Supplemental Material: Theory

Computational Methods and additional discussion of the avidin/Re-functionalized biotin derivatives is available in the article text. This section describes the methodology and results from intermediate modeling studies of the avidin-biotin complex and related structures. Reference numbers refer back to the full article.

Biotin-derivative accommodation within the avidin binding pocket was examined by stepwise replacement of the biotin valeryl carboxyl group with a series of linear hydrocarbon chains. The avidin binding pocket for biotin is composed of three regions, two of which are centered about the fused ureido and thiophene rings and a third that is localized to the valeryl carboxyl group (**Figure Supp1A**).[24] The biotin ureido ring is the site of sterically complementary hydrogen bonding interactions to residues Asn118, Tyr33 and Ser16, and Thr35 (**Figure Supp1B**). The remarkable avidin binding affinity for biotin (Kd ~ 10^{-15} M [T12]) is in no small part due to the localization of this hydrogen bonding framework to the inside of the avidin β-barrel. The thiophene region of the binding pocket for the monomeric avidin is comprised of aromatic residues Trp97, Phe79, and Trp70 (**Figure Supp1C**). In the tetrameric avidin crystal structure, the aromatic pocket includes Trp110 from a second monomeric avidin unit. These seven (or eight) residues serve to virtually engulf the fused ureido and thiophene rings within the β-barrel binding pocket, serving the dual roles of biotin binding and alignment. The valeryl carboxyl tail is bound within the third binding region to residues Ala39, Thr40, and Ser73 (**Figure Supp1D**). Residues Ala39 and Thr40 are located on the avidin 3-4 loop and each contributes one N-H donor to a hydrogen bonding interaction with the hydroxyl oxygen of the carboxyl tail. Residue Ser73, located on the avidin 5-6 loop, forms a single hydrogen bond between the serine hydroxyl group and the carbonyl end of the biotin tail. Ordering of the 3-4 and (to a lesser extent) 5-6 loops by biotin hydrogen bonding leaves enough room only for a chain of waters that hydrogen bond to the biotin carboxyl end and trail into the solvent. The biotin fatty acid tail is also in close contact with the indole ring of Trp70, which runs along the long axis of the hydrocarbon chain.

An MD simulation of the avidin-biotin complex was performed to determine what modifications to the structure of the binding pocket occur in the absence of crystal waters and the remainder of the avidin-biotin tetramer. By reproducing the binding pocket in the native complex, it could be demonstrated that either the biotin derivatives were retained in the binding pocket in a similar manner and that subsequent steric deformations of the binding pocket were, in a sense, allowed structurally, or that these large modifications served to unbind biotin from the inner-β-barrel region and, therefore, destabilize the modified avidin-biotin complex in a manner inconsistent with the experimental results.

A 5000 ps MD simulation of the avidin-biotin complex revealed two changes to the structure of the biotin binding pocket. The first change was a breaking of the H-O···O=C hydrogen bond in the biotin carboxyl tail to Ser73 in the 5-6 loop and the formation of a local hydrogen bond between the N-H group of Ser73 and the backbone oxygen of Trp70 (**Figures Supp2A and Supp2B**). The largely unrestricted motions of the 3-4 and 5-6 loops resulted in their increased separation relative to the initial position of the carboxyl end of the biotin tail, which followed the 3-4 loop by the retention of its two hydrogen bonds. Over the course of the

simulation, however, the hydroxyl oxygen is exchanged with the carbonyl oxygen as the hydrogen bond acceptor to Ala39 and Thr40. The 3-4 and 5-6 loop interactions with biotin are, of course, very dependent on the loop positions. The specific flex of the 3-4 loop is familiar to both theoretical and experimental studies, as its motion is the only major structural change to occur upon biomolecule binding to biotin[24] and the only significant motion required in the modeling of the unbinding of biotin from avidin.[25]

