

# Identification of an Evolutionarily Conserved Allosteric Network in Steroid Receptors

Namita Dube, Sabab Hasan Khan, Riley Sasse, and C. Denise Okafor\*



Cite This: *J. Chem. Inf. Model.* 2023, 63, 571–582



Read Online

ACCESS |



Metrics & More

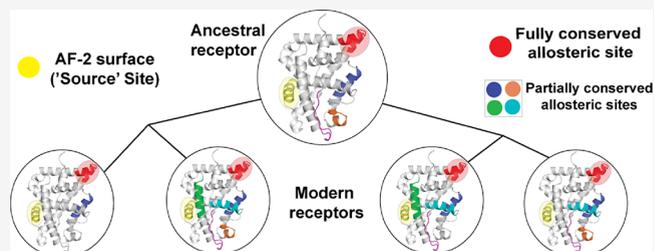


Article Recommendations



Supporting Information

**ABSTRACT:** Allosteric pathways in proteins describe networks comprising amino acid residues which may facilitate the propagation of signals between distant sites. Through inter-residue interactions, dynamic and conformational changes can be transmitted from the site of perturbation to an allosteric site. While sophisticated computational methods have been developed to characterize such allosteric pathways linking specific sites on proteins, few attempts have been made to apply these approaches toward identifying new allosteric sites. Here, we use molecular dynamics simulations and suboptimal path analysis to discover new allosteric networks in steroid receptors with a focus on evolutionarily conserved pathways. Using reconstructed ancestral receptor, we identify networks connecting several sites to the activation function surface 2 (AF-2), the site of coregulator recruitment. One of these networks is conserved across the entire family, connecting a predicted allosteric site located between helices 9 and 10 of the ligand-binding domain. We investigate the basis of this conserved network as well as the importance of this site, discovering that the site lies in a region of the ligand-binding domain characterized by conserved inter-residue contacts. This study suggests an evolutionary importance of the helix 9–helix 10 site in steroid receptors and identifies an approach that may be applied to discover previously unknown allosteric sites in proteins.



## INTRODUCTION

The dynamic nature of proteins is an essential characteristic for biological functions, permitting catalysis, signaling, and other critical protein actions.<sup>1,2</sup> As protein motions and conformational transitions are intricately linked to their function, understanding how these dynamics are altered over long evolutionary timescales could illuminate the biophysical principles that govern the function of extant proteins. Molecular dynamic (MD) simulations offer an accessible way to investigate macromolecular dynamics and learn how proteins evolve dynamic functions.<sup>3–5</sup> Here, we seek to understand the evolution of allosteric signaling pathways in steroid receptors.

Steroid receptors (SR) are ligand-activated transcription factors that belong to the nuclear receptor (NR) superfamily of proteins.<sup>6,7</sup> As is common to nuclear receptors, allostery plays a key role in regulating SR function.<sup>8</sup> In complex with cellular factors broadly known as coregulators, SRs bind specific enhancer/promoter DNA sequences to activate gene transcription.<sup>9</sup> Exponentially large numbers of coregulatory binding partners of SRs have been reported.<sup>10,11</sup> Differing local concentrations of potential coregulatory partners coupled with structurally dynamic SRs whose conformations are influenced by ligand binding result in myriad possibilities of dynamically composed transcription complexes.<sup>11,12</sup> Thus, binding of agonist or antagonist ligands to the ligand-binding domain (LBD) allosterically regulates SRs by initiating

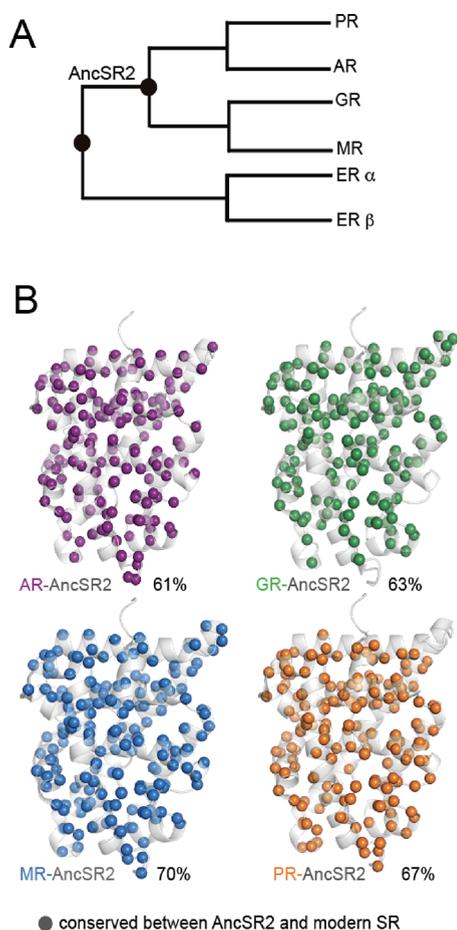
dynamic interdomain and intradomain effects.<sup>13–16</sup> For example, conformational effects propagate to multiple sites on the LBD, including a dimerization interface and the coregulator binding surface known as activation function surface-2 (AF-2). The SR family comprises six members in vertebrates: estrogen receptors  $\alpha$  and  $\beta$  (ER $\alpha$ /ER $\beta$ ), progesterone receptors (PRs), androgen receptors (ARs), glucocorticoid receptors (GRs), and mineralocorticoid receptors (MRs).<sup>4,17</sup> While all SRs derive from a common ancestral gene, extant SRs group into two phylogenetic classes: the estrogen receptors and oxosteroid (or 3-ketosteroid) receptors:<sup>18</sup> progesterone, androgen, mineralocorticoid, and glucocorticoid receptors (PR/AR/MR/GRs). One key distinction between both classes is their preference for steroid hormones with an aromatized A-ring versus 3-ketosteroids with a keto substituent at C3. The reconstruction and functional characterization of ancestral SR genes,<sup>19–21</sup> including ancestral 3-ketosteroid receptor AncSR2 (Figure 1A), was beneficial for

Received: August 29, 2022

Published: January 3, 2023



describing the evolutionary history and proliferation of the SR family.



**Figure 1.** Evolutionary history and conservation in oxosteroid receptors. (A) Phylogeny of steroid receptors identifies two families, namely, oxosteroid/3-ketosteroid receptors and estrogen receptors. AncSR2 is the reconstructed ancestor of the oxosteroid clade. (B) Conservation between AncSR2 and human oxosteroid receptors, which are mapped onto the crystal structure of AncSR2 LBD (PDB 4LTW). Spheres highlight residue positions conserved between AncSR2 and extant SRs.

While the extant SR LBDs are only moderately conserved (~50% identity) between one another, they each share higher identity with their ancestor AncSR2 (Figure 1B). AncSR2 is promiscuously activated by corticosteroids, progestogens, and androgens.<sup>18</sup> To a lesser extent, promiscuity is also observed in modern SRs, such as PR, which is potently activated by glucocorticoids, and MR, which binds glucocorticoids with the same affinity as the cognate hormone aldosterone. This cross-reactivity results from the conservation within the SR family.<sup>22–26</sup> We hypothesize that this conservation extends to allosteric signaling networks that may be common to all SRs and permit communication across the receptors. Here, we use allostery to describe the phenomenon where perturbation at a protein site (e.g., mutagenesis, ligand binding, and post-translational modification) induces a functional and/or structural effect at some distant site of the same protein.<sup>27</sup> This work focuses on identifying and describing networks in SRs that mediate allosteric signaling.

