Kinetic Characterization and Computational Modeling of the *Escherichia coli* Heptosyltransferase II: Exploring the Role of Protein Dynamics in Catalysis for a GT-B Glycosyltransferase

Bakar A. Hassan[1], Zhiqi A. Liu[1], Jozafina Milicaj[1], Mia S. Kim[1], Meka Tyson[1], Yuk Y. Sham[2,3], Erika A. Taylor[1,*]

[1]Department of Chemistry, Wesleyan University, Middletown, CT, 06459

[2]Bioinformatics and Computational Biology Program, University of Minnesota, Minneapolis, MN 55455

[3]Department of Integrative Biology and Physiology, Medical School, University of Minnesota, Minneapolis, MN 55455

*Corresponding Author Information - Tel: 860-685-2739. Fax: 860-685-2211; E-mail: eataylor@wesleyan.edu.

Keywords

Heptosyltransferase, glycosyltransferase, GT-B, tryptophan fluorescence, structural rearrangement, circular dichroism, protein dynamics, molecular dynamics simulations, kinetics, binding affinities, docking, multiple sequence alignment

Abbreviations

glycosyltransferase (GT), heptosyltransferase II (HepII), L-*glycero*-D-*manno*-heptose (Hep), ADP-L-*glycero*-D-*manno*-heptose (ADP-Hep), lipopolysaccharide (LPS), *E. coli* Hep-Kdo_2-Lipid A (HLA), O-deacylated *E. coli* Hep-Kdo_2-Lipid A (ODHLA), tryptophan (Trp), phenylalanine (Phe), circular dichroism (CD), molecular dynamics (MD), toll-like receptor 4 (TLR4), Carbohydrate-Active enZYme (CaZY), ADP-β-mannose (ADP-Man), molecular mechanics generalized Born surface area (MMGBSA), Generalized Amber Forcefield (GAFF2), Luria-Bertani (LB), column volume (CV), Pyruvate Kinase/Lactate Dehydrogenase (PK/LDH), multiple sequence alignment (MSA), melting temperature (T_M)

Abstract

Glycosyltransferases (GTs) are enzymes that are uniquely adapted to promote the formation of a glycosidic bond between a sugar molecule and a wide variety of substrates. Heptosyltransferase II (HepII) is a GT involved in the lipopolysaccharide (LPS) biosynthetic pathway that transfers the seven-carbon sugar (L-*glycero*-D-*manno*-heptose; Hep) onto a lipid anchored glycopolymer (heptosylated Kdo_2-Lipid A, Hep-Kdo_2-Lipid A or HLA). LPS plays a key role in Gram-negative bacterial sepsis as a stimulator of the human immune response and has been used as an adjuvant in vaccines. As such, ongoing efforts towards inhibition of LPS biosynthetic enzymes to aid development of novel antimicrobial therapeutics has driven significant effort towards the characterization of these enzymes. Three heptosyltransferases are involved in the inner-core biosynthesis, with *E. coli* HepII being the last to be quantitatively characterized *in vivo,* as described herein. HepII shares modest sequence similarity with heptosyltransferase I (HepI) while maintaining a high degree of structural homology. Here we report the first kinetic and biophysical characterization of HepII and demonstrate the properties of HepII that are shared by HepI to include sugar donor promiscuity, and sugar acceptor induced secondary structural changes which results in significant thermal stabilization. HepII also has an increased catalytic efficiency and a significantly tighter binding affinity for both of its substrates, with an insensitivity to the number of acyl chains on the sugar acceptor. Additionally, a structural model of the HepII ternary complex, refined by molecular dynamics simulations, was developed to probe potentially important substrate-protein contacts and revealed the potential of Tryptophan (Trp) residues responsible for reporting on ligand binding. As was previously described for HepI, Tryptophan fluorescence in HepII allowed observation of

substrate induced changes in Trp fluorescence intensity which enabled determination of substrate dissociation constants. Combined, these efforts meaningfully enhance our understanding of the Heptosyltransferase family of enzymes and will aid in future efforts to design novel, potent and specific inhibitors for this family of enzymes.

Introduction

The ever growing concern of multi-drug resistant bacterial infections is invigorating efforts towards the development of novel antimicrobials that could be used as standalone treatments or as part of combinatorial therapeutics.[1-3] In Gram-negative bacteria, the lipopolysaccharide (LPS) is a major contributor to virulence and it acts as a barrier to protect against xenobiotics including hydrophobic antibiotics.[4] The LPS is a lipid anchored glycopolymer that is incorporated into the outer leaflet of the outer membrane of Gram-negative bacteria. LPS can elicit the human immune response through interactions with toll-like receptor 4 (TLR4) and plays a role in bacterial sepsis.[5] This modulatory behavior has inspired some to use LPS and various analogs as vaccine adjuvants.[6] The fully synthesized LPS is a major component of the outer leaflet of the outer membrane, making up approximately 75% of the bacterial membrane surface,[17] with LPS having been shown to be important for the formation of bacterial biofilms in multiple organisms,[7] making the LPS biosynthetic enzymes targets for inhibitor development to enable multiple medical applications.[8-12] The LPS consist of the acylated Lipid A base, an inner core and outer core oligosaccharide region, and O-antigen repeating region (Figure 1A). The LPS Lipid A and core region are both assembled on the inner leaflet of the inner membrane, while the O-antigen is added after the truncated LPS is transported onto the bacteria's surface via