The second change to the binding pocket resulted in changes to the thiophene/ureido binding regions by the removal of the hydrogen bond between biotin and Asn118 and the formation of a single hydrogen bond between Asn118 and Trp97 (**Figures Supp2C and Supp2D**). The treatment of the tetrameric structure by the same methods results in this Asn118-biotin interaction being retained,[25] leading to the conclusion that this broken interaction is a result of the treatment of only the monomeric avidin-biotin complex. Neglecting possible parameterization limitations, two explanations related to structural features can be considered. The first is the long distance of this β-barrel interaction. Asn118 lies across the β-barrel from residues Ser16, Tyr33, and Thr35. As restricted in the crystal lattice, the β-barrel may undergo some slight structural change in strand 8 (which contains the Asn118 residue) that leads to interaction breaking. The removal of this interaction in favor of a closer range hydrogen bonding interaction between Asn118 and Trp97 in the monomeric form may also be due to the removal of more significant inter-monomeric interactions. The tetrameric avidin-biotin complex packs by van der Waals and hydrogen bonding interactions along strand 8.[25] The absence of packing interactions may relieve strand strain from inter-monomeric alignment, weakening the Ser16-biotin hydrogen bond enough to break it over the MD simulation. While removing two hydrogen bonding interactions and weakening one hydrophobic interaction in the binding region, these changes result in only a marginal shifting of the position of the biotin in the binding pocket due to the three strong hydrogen bonds remaining in the β-barrel, the overall retention of shape in this region of the avidin, and, over the time scale of the simulation, the only transient disruption of the hydrogen bond between the biotin carboxyl end and Ser73 (5-6 loop).

Biotin Decarboxylation MM/MD. Biotin decarboxylation removes the three hydrogen bonds to the avidin 3-4 and 5-6 loops. Accordingly, this modification is expected to have a significant impact on the motion at the fatty acid end while affecting inner-β-barrel binding negligibly. The MD simulation results for the decarboxylated biotin are nearly identical to those of the avidin-biotin complex, with the exception of those changes expected at the 3-4 and 5-6 loops. The thiophene/ureido ring binding pockets change as observed in the unaltered biotin case. The removal of the carboxyl tail results in disorder between Ala39 and Thr40, as noted by the new disorder of the two N-H hydrogen bond donor orientations. The retention of shape and biotin binding overall within the β-barrel in the absence of interactions to the flexible 3-4 and 5-6 loops is consistent with the results required to observe Re complex binding in the experimental study presented here. The flex in the 3-4 loop has been considered previously in theoretical studies of the unbinding of biotin from avidin and was found to be an integral part of the unbinding process.[25]

Biotin-Hydrocarbon Derivatives MM/MD. The localization of binding pocket deformation with biotin derivative enlargement is shown for a number of the hydrocarbon tail lengths in **Figure Supp3**. Extension of the hydrocarbon tails into the underlying pocket of the 3-

4 loop (and, to a lesser extent, the 5-6 loop) results in one of the tails protruding into the groove previously occupied by the crystal waters, a second tail sliding along the length of strand 8 of the β-barrel, and the third tail extending into the gap formed by the incomplete closure of the 3-4 loop around this end of the β-barrel. With the removal of the hydrogen bonding interactions between the biotin and the 3-4 loop, there is no electrostatic interaction to impose any order on this surface region. Consequently, the deformation of the 3-4 loop at each step in the CH_3 addition process follows the hydrocarbon tails consistent with close-packing interactions.

The closure of the longest three hydrocarbon chains to form a single, cryptand-like structure removes the flexibility of these chains and localizes the bulk/volume of the modified biotin end to a single location along the 3-4 loop (**Figure Supp3E**). This localization results in the largest 3-4 loop shift and the most pronounced avidin structural changes from among the series of hydrocarbon tail simulations. Despite these large shifts, however, the ureido and thiophene binding pocket shifts negligibly from the cases of unrestricted hydrocarbon tail motion. All avidin deformations that occur to accommodate the larger ligands occur as surface deformations in the 3-4 loop.

Biotin-Re Complex MM. The smaller of the two Re complexes is shown in surface (avidin) and van der Waals (biotin-Re) renderings in **Figure Supp4**. The MM energy minimization results in 3-4 loop disorder relative to the unsubstituted biotin structure that is consistent with the larger hydrocarbon chain lengths. The opening of the 3-4 loop pocket from steric interactions is again observed with negligible shift in the position of the ureido ring/thiophene ring binding pocket inside the β-barrel structure, although some slight deformations of the local β-sheets themselves can be observed relative to the avidin-biotin case.

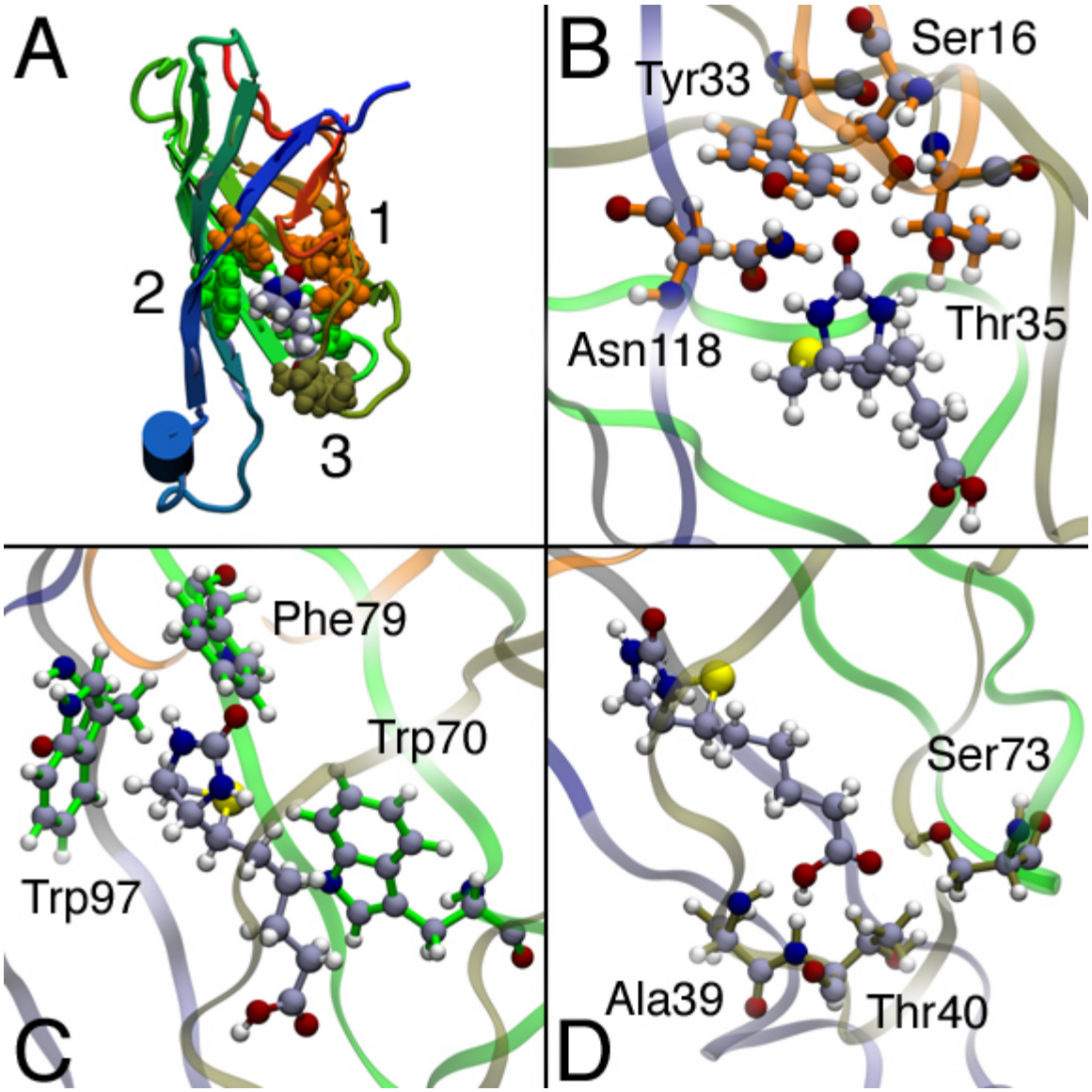

Figure Supp1. The three regions of the biotin binding pocket in the avidin-biotin complex. The entire pocket, shown in **A**, identifies the positions of the three regions (by color) expanded in **B** through **D**. Region 1, shown in B, identifies the four residues involved in hydrogen bonding within the β-barrel. Region 2, shown in C, identifies the three residues involved in hydrophobic interactions. Region 3, shown in D, identifies the three residues involved in carboxyl tail hydrogen bonding.

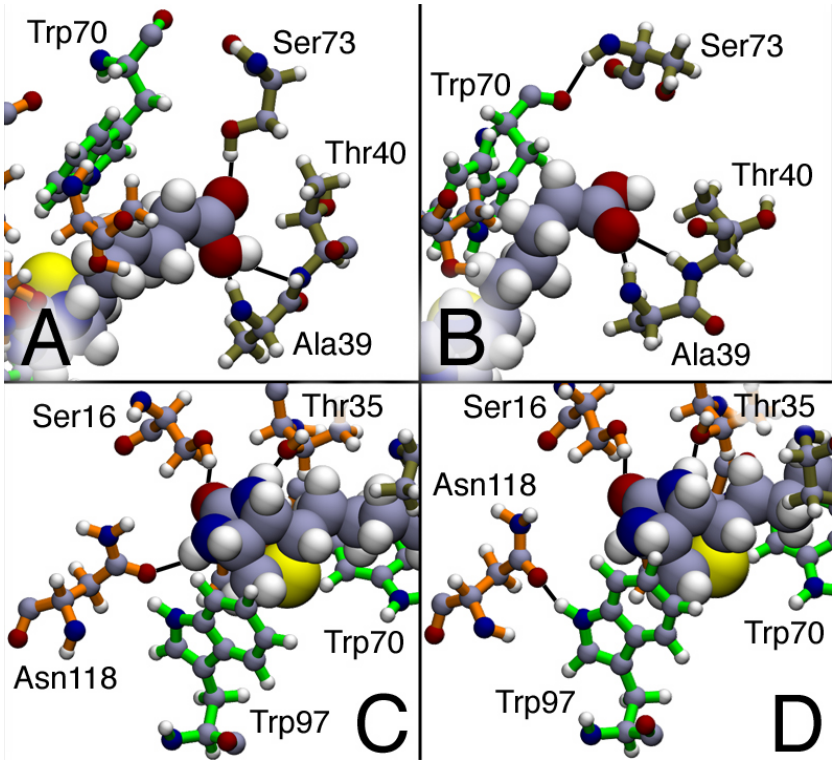

Figure Supp2. The changes to the biotin binding pocket over the course of the 5000 ps MD simulation. **A** (0 ps) and **B** (5000 ps) identify the changes to the carboxyl tail hydrogen bonding that occur during the simulation. **C** (0 ps) and **D** (5000 ps) identify the β-barrel change between Asn118 and Trp97 that occurs and persists during the simulation.

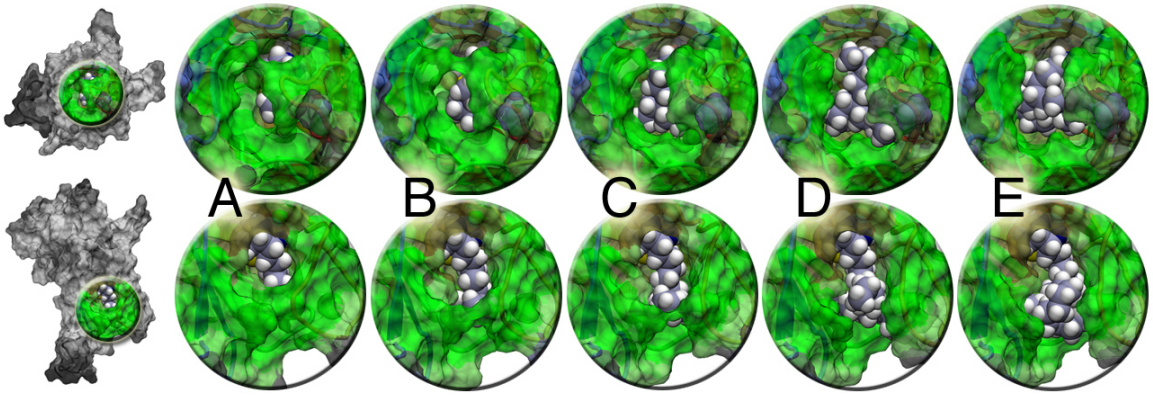

Figure Supp3. Two views of the enlargement of the biotin binding pocket by hydrocarbon tail elongation. The entire avidin-biotin structure and highlighted biotin binding pocket are shown at left. The functionalized biotins are shown in van der Waals representation and the avidin is rendered as a surface. The five shown pendant groups are R = H (A), R = CH_2CH_3 (B), R = $(CH_2)_2CH_3$ (C), R = $(CH_2)_3CH_3$ (D), and CR_3 = $C(CH_2)_5)_3CH$ (E).

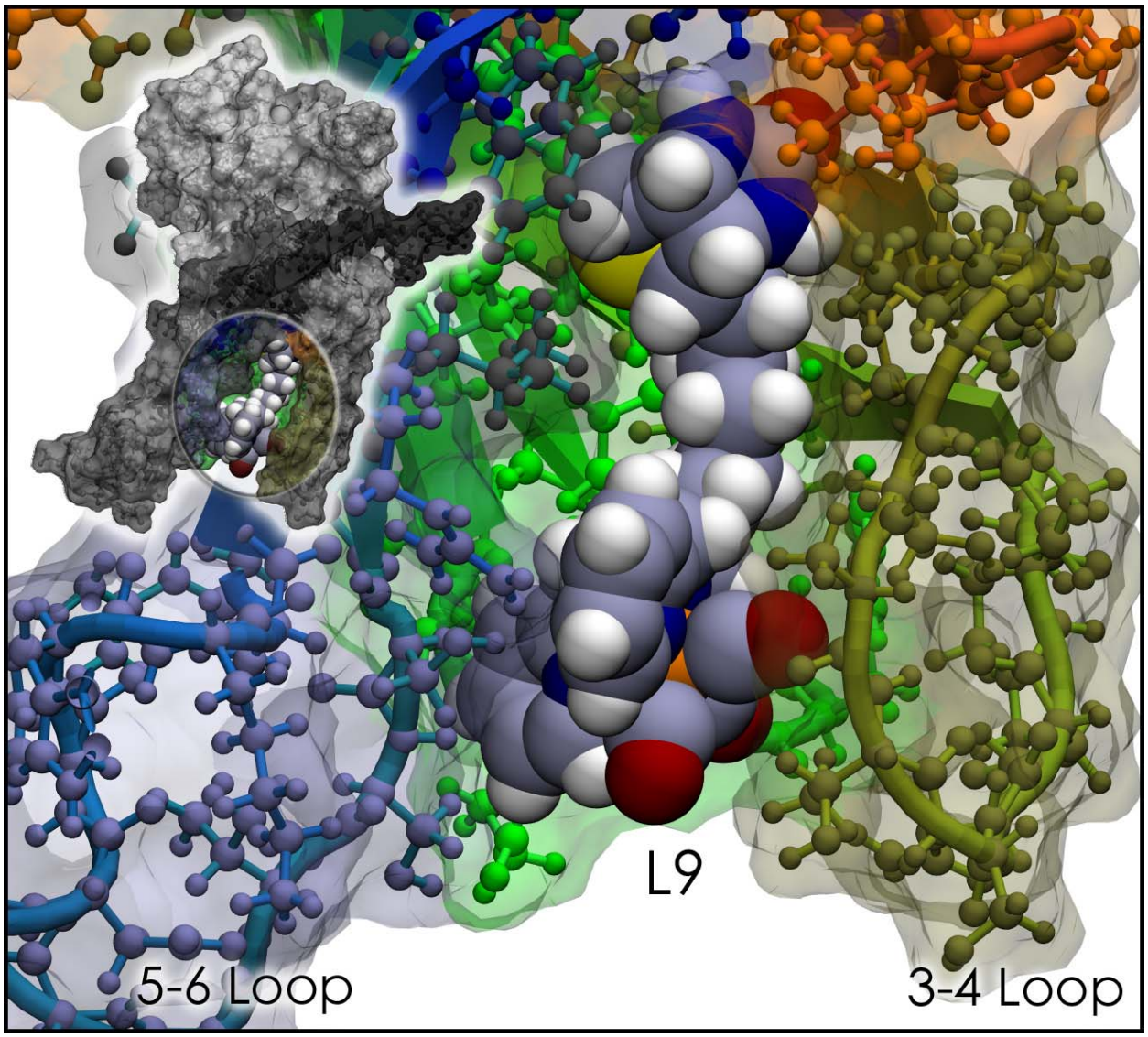

Figure Supp4. A biotin-Re ligand bound within the avidin-biotin binding pocket. The entire avidin-biotin-Re complex is shown (inset) with the binding pocket highlighted. The biotin-Re ligand is shown in van der Waals representation. The 3-4 loop is at bottom-right (gold).