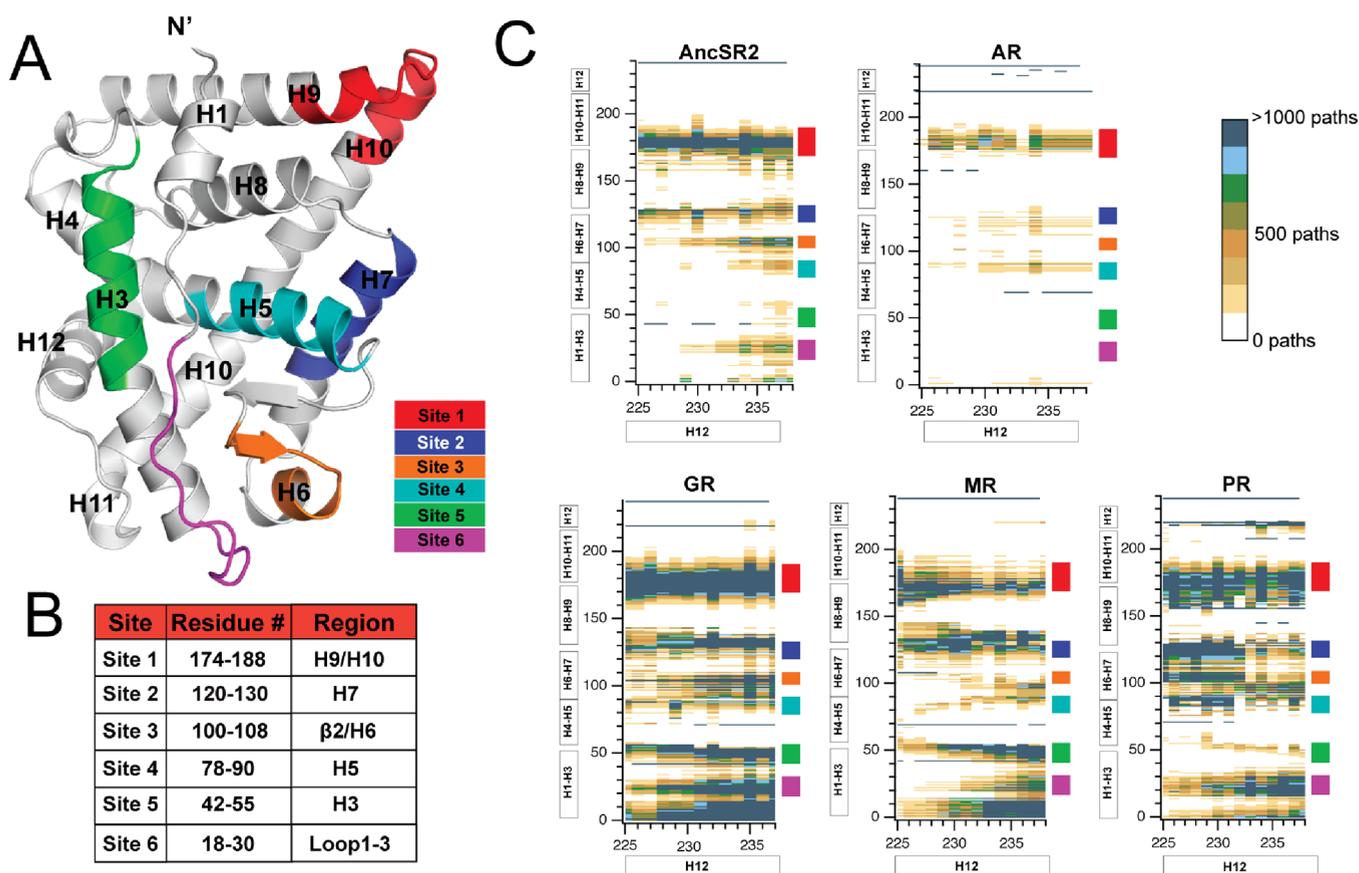
Communication between distant sites within proteins is proposed to occur via allosteric signaling pathways.<sup>28</sup> There has been a recent surge in computational methods geared at revealing pathways of amino acids that permit perturbations to be transmitted between two allosterically linked protein sites.<sup>29–31</sup> Here, we characterize allosteric signaling using suboptimal paths analysis.<sup>32,33</sup> This approach builds on a graph-based representation of proteins as amino acid networks where each node is a residue.<sup>34</sup> Nodes within spatial proximity are connected by edges whose weights correspond to measurable attributes of the amino acids. The shortest chain of amino acids connecting the two sites (or the optimal path) is presumed to contain the most critical residues for mediating allostery between both regions. However, a subset of slightly longer “suboptimal” paths conveys a substantial amount of allosteric information. Thus, distant sites connected by a large number of suboptimal paths relative to other regions is an indicator of a strong allosteric connection.

Here, we combine MD simulations with the suboptimal paths approach to discover allosterically linked sites in SRs in an unbiased manner. To identify paths that are evolutionarily conserved, we perform our investigation using oxosteroid receptors AR, PR, MR, and GR along with their reconstructed ancestor AncSR2. This approach provides a perspective on how allosteric signaling mechanisms have evolved in SRs as a result of residue substitutions, proposing rationales for how allosteric signaling networks changed over time to permit functional diversification in SRs.

Focusing our analysis on the LBDs of the five SRs, we uncover multiple sites that are allosterically linked with the activation function helix, Helix 12 (H12). Of these sites, we observe that only one site located at the H9–H10 loop (L9–10) is conserved among all SRs. This site is implicated in SR dimerization and holds functional relevance for nuclear receptors. To understand the structural and dynamic basis of this conserved signaling network, we identified the most prevalent amino acids in suboptimal paths as well as other allosteric sites connected to L9–10. We determined that L9–10 lies within the localized region of SR LBDs possessing the most conserved edges, allowing this signaling network to be preserved in all SRs. We identified helices 10 and 5 (H10 and H5) as crucial signaling conduits which also play a role in the conserved signaling network between H12 and L9–10. These studies suggest that L9–10 is an allosteric site that may influence key functions in SRs beyond known roles in dimerization and motivate future experimental work to define these functions.

## RESULTS AND DISCUSSION

**Six Conserved Allosteric SR Hotspots Communicate with H12.** The AF-2 surface, comprising H12 with parts of H3 and H4, is a key functional surface on nuclear hormone receptors that mediates coregulator recruitment.<sup>35</sup> To identify signaling pathways across PR, AR, MR, GR, and AncSR2 that could mediate communication to AF-2, we began by quantifying communication between all LBD residues and H12 using a suboptimal-path analysis. Five-hundred-nanosecond MD trajectories were obtained in triplicate for each unliganded SR and combined followed by the generation of dynamic networks (nodes and edges) via NetworkView<sup>36</sup> in VMD.<sup>37</sup> In this method, edges are weighted by calculated correlations between residue pairs. Distal sites are connected by chains of edges (paths) formed between proximal amino



**Figure 2.** Six allosteric sites in SRs. (A) Structural model of AncSR2 LBD (PDB 4LTV) with helices labeled and Sites 1–6 colored accordingly. Site 1 (res 174–188) is L9–10 with a few residues from each helix. Site 2 (res 120–130) is the C-terminal half of H7. Site 3 (res 100–108) is one  $\beta$  strand and part of H6. Site 4 (res 78–90) is H5. Site 5 (res 42–55) is the C-terminal half of H3. Site 6 (res 18–30) is L1–3, the loop linking helices 1 and 3. (B) Amino acid residues corresponding to locations of Sites 1–6. (C) Matrices showing the number of suboptimal paths between all residues and H12 (residues 225–235) for AncSR2, AR, GR, PR, and MR. Positions of Sites 1–6 are identified.

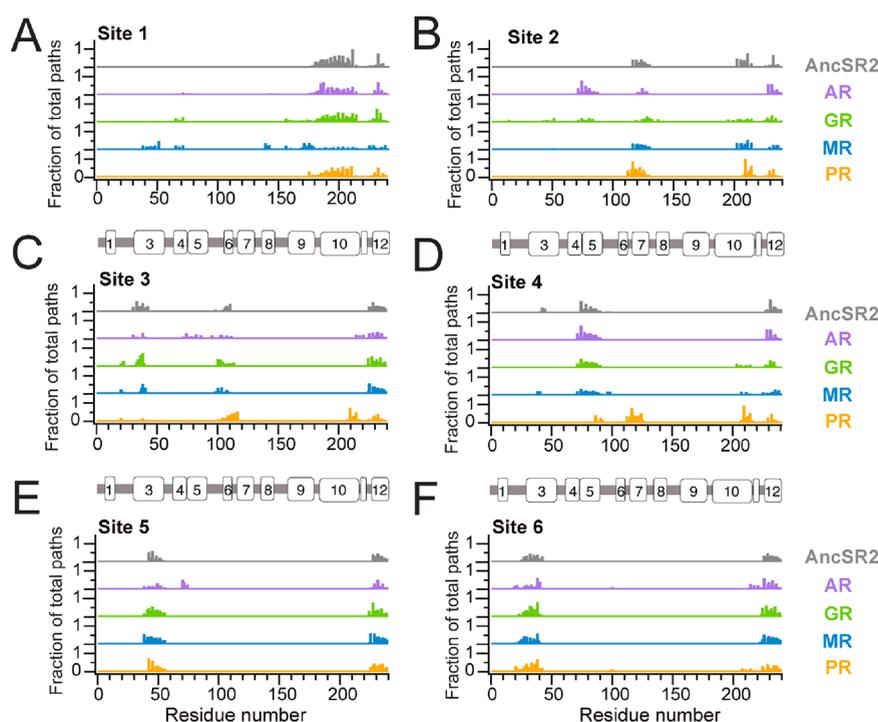
acids. Focusing on distant residue pairs, we describe the communication between any two amino acids as the number of suboptimal paths between them. By identifying residues with strong allosteric communication to AF-2 (i.e., connected by large numbers of suboptimal paths), we aimed to discover new allosteric sites within SRs. To optimize the cutoff for calculating suboptimal paths (see [Methods](#)), we tested cutoff values of 25, 50, and 75 ([Figure S1](#)), identifying 50 as an ideal value to distinguish between allosteric sites in the five SRs studied here.

We observe six distinct LBD sites in AncSR2 that are allosterically connected to H12 ([Figure 2A,B](#)). Designated as sites 1–6 and numbered from the C- to the N-terminal, these positions are conserved in the extant SRs to varying extents ([Figure 2C](#)). Site 1, located on H9–H10 and the intervening loop L9–10, is the only site that is equally prominent in all five SRs. Site 2, located on the C-terminal end of H7, is present in all SRs except for AR where it appears extremely attenuated. Site 3 ( $\beta$ -strand and the N-terminal of H6) is found in AncSR2, GR, and PR with much lower path numbers in MR and AR. Site 4 is found on H5 and contains more paths in extant SRs GR and PR than AncSR2. We note that Site 5, unlike the other sites that are distal to AF-2, is located on H3, which is part of the AF-2, close to H12. Interestingly, Site 5 is non-existent in AR and is weak (low number of suboptimal paths) in AncSR2 and MR. Finally, Site 6, located on the loop

linking helices 1 and 3 (L1–3), is conserved in all SRs except AR.

Signals may propagate from allosteric sites to H12 via several potential routes. To characterize the extent to which signaling pathways between Sites 1–6 and H12 are conserved across SRs, we identified and quantified the residues that constitute suboptimal paths ([Figure 3](#)). For each SR site, we quantify the utilization of each residue as a fraction of the total number of paths. For example, a residue appearing in 50% of given paths between any site and H12 is assigned fraction 0.5. In AncSR2, communication between Site 1 and H12 proceeds almost exclusively through H10 ([Figure 3A](#)). This pattern is conserved in AR/PR, while GR/MR both incorporate H9 and H4 residues as well as H3 residues in MR only. Conversely, Site 2 signaling from H7 is conserved between AncSR2, MR, and PR with paths traveling to H12 via the C-terminus of H10 ([Figure 3B](#)). GR and AR incorporate additional residues in paths, including H3 and H5.

Signaling paths for Site 3, located on the H6/ $\beta$ -strand, primarily proceed through the bottom of H3 to the AF-2 in all SRs, except PR, which uniquely display significant incorporation of H10 residues ([Figure 3C](#)). AR also specifically utilizes H4/H5 in paths—residues that are not observed in any other SR. Site 4 signaling (res 78–90) is also remarkably conserved in all SRs but PR ([Figure 3D](#)). While signaling proceeds directly through H4 to H12 in AncSR2, AR, MR, and GR, PRs include a strong incorporation of H7. Some contributions from



**Figure 3.** Residues in allosteric signaling pathways. (A–F) For each site (Sites 1–6), residues involved in suboptimal paths to H12 are shown. The Y axis indicates a normalized percentage utilization of each residue relative to the total combined number of paths between each site and H12.

H10 residues are also observed in MR, GR, and PR. Unsurprisingly, signaling from Site 5 (res 42–55, located on H3) is conserved across all SRs (Figure 3E). The proximity of H3 to H12 permits interhelical contacts to easily form between the helices. Other than minor H4/H5 contributions in AR, no other residues are observed in these paths. Finally, Site 6 (L1–3) signaling, most prominent in GR, MR, and PR, proceeds to H12 through H3 residues (Figure 3F).

**Ligand-Bound Complexes Confirm Site 1 as a Crucial Allosteric Site.** To reveal the effects of ligands on SR signaling networks, we performed MD simulations on all SRs in the presence of multiple steroidal ligands with a range of activation profiles: corticosterone (Cor), dexamethasone (Dex), dihydrotestosterone (DHT), progesterone (Prog), and cortisol (HCY) (Figure 4B). Following MD simulations, we repeated the suboptimal paths analysis, calculating paths between all pairs of residues for each SR–ligand complex. We observe that, in the presence of ligands, Site 1 signaling remains the only conserved feature across all complexes (Figure 4A).

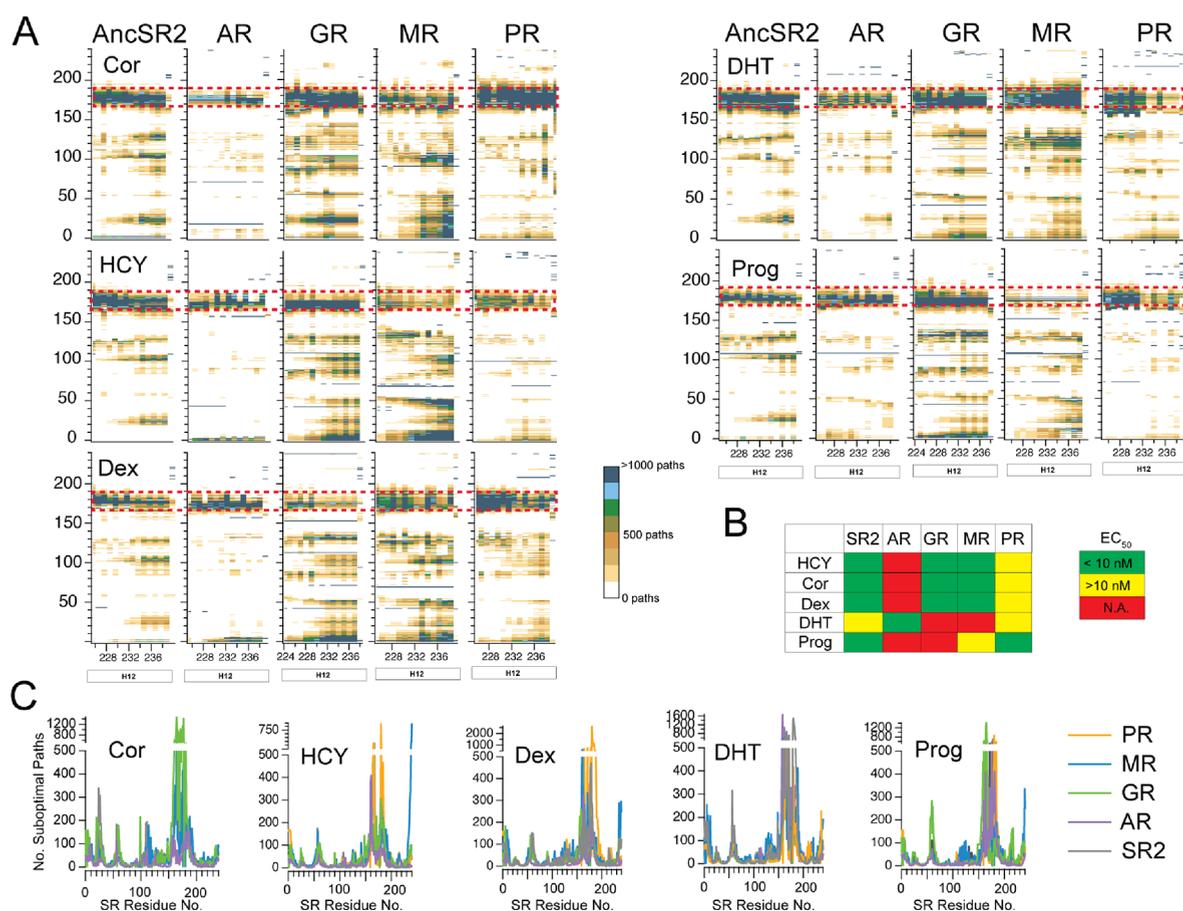
To identify the residues in each SR with the strongest allosteric connection to the bound ligands, we generated and quantified suboptimal paths between the ligand and all SR residues (Figure 4C). A clear trend is observed in all SR–ligand complexes: the highest number of paths appears between the ligand and residues 155–185, encompassing the entirety of H9, L9–10, and part of H10. This observation confirms H9/H10 as a crucial allosteric site conserved across all SRs.

We sought to further investigate the effects of ligands on signaling between AF-2 and Site 1. To characterize the preferred signaling paths and compare them to those used by unliganded SRs, we identified and quantified the residues that constitute suboptimal paths in ligand-bound complexes (Figure 5A). We observe that ligands tend to introduce a

larger fraction of H4 and H9 residues in paths. In contrast, in unliganded complexes, H4 residues were only observed for GR and MR (Figure 3A). With the addition of ligands, H4 residues play a prominent role in multiple AR and GR complexes with very minor contributions in AncSR2 and MR. H9 also plays a more prominent role in AR and GR complexes than previously seen in the absence of ligands. Overall, we observe that AncSR2 and PR signaling paths mostly remain unchanged by ligands.

To quantify how ligand addition drives changes in edge weights (see Methods for description of edge weights) that may alter signaling paths, we first identified the edges primarily utilized for Site 1 signaling in the unliganded complexes (Figure 3). We then determined the weights of the same edges in ligand-bound states and used Box plots to compare distributions across all complexes (Figure 5B–F). Finally, to gain a better sense of which weights increase versus decrease to drive changes in communication pathways, we identified the amino acid pair involved in each edge and created heat maps to compare their weights (Tables S1–S5). These edges have been grouped as intra- or interhelical and by their location in SRs. In AncSR2 (Figure 5B) and PR (Figure 5F), we observe that most edge weights lie within a similar range (20–160), suggesting that ligand addition does not drastically affect edge weights. Similarly, heat maps of AncSR2 and PR edge weights show minimal differences between ligand states (including unliganded) for regions across the receptor (Tables S1 and S5). These modest differences may explain why signaling paths are mostly unchanged by ligand addition in these two SRs (Figure 5A). Notably, AncSR2 and PR are the only SRs in this study, which are activated by all ligands to varying extents in vitro assays (Figure 4B).

In AR, ligand addition leads to large increases ( $\Delta \approx 120$ –150) in the range of edge weights (Figure 5C). We also observe a larger interquartile range in AR complexes compared



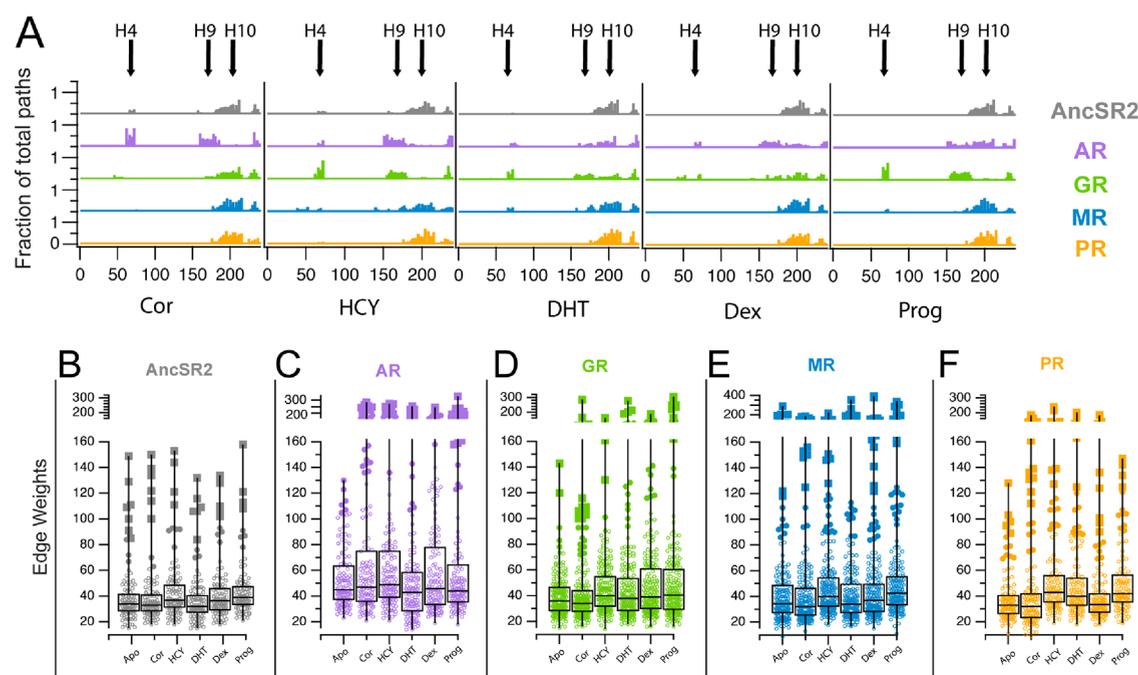
**Figure 4.** Site 1 is allosterically linked to ligand binding. (A) Matrices showing the number of suboptimal paths between all residues and H12 in SR–ligand complexes (five steroid hormones with each SR). For each complex, the signaling network between Site 1 and H12 is conserved as evidenced by large numbers of suboptimal paths (red dashed box). (B) Steroids used for the study display varying activation profiles in each SR. N.A. = no activation observed.<sup>38–42</sup> (C) Number of suboptimal paths calculated between the ligand and all residues for all SR–hormone complexes. The highest path numbers are consistently observed between hormones and H9/H10/L9–10, indicating an allosteric connection between the ligand-binding pocket and Site 1.

to other SRs. The heat map of AR edge weights (Table S2) shows that this increase is driven by edges between H4–H12 (i.e., edges connecting H4 and H12), H4/5–H8, and H10–H12. Interestingly, we note that, for most of the edges, AR–DHT weights are the same or lower than the unliganded, which is also reflected in the lower box range of DHT (Figure 5C). Despite maintaining similar edge weights, the DHT complex uses different residues than the unliganded complex (Figure 5A). This observation suggests that, like other ligands, DHT induces strong edges in H4 and H9 that favor signaling and drive predicted pathways (Figure 5A). Additionally, we note that DHT is the only ligand that activates AR *in vitro*, and similar to observations in AncSR2 and PR, it preserves similar or lower (compared to unliganded ARs) edge weights across various regions of AR (Table S2). These combined observations suggest that activating ligands may induce strong edges in certain regions of the receptor compared to inactive ligands. Future studies with a larger number of ligands would be necessary to further explore this hypothesis.

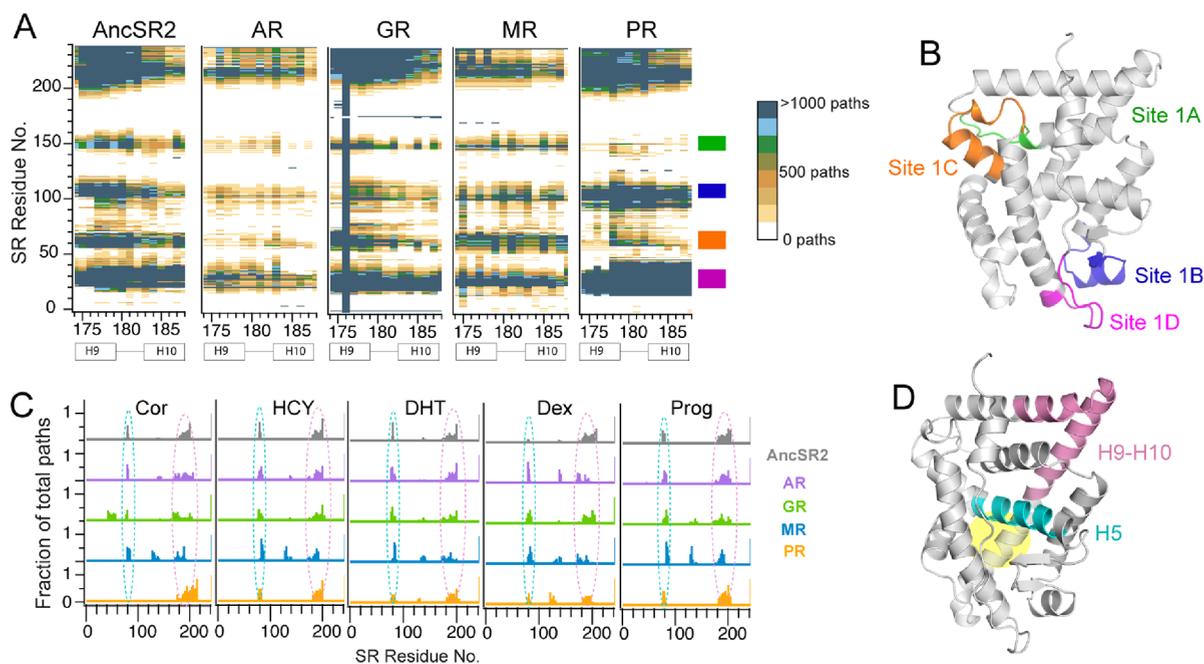
GR (Figure 5D) and MR (Figure 5E) show more variable distributions in the range of edges. Dex and Cor complexes were observed to maintain the same path residues as unliganded GR, while DHT, prog, and HCY incorporated H9 residues (Figure 5A). Heat maps show that interhelical edge weights between residues 181 and 185 (located on N-

terminal of H10) are increased in progesterone and DHT complexes only (Table S3). The weakened H10 edges may drive the preferences for H9 residues in the paths for DHT and prog complexes. We also note that these two ligands are the only GR-inactive hormones as the other three ligands (Cor, HCY, and Dex) are glucocorticoids that activate GR. This observation provides additional evidence of differential ligand modulation of edge weights by active versus inactive ligands. Finally, MR are the only SR for which no ligand-bound complexes utilize the same paths as unliganded MR (Figure 3A) as all ligands impose H10 signaling (Figure 5A). Heat map analysis of edge weights do not provide a clear rationale for this observation as edge weights between unliganded and ligand-bound forms remain similar across the receptor (Table S4). While further studies will be necessary to elucidate the effect of ligands on MR signaling, it is clear that all ligands used here confer structural and/or dynamical changes that alter preferred signaling paths.

**Why Is Site 1 So Important for All SRs?** To explore the basis of the conservation of Site 1 signaling in unliganded and ligand-bound SR complexes, we used the suboptimal-path analysis to identify SR regions allosterically linked to Site 1. Excluding the allosteric connection with H12, which was previously established, we discovered four regions that are strongly connected to Site 1, which are designated as auxiliary



**Figure 5.** Differential effects of ligands on suboptimal paths. (A) Residues in Site 1 allosteric signaling pathways for liganded complexes. For each complex, residues involved in suboptimal paths from AF-2 to Site 1 are shown. The Y axis indicates a normalized percentage utilization of each residue relative to the total combined number of paths between Site 1 and H12. (B–F) Box plots showing weights of edges utilized by unliganded SRs. For each SR, weights in the unliganded (apo) complex are presented. Weights of the same edges in liganded complexes are also presented. Axes for box plots are split to separate edge weights below and above 120.



**Figure 6.** Site 1 is linked to non-AF-2 SR regions. (A) Matrices showing the number of suboptimal paths between all residues and Site 1 (residues 174–188). Other than H12 (residues 225–235), four new auxiliary sites are identified, which are labeled Sites 1A, 1B, 1C, and 1D. (B) Structural model of AncSR2 showing positions of Sites 1A–1D. (C) For ligand-bound SR complexes, residues involved in suboptimal paths between the ligand and Site 1 are shown. All complexes utilize H5 (highlighted in cyan) and H9–H10 (highlighted in pink) for signaling to Site 1. (D) Structural model of AncSR2 highlighting locations of H5 and H9–H10. The position of the ligand-binding pocket is highlighted in yellow.