cross membrane transporters (Figure 1B).[15-16] The Lipid A and inner core are highly conserved among Gram-negative bacteria making them good targets for the design of potential LPS biosynthesis inhibitors that would have broad spectrum activity. Previous *in vivo* experiments have demonstrated that mutations that lead to truncation of LPS at the inner core exhibit a phenotype that includes an increased susceptibility to hydrophobic antibiotics due to a compromised outer membrane.[18] Heptosyltransferase I and II (HepI, HepII), enzymes involved in the inner core biosynthesis, are found in a wide genus of clinically relevant bacteria including Enterobacteriaceae, Pasteurellaceae, Pseudomonadaceae, Moraxellaceae, Vibrionaceae, Burkholderiaceae and Neisseriaceae.[19] Therefore, inhibition of HepI or HepII *in vivo* would lead to a truncated LPS[20] making these enzymes highly desirable drug targets.

HepI has been thoroughly characterized both *in vivo* and *in vitro*, with some nanomolar inhibitors successfully designed for it to date.[21-31] HepII, however, had previously only been functionally characterized *in vivo*, with an apo crystal structure (PDB: 1PSW; Uniprot: P37692) with no corresponding manuscript.[26, 32-33] HepI and HepII function consecutively in the LPS inner core biosynthetic pathway. HepI utilizes an ADP-L-*glycero*-β-D-*manno*-heptose (ADP-Hep) and transfers the heptose moiety onto the Kdo_2-Lipid A forming an α(1→5) glycosidic bond yielding ADP and Hep-Kdo_2-Lipid A (Figure 1B). HepII utilizes an identical sugar donor ADP-Hep and transfers the heptose moiety onto the Hep-Kdo_2-Lipid A (product of HepI reaction) forming an α(1→3) glycosidic bond to form ADP and Hep_2-Kdo_2-Lipid A (Figure 1-2). These enzymes catalyze near identical reactions while having a surprisingly low degree of sequence identity (32%; Figure S1). They are both part of the GT-B structural class of glycosyltransferases and the GT-9 structural family in the Carbohydrate-Active enZYme (CaZY) database (www.cazy.org). As

members of the GT-B structural class, they have two independent α/β/α Rossmann-like domains connected by a linker region (Figure 3). As in all GT-B enzymes, the N-terminal domain binds the nucleophilic, sugar acceptor substrate and the C-terminal domain binds the activated sugar donor. It was previously demonstrated that HepI can utilize a hexose analog sugar donor, ADP-β-mannose (ADP-Man),[31] while also tolerating removal of acyl chains from the acceptor;[24] these two prior results motivated our investigation of the substrate selectivity of HepII. Additionally, qualitative *in vivo* data has demonstrated HepII's capabilities of utilizing ADP-Man as an alternative sugar donor while also demonstrating that the enzyme has a greater catalytic efficiency relative to HepI; however, these findings have never before been validated *in vitro*.[26] Herein, we report the first characterization of *Escherichia coli* HepII including kinetic and fluorescence based *in vitro* studies, as well as computational modeling and molecular dynamics *in silico* methods to gain a greater insight into the structural properties that govern the functional characteristics of this enzyme. The study of HepII will enable the future design of inhibitors that act as potent and specific antimicrobials with synergistic properties in combinatorial therapeutics with other hydrophobic antibiotics.

Materials and Methods

Multiple Sequence Alignment and Sequence Conservation Analysis:

Multiple Sequence alignments and sequence conservation of HepI and HepII were obtained from the consurf webserver.[34-37] Briefly, *E. coli* HepI and HepII amino acid sequences were obtained from crystal structures (HepI: PDB 6DFE, HepII: PDB 1PSW). Homologous sequences for each protein were retrieved from the uniref90 server and yielded 2028 unique sequences for HepI and 1994 unique sequences for HepII using a sequence similarity cutoff

including only sequences with 95% to 35% sequence identity. Sequences were ordered by ascending E values and 150 representative sequences were chosen equally spaced from the list of available sequences. Multiple sequence alignment for homologues and pairwise sequence alignment between HepI and HepII were performed with Clustal Omega.[38] Resulting sequence alignment figures were made with the ENDscript server.[39]

Molecular Modeling and Molecular Dynamic Simulation Refinement:

The ternary model of HepII was made from the available apo crystal structure (PDB: 1PSW). Missing sidechains were modeled with Prime from the Schrodinger suite.[40-41] The previously developed ternary model of HepI[42] served as template for the placement for ADP-Hep, due to the high degree of sequence similarity occurring in the C-terminal, