; Löffler, B.; Stihle, M.; et al. Structural basis of proline-specific exopeptidase activity as observed in human dipeptidyl peptidase-IV. *Structure* **2003**, *11* (8), 947−959.

(77) Nabeno, M.; Akahoshi, F.; Kishida, H.; et al. A comparative study of the binding modes of recently launched dipeptidyl peptidase IV inhibitors in the active site. *Biochem. Biophys. Res. Commun.* **2013**, *434* (2), 191−196.