Sites 1A, 1B, 1C, and 1D (Figure 6A,B). We note that Sites 1A–D are conserved in all five SRs, compared to the less conserved Sites 2–5 (Figure 2A), which suggests that Site 1 may hold higher evolutionary relevance than H12 in SRs.

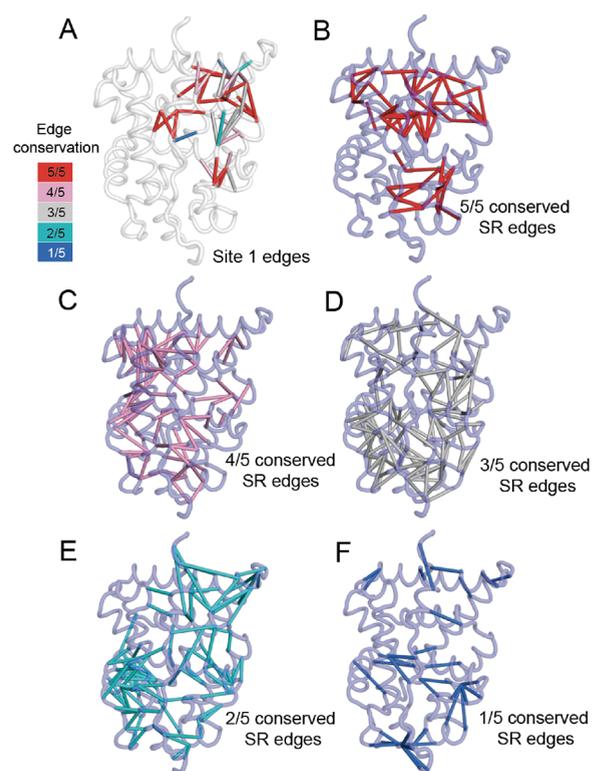
We then identified the residues involved in mediating signaling between Site 1 and the ligand-binding pocket in our SR–hormone complexes (Figure 6C). Consistently, H9/H10 and H5 residues emerge to be crucial for mediating signaling between Site 1 and the bound ligand (Figure 6C,D). We

hypothesized that H9/H10 and H5 function as conduits conserved across SRs that drive strong allosteric connections to Site 1. To test this hypothesis, we sought to identify the residues and edges that allow adjacent SR regions to easily access H9/H10 and H5. We determined all interhelical edges involved in Site 1 signaling, that is, edges formed between H9/H10/H5 and amino acids outside the helices.

We describe the conservation of these edges as 1/5, 2/5, 3/5, 4/5, or 5/5 based on the extent to which residue identities are conserved across the five SRs studied here. For example, the Leu 130–Arg 188 edge connecting H7 and H10 is conserved in AncSR2 (Leu127–Arg185), AR (Leu797–Arg855), GR (Leu656–Arg714), MR (Leu862–Arg920), and PR (Leu811–Arg869), that is, 5/5 conservation. For the 41 interhelical H5/H9/H10 edges identified, we observed that 19 show 5/5 conservation, that is, they are formed by amino acid pairs that are fully conserved across all five SRs. In 10 of the 41 edges, 4/5 residue pairs are conserved. In 7 of the 41 edges, 3/5 residue pairs are conserved. In 2 of the 41 edges, 2/5 conservation is observed. A list of all edges and residue identities is provided (Table S6). To understand how this pattern of conservation compares to all other edges in the SR LBD, we identified all interhelical edges that exist in the SRs studied here, that is, not limited to those involved in Site 1 signaling. We grouped and colored these by amino acid conservation (Figure 7A). Interestingly, the majority of 5/5 conserved edges are localized bridging of H10, H9, H8, and H5 (Figure 7B). For all other edges (4/5–1/5 conserved), we observe a more dispersed distribution (Figure 7C–F). Thus, our identified conduits of Site 1 signaling, that is, H5/H9/H10, also possess the most evolutionarily conserved interhelical connections across the LBD.

Finally, to confirm the unique character of Site 1, we performed a similar identification of SR regions (i.e., auxiliary sites) allosterically linked to Sites 2, 3, 4, and 6 (Figure 8), omitting Site 5 (located on H3) because of its proximity to AF-2 (Figure 2A). We observe that Sites 3 and 6 reveal identical auxiliary sites (Figure 8B,D), while Sites 2 and 4 also possess nearly identical auxiliary sites (Figure 8A,C). This redundancy in auxiliary sites is likely to result from the proximity between the paired sites. Sites 3 and 6, located on  $\beta$ 2/H6 and Loop1–3, respectively, are adjacent to one another in their physical space as is the case with Sites 2 and 4 on H7 and H5, respectively (Figure 2A,B). We observe that, unlike Sites 1A–1D (Figure 6A), auxiliary sites for Sites 2 and 4 are not fully conserved across SRs. Moreover, we notice that two of the predicted auxiliary sites (sites 2C/4C and 2D/4D) are located just a few residues away from Sites 2/4. Site 2C/4C, located on H6, is very close to Site 2 on H7, while Site 2D/4D on H4 is only a few amino acids away from Site 4, which is on H5. Because of their proximity, these auxiliary sites do not represent true allosteric sites with respect to Sites 2 and 4. A similar observation is made for Site 3/6 auxiliary sites as a segment of site 3B/6B (H3) is very close to Site 6 (L1–3). In summary, Site 1 is unique compared to Sites 2–6 identified in our analysis because it is not redundant with other sites and also because all Site 1 auxiliary sites are distant from Site 1, suggesting that they are likely to be true allosteric sites.

**Functional Relevance of Site 1.** The LBD represents the largest known dimerization interface of SRs with H5, H9, H10, and H11 implicated<sup>16,43</sup> along with other structural elements.<sup>44</sup> While there is not a consensus mode of dimerization reported for NRs or even SRs,<sup>7,45–49</sup> H9, H10, and L9–10, which

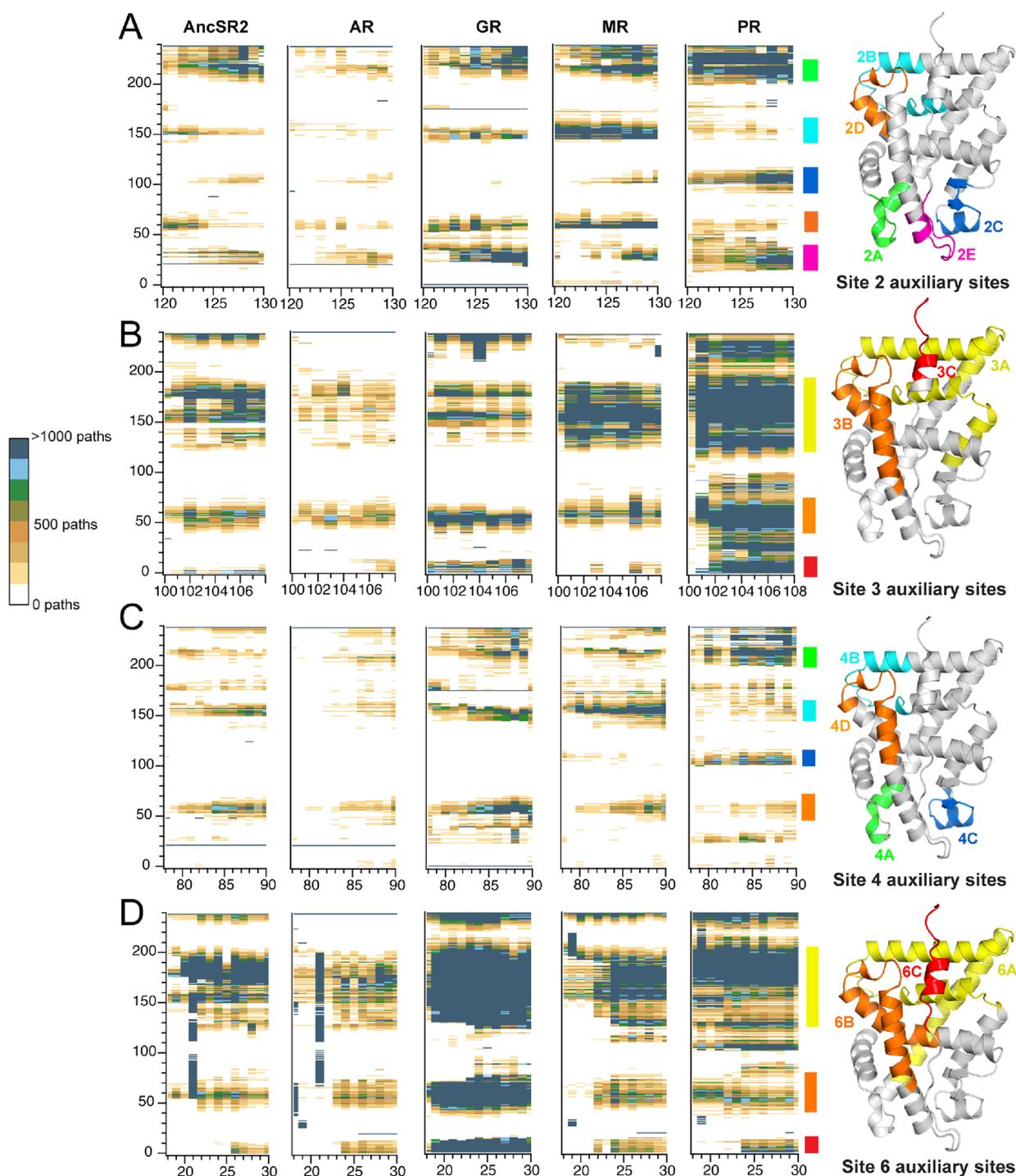


**Figure 7.** Site 1 signaling utilizes conserved edges. (A) Interhelical edges involved in Site 1 signaling are colored by conservation; 5/5 indicates the edge is constituted by a residue pair that is conserved in all five SRs, while 1/5 indicates that the residue pair is non-conserved with a unique identity in each SR. The majority of edges involved in Site 1 signaling are 5/5 conserved. (B) Edges across SRs that are 5/5 conserved. Majority of these are localized in the H5/H8/H10 region. (C) Edges across SRs that are 4/5 conserved. (D) Edges across SRs that are 3/5 conserved. (E) Edges across SRs that are 2/5 conserved. (F) Edges across SRs that are 1/5 conserved.

constitute Site 1 in our study, is the most observed feature associated with receptor dimerization.<sup>50,51</sup> In addition to dimerization functions of Site 1, prior work has identified connections between Site 1 and important LBD sites. A disease-associated H10 mutation in GR located within Site 1 reduced the ligand binding ability, nuclear translocation, and impaired AF-2 surface.<sup>52</sup> Computational analysis predicted a mechanism by which this mutation transmitted conformational changes to AF-2. Other H10 mutations outside of Site 1 have also influenced H12 positioning in ER.<sup>53</sup> L9–10 may also be implicated in LBD–DBD communication in NRs.<sup>54</sup>

## CONCLUSIONS

The SR family of transcription factors displays high structural and functional conservation. Allosteric signaling is also a hallmark of SRs with several ligand-mediated allosteric events shown to be conserved across multiple SRs.<sup>8</sup> In this work, we have used MD simulations to reveal allosterically linked regions in SR ligand-binding domains. By focusing on the four oxosteroid receptors and their resurrected common ancestor, AncSR2, we can uncover sites that are evolutionarily conserved in the family. We identified a site localized between H9 and H10, designated as Site 1, which connects to several critical regions of the receptor, including H12 and the ligand-binding pocket. This allosteric connection is conserved in all SRs studied here, observed in both ligand-bound and unliganded



**Figure 8.** Auxiliary sites linked to Sites 2–6. (A–D) Matrices showing the number of suboptimal paths between all residues and (A) Site 2 (residues 120–130), (B) Site 3 (residues 100–108), (C) Site 4 (residues 78–90), and (D) Site 6 (residues 18–30). Identified auxiliary sites are mapped onto the structural model of AncSR2 and shown to the right of matrices. H12 (residues 225–235) is not identified as an auxiliary site to Sites 2–6 as the connection to H12 is previously established.

complexes. We investigate other sites identified in this work and demonstrate the uniqueness of Site 1. Site 1 holds important significance for SRs and is known to play a role in dimerization of ER, MR, and GR. However, little is known about the importance of Site 1 beyond dimerization. Additionally, we note that our simulations show differences in the strength of Site 1 signaling, that is, the number of suboptimal paths between AF-2 and Site 1. While the reason for these differences are not currently understood, future work

will be geared at determining whether this predicted strength of Site 1 signaling correlates with the functional profile of SRs.

We identified four additional sites that are allosterically linked to Site 1 (Figure 5A). These auxiliary sites are conserved in all SRs. By identifying residues in allosteric paths between the binding pocket and Site 1, we determined that all ligands utilize H9/H10 and H5 for signaling, a potential reason that Site 1 is the only one of the six Sites conserved in all SRs. We observed high levels of conservation

in the H9/H10 and H5 edges used to signal from auxiliary sites to Site 1. We further tested this hypothesis by identifying and classifying all edges that permit interhelical signaling in SR LBDs, finding that the most conserved edges are those connecting H5, H8, H9, and H10. Combined, our investigations suggest that the conservation of Site 1 signaling arises in part from the proximity of these highly conserved edges.

The allosteric networks identified here between Site 1 and both AF-2 and the ligand-binding pocket imply that both ligand and coregulator binding may allosterically regulate SR dimerization and other Site 1-related SR functions. Conversely, modulation of Site 1 may influence ligand binding and coregulator recruitment. Future work will be necessary to define the roles of Site 1 as an allosteric SR modulator.

## METHODS

**Model Preparation.** Five steroid receptor LBD crystal structures were used to prepare starting structures for MD simulations: AncSR2 (PDB: 4LTW),<sup>55</sup> PR (PDB: 1A28),<sup>56</sup> MRs (PDB: 2AA5),<sup>57</sup> GR (PDB: 4P6W),<sup>58</sup> and AR (PDB: 1E3G).<sup>44</sup> All SRs were stripped on N- and C-termini to obtain the same number of aligned amino acids in each LBD: PR (Q682–A922), MR (A733–S973), GR (Q527–N766), AR (E668–S908), and AncSR2 (S-2–A238). Each SR contains 241 amino acids except for GR with 240, resulting from H11 deletion. All water and surface-bound molecules from crystallization buffers were deleted. For clarity, residue numbers 1–241 are used in this manuscript for all SRs. In addition to unliganded SR models, liganded models were generated for each SR using progesterone, cortisol, dihydrotestosterone, corticosterone, and dexamethasone. Ligand complexes were obtained by modifying steroidal cores in each complex to obtain the desired ligand. Our previous use of this approach yielded the same results when compared to complexes obtained by docking ligands into the SR pockets.<sup>4</sup>

**MD Simulations.** All complexes were solvated in an octahedral box using TIP3P water with a cut-off 10 Å buffer around the protein complex. Na<sup>+</sup> and Cl<sup>−</sup> ions were introduced to neutralize the protein and achieve physiological conditions. All systems were set up using tleap in AmberTools<sup>59</sup> with the protein.ff14SB forcefield.<sup>60,61</sup> Parameters for the hormones were obtained using Antechamber<sup>62,63</sup> in AmberTools. Minimizations and simulations were performed with Amber with GPU acceleration.<sup>64,65</sup> Four phases of minimization were performed. First, systems were minimized with 5000 steps of the steepest descent followed by 5000 steps of conjugate gradient minimization with 500 kcal/mol·Å<sup>2</sup> restraints on all atoms. Second, restraints were reduced to 100 kcal/mol·Å<sup>2</sup> and the minimization was repeated. Third, restraints were removed from all atoms excluding ligand atoms and the previous minimization was repeated. Finally, restraints were removed from all atoms for a final minimization. The systems were heated from 0 to 300 K using a 100 ps run with constant volume periodic boundaries and 5 kcal/mol·Å<sup>2</sup> restraints on all protein and ligand atoms. Ten nanoseconds of MD equilibration was performed with 10 kcal/mol·Å<sup>2</sup> restraints on protein and ligand atoms using the *NPT* ensemble. Equilibration was repeated for another 10 ns, keeping very weak restraints of 1 kcal/mol·Å<sup>2</sup> on both the protein and ligand followed by another repeat of 1 kcal/mol·Å<sup>2</sup> weak restraint equilibrium on only ligands. Restraints were finally removed, and triplicate 500 ns production simulations were performed for each system. A 2 fs time step was used, and all

bonds between heavy atoms and hydrogens were fixed with the SHAKE algorithm.<sup>66</sup> A cut-off distance of 10 Å was used to evaluate long-range electrostatics with Particle Mesh Ewald (PME) and for van der Waals forces. The “strip” and “trajout” commands of the CPPTRAJ module<sup>67</sup> were used to remove solvent atoms and extract 25,000 evenly spaced frames from each simulation (75,000 frames total per complex) for analysis.

**Contact Maps and Network Analysis.** The Network View plugin in VMD<sup>36,37</sup> and the Carma program<sup>68</sup> were used to analyze contacts and produce dynamic networks for each system.<sup>32</sup> Residue contact maps were used to determine how dynamic contacts are altered across various complexes.<sup>32</sup> To generate contact maps, all solvent atoms were stripped, leaving ligand and protein atoms. Protein residues are defined as nodes. Edges (or contacts) are created between two non-neighboring nodes if any heavy atoms of the two residues are within 4.5 Å of each other for 75% of the trajectory. To produce dynamic networks, edges in residue contact maps were weighted by covariance calculated from MD simulations (following the protocol described in ref 32). The edge weight is a unitless metric that is inversely proportional to the absolute value of the calculated pairwise correlation between the nodes. A high edge weight connecting two residues corresponds to a low correlation between the residue pair and vice versa. In comparing the same edge between two complexes, a substantially lower or higher edge weight indicates an increase or decrease, respectively, in the correlation between residues comprising the edge, which we interpret as a corresponding increase or decrease in the strength of the edge for mediating communication, respectively.

**Suboptimal Paths.** Communication between the hormone and AF-2 surface was described by generating suboptimal paths between these sites using the Floyd–Warshall algorithm.<sup>69</sup> Communication paths are drawn as a chain of edges connecting the ligand (source node) with a “sink” node on helix 12 (E230). Due to the inverse correlation between the correlation and edge weights, the sum of edges along a path between two distant nodes becomes lower as the strength of communication (i.e., correlation) increases. The optimal path is defined as the path for which the sum of edges is the lowest. For each complex analyzed here, a set of suboptimal paths are obtained by adding a cutoff (50) to the optimal path length and extracting all paths that lie within this length. The choice of cutoff was optimized by performing initial analysis with 25, 50, and 75 (Figure S1).

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.2c01096>.

Additional details on cut-off selection for suboptimal paths analysis, edge weights of unliganded and ligand-bound AncSR2, AR, GR, MR, and PR complexes, and list of interhelical edges/residue pairs identified to be involved in Site 1 signaling (PDF)

## AUTHOR INFORMATION

### Corresponding Author

C. Denise Okafor – Department of Biochemistry and Molecular Biology and Department of Chemistry, Pennsylvania State University, State College, Pennsylvania

16802, United States;  [orcid.org/0000-0001-7374-1561](https://orcid.org/0000-0001-7374-1561);  
Email: [cdo5093@psu.edu](mailto:cdo5093@psu.edu)

## Authors

**Namita Dube** – Department of Biochemistry and Molecular Biology, Pennsylvania State University, State College, Pennsylvania 16802, United States

**Sabab Hasan Khan** – Department of Biochemistry and Molecular Biology, Pennsylvania State University, State College, Pennsylvania 16802, United States

**Riley Sasse** – Department of Chemistry, Pennsylvania State University, State College, Pennsylvania 16802, United States

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.jcim.2c01096>

## Author Contributions

C.D.O. conceived and designed the study. N.D. set up and performed simulations. N.D. and C.D.O. analyzed simulations. N.D., S.H.K., R.S., and C.D.O. wrote the manuscript. All authors have given approval to the final version of the manuscript.

## Funding

This work was supported by a Burroughs Wellcome Fund Award to C.D.O.

## Notes

Amber software necessary for running the MD simulations is publicly available (<https://ambermd.org/>). NetworkView, VMD, Carma, and CatDCD software for generating dynamic networks and performing suboptimal paths analysis are publicly available. Complete MD trajectories utilized for dynamic network analysis are available from the author upon request.

The authors declare no competing financial interest.

## REFERENCES

- (1) Campitelli, P.; Modi, T.; Kumar, S.; Ozkan, S. B. The Role of Conformational Dynamics and Allostery in Modulating Protein Evolution. *Annu. Rev. Biophys.* **2020**, *49*, 267–288.
- (2) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450*, 964–972.
- (3) Elber, R.; Karplus, M. Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. *Science* **1987**, *235*, 318–321.
- (4) Okafor, C. D.; Hercules, D.; Kell, S. A.; Ortlund, E. A. Rewiring Ancient Residue Interaction Networks Drove the Evolution of Specificity in Steroid Receptors. *Structure* **2020**, *28*, 196–205.e3.
- (5) Melo, M. C. R.; Bernardi, R. C.; de la Fuente-Nunez, C.; Luthey-Schulten, Z. Generalized correlation-based dynamical network analysis: a new high-performance approach for identifying allosteric communications in molecular dynamics trajectories. *J. Chem. Phys.* **2020**, *153*, 134104.
- (6) Tenbaum, S.; Baniahmad, A. Nuclear receptors: structure, function and involvement in disease. *Int. J. Biochem. Cell Biol.* **1997**, *29*, 1325–1341.
- (7) Weikum, E. R.; Liu, X.; Ortlund, E. A. The nuclear receptor superfamily: A structural perspective. *Protein Sci.* **2018**, *27*, 1876–1892.
- (8) Okafor, C. D.; Colucci, J. K.; Ortlund, E. A. Ligand-Induced Allosteric Effects Governing SR Signaling. *Nucl. Recept. Res.* **2019**, *6*. DOI: [DOI: 10.32527/2019/101382](https://doi.org/10.32527/2019/101382).
- (9) Aranda, A.; Pascual, A. Nuclear hormone receptors and gene expression. *Physiol. Rev.* **2001**, *81*, 1269–1304.
- (10) Kumar, R.; McEwan, I. J. Allosteric modulators of steroid hormone receptors: structural dynamics and gene regulation. *Endocr. Rev.* **2012**, *33*, 271–299.
- (11) Shao, D.; Lazar, M. A. Modulating nuclear receptor function: may the phos be with you. *J. Clin. Invest.* **1999**, *103*, 1617–1618.
- (12) Nettles, K. W.; Greene, G. L. Ligand control of coregulator recruitment to nuclear receptors. *Annu. Rev. Physiol.* **2005**, *67*, 309–333.
- (13) Mackinnon, J. A.; Gallastegui, N.; Osguthorpe, D. J.; Hagler, A. T.; Estébanez-Perpiñá, E. Allosteric mechanisms of nuclear receptors: insights from computational simulations. *Mol. Cell. Endocrinol.* **2014**, *393*, 75–82.
- (14) Rastinejad, F.; Ollendorff, V.; Polikarpov, I. Nuclear receptor full-length architectures: confronting myth and illusion with high resolution. *Trends Biochem. Sci.* **2015**, *40*, 16–24.
- (15) Estebanez-Perpina, E.; Arnold, L. A.; Nguyen, P.; Rodrigues, E. D.; Mar, E.; Bateman, R.; Pallai, P.; Shokat, K. M.; Baxter, J. D.; Guy, R. K.; Webb, P.; Fletterick, R. J. A surface on the androgen receptor that allosterically regulates coactivator binding. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16074–16079.
- (16) Bledsoe, R. K.; Montana, V. G.; Stanley, T. B.; Delves, C. J.; Apolito, C. J.; McKee, D. D.; Consler, T. G.; Parks, D. J.; Stewart, E. L.; Willson, T. M.; Lambert, M. H.; Moore, J. T.; Pearce, K. H.; Xu, H. E. Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. *Cell* **2002**, *110*, 93–105.
- (17) Connelly, P. J.; Casey, H.; Montezano, A. C.; Touyz, R. M.; Delles, C. Sex steroids receptors, hypertension, and vascular ageing. *J. Hum. Hypertens.* **2022**, *36*, 120–125.
- (18) Eick, G. N.; Colucci, J. K.; Harms, M. J.; Ortlund, E. A.; Thornton, J. W. Evolution of minimal specificity and promiscuity in steroid hormone receptors. *PLoS Genet.* **2012**, *8*, No. e1003072.
- (19) Thornton, J. W. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 5671–5676.
- (20) Eick, G. N.; Thornton, J. W. Evolution of steroid receptors from an estrogen-sensitive ancestral receptor. *Mol. Cell. Endocrinol.* **2011**, *334*, 31–38.
- (21) Ortlund, E. A.; Bridgham, J. T.; Redinbo, M. R.; Thornton, J. W. Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* **2007**, *317*, 1544–1548.
- (22) Issar, M.; Sahasranaman, S.; Buchwald, P.; Hochhaus, G. Differences in the glucocorticoid to progesterone receptor selectivity of inhaled glucocorticoids. *Eur. Respir. J.* **2006**, *27*, 511–516.
- (23) Baxter, J. D.; Funder, J. W.; Apriletti, J. W.; Webb, P. Towards selectively modulating mineralocorticoid receptor function: lessons from other systems. *Mol. Cell. Endocrinol.* **2004**, *217*, 151–165.
- (24) Lange, C. A. Making sense of cross-talk between steroid hormone receptors and intracellular signaling pathways: who will have the last word? *Mol. Endocrinol.* **2004**, *18*, 269–278.
- (25) Krasowski, M. D.; Drees, D.; Morris, C. S.; Maakestad, J.; Blau, J. L.; Ekins, S. Cross-reactivity of steroid hormone immunoassays: clinical significance and two-dimensional molecular similarity prediction. *BMC Clin. Pathol.* **2014**, *14*, 33.
- (26) Kohn, J. A.; Deshpande, K.; Ortlund, E. A. Deciphering Modern Glucocorticoid Cross-pharmacology Using Ancestral Corticosteroid Receptors. *J. Biol. Chem.* **2012**, *287*, 16267–16275.
- (27) Liu, J.; Nussinov, R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLoS Comput. Biol.* **2016**, *12*, No. e1004966.
- (28) Wang, J.; Jain, A.; McDonald, L. R.; Gambogi, C.; Lee, A. L.; Dokholyan, N. V. Mapping allosteric communications within individual proteins. *Nat. Commun.* **2020**, *11*, 3862.
- (29) Schueler-Furman, O.; Wodak, S. J. Computational approaches to investigating allostery. *Curr. Opin. Struct. Biol.* **2016**, 159–171.
- (30) Collier, G.; Ortiz, V. Emerging computational approaches for the study of protein allostery. *Arch. Biochem. Biophys.* **2013**, *538*, 6–15.
- (31) Feher, V. A.; Durrant, J. D.; Van Wart, A. T.; Amaro, R. E. Computational approaches to mapping allosteric pathways. *Curr. Opin. Struct. Biol.* **2014**, 98–103.

- (32) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *16*, 6620–6625.
- (33) Van Wart, A. T.; Durrant, J.; Votapka, L.; Amaro, R. E. Weighted Implementation of Suboptimal Paths (WISP): An Optimized Algorithm and Tool for Dynamical Network Analysis. *J. Chem. Theory Comput.* **2014**, *2*, 511–517.
- (34) Kannan, N.; Vishveshwara, S. Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein Eng.* **2000**, *13*, 753–761.
- (35) Wärnmark, A.; Treuter, E.; Wright, A. P.; Gustafsson, J.-A. k. Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol. Endocrinol.* **2003**, *10*, 1901–1909.
- (36) Eargle, J.; Luthey-Schulten, Z. NetworkView: 3D display and analysis of protein-RNA interaction networks. *Bioinformatics* **2012**, *22*, 3000–3001.
- (37) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (38) Sedláč, D.; Paguio, A.; Bartuňek, P. Two panels of steroid receptor luciferase reporter cell lines for compound profiling. *Comb. Chem. High Throughput Screening* **2011**, *14*, 248–266.
- (39) Katsu, Y.; Oka, K.; Baker, M. E. Evolution of human, chicken, alligator, frog, and zebrafish mineralocorticoid receptors: Allosteric influence on steroid specificity. *Sci. Signaling* **2018**, *11*, No. ea01520.
- (40) Krishnan, A. V.; Zhao, X.-Y.; Swami, S.; Brive, L.; Peehl, D. M.; Ely, K. R.; Feldman, D. A Glucocorticoid-Responsive Mutant Androgen Receptor Exhibits Unique Ligand Specificity: Therapeutic Implications for Androgen-Independent Prostate Cancer. *Endocrinology* **2002**, *143*, 1889–1900.
- (41) Katsu, Y.; Kohno, S.; Oka, K.; Baker, M. E. Evolution of corticosteroid specificity for human, chicken, alligator and frog glucocorticoid receptors. *Steroids* **2016**, 38–45.
- (42) Lin, X.; Takagi, W.; Hyodo, S.; Ijiri, S.; Katsu, Y.; Baker, M. E. Regulation by Progestins, Corticosteroids, and RU486 of Transcriptional Activation of Elephant Shark and Human Progesterone Receptors: An Evolutionary Perspective. *ACS Pharmacol. Transl. Sci.* **2022**, *5*, 52–61.
- (43) Gampe, R. T., Jr.; Montana, V. G.; Lambert, M. H.; Miller, A. B.; Bledsoe, R. K.; Milburn, M. V.; Kliewer, S. A.; Willson, T. M.; Xu, H. E. Asymmetry in the PPAR $\gamma$ /RXR $\alpha$  crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. *Mol. Cell* **2000**, *5*, 545–555.
- (44) Matias, P. M.; Donner, P.; Coelho, R.; Thomaz, M.; Peixoto, C.; Macedo, S.; Otto, N.; Joschko, S.; Scholz, P.; Wegg, A.; Bäsler, S.; Schäfer, M.; Egner, U.; Carrondo, M. A. Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *J. Biol. Chem.* **2000**, *275*, 26164–26171.
- (45) Timmermans, S.; Vandewalle, J.; Libert, C. Dimerization of the Glucocorticoid Receptor and Its Importance in (Patho)physiology: A Primer. *Cell* **2022**, *11*, 683.
- (46) Beinstainer, B.; Markov, G. V.; Erb, S.; Chebaro, Y.; McEwen, A. G.; Cianferani, S.; Laudet, V.; Moras, D.; Billas, I. M. L. A structural signature motif enlightens the origin and diversification of nuclear receptors. *PLoS Genet.* **2021**, *17*, No. e1009492.
- (47) Louw, A. GR Dimerization and the Impact of GR Dimerization on GR Protein Stability and Half-Life. *Front. Immunol.* **2019**, *10*, 1693.
- (48) Wasmuth, E. V.; Broeck, A. V.; LaClair, J. R.; Hoover, E. A.; Lawrence, K. E.; Paknejad, N.; Pappas, K.; Matthies, D.; Wang, B.; Feng, W.; Watson, P. A.; Zinder, J. C.; Karthaus, W. R.; de la Cruz, M. J.; Hite, R. K.; Manova-Todorova, K.; Yu, Z.; Weintraub, S. T.; Klinge, S.; Sawyers, C. L. Allosteric interactions prime androgen receptor dimerization and activation. *Mol. Cell* **2022**, *82*, 2021–2031.e5 e2025.
- (49) Ruff, M.; Gangloff, M.; Wurtz, J. M.; Moras, D. Estrogen receptor transcription and transactivation: Structure-function relationship in DNA- and ligand-binding domains of estrogen receptors. *Breast Cancer Res.* **2000**, *2*, 353–359.
- (50) Jiménez-Panizo, A.; Pérez, P.; Rojas, A. M.; Fuentes-Prior, P.; Estébanez-Perpiñá, E. Non-canonical dimerization of the androgen receptor and other nuclear receptors: implications for human disease. *Endocr.-Relat. Cancer* **2019**, *26*, R479–R497.
- (51) Bianchetti, L.; Sinar, D.; Depenveiller, C.; Dejaegere, A. Insights into mineralocorticoid receptor homodimerization from a combined molecular modeling and bioinformatics study. *Proteins* **2021**, 952.
- (52) Nader, N.; Bachrach, B. E.; Hurt, D. E.; Gajula, S.; Pittman, A.; Lescher, R.; Kino, T. A novel point mutation in helix 10 of the human glucocorticoid receptor causes generalized glucocorticoid resistance by disrupting the structure of the ligand-binding domain. *J. Clin. Endocrinol. Metab.* **2010**, *95*, 2281–2285.
- (53) Nichols, M.; Rientjes, J. M.; Stewart, A. F. Different positioning of the ligand-binding domain helix 12 and the F domain of the estrogen receptor accounts for functional differences between agonists and antagonists. *EMBO J.* **1998**, *17*, 765–773.
- (54) Broekema, M. F.; Massink, M. P. G.; Donato, C.; de Ligt, J.; Schaarschmidt, J.; Borgman, A.; Schooneman, M. G.; Melchers, D.; Gerding, M. N.; Houtman, R.; Bonvin, A. M. J. J.; Majithia, A. R.; Monajemi, H.; van Haaften, G. W.; Soeters, M. R.; Kalkhoven, E. Natural helix 9 mutants of PPAR $\gamma$  differently affect its transcriptional activity. *Mol. Metab.* **2019**, 115–127.
- (55) Colucci, J. K.; Ortlund, E. A. X-Ray Crystal Structure of the Ancestral 3-Ketosteroid Receptor–Progesterone–Mifepristone Complex Shows Mifepristone Bound at the Coactivator Binding Interface. *PLoS One* **2013**, *8*, No. e80761.
- (56) Williams, S. P.; Sigler, P. B. Atomic structure of progesterone complexed with its receptor. *Nature* **1998**, *393*, 392–396.
- (57) Bledsoe, R. K.; Madauss, K. P.; Holt, J. A.; Apolito, C. J.; Lambert, M. H.; Pearce, K. H.; Stanley, T. B.; Stewart, E. L.; Trump, R. P.; Willson, T. M.; Williams, S. P. A Ligand-mediated Hydrogen Bond Network Required for the Activation of the Mineralocorticoid Receptor. *J. Biol. Chem.* **2005**, *280*, 31283–31293.
- (58) He, Y.; Yi, W.; Suino-Powell, K.; Zhou, X. E.; Tolbert, W. D.; Tang, X.; Yang, J.; Yang, H.; Shi, J.; Hou, L.; Jiang, H.; Melcher, K.; Xu, H. E. Structures and mechanism for the design of highly potent glucocorticoids. *Cell Res.* **2014**, *24*, 713–726.
- (59) Case, D.; Belfon, K.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham, III, T.; Cruzeiro, V.; Darden, T.; Duke, R.; Giambasu, G.; Gilson, M.; Gohlke, H.; Goetz, A.; Harris, R.; Izadi, S.; Izmailov, S.; Kasavajhala, K.; Kovalenko, A.; Krasny, R.; Kurtzman, T.; Lee, T.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Man, V.; Merz, K.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Roe, D.; Roitberg, A.; Sagui, S.; Schott-Verdugo, S.; Shen, J.; Simmerling, C.; Skrynnikov, N.; Smith, J.; Swails, J.; Walker, R.; Wang, J.; Wilson, L.; Wolf, R.; Wu, X.; Xiong, Y.; Xue, Y.; York, D.; Kollman, P. *AMBER 2020*. University of California: San Francisco 2020.
- (60) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (61) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Miguez, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16*, 528–552.
- (62) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* **2001**, U403.
- (63) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (64) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555.

(65) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.

(66) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(67) Roe, D. R.; Cheatham, T. E., III PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(68) Glykos, N. M. Software news and updates carma: A molecular dynamics analysis program. *J. Comput. Chem.* **2006**, 1765–1768.

(69) Floyd, R. W. Algorithm 97: shortest path. *Commun. ACM* **1962**, *5*, 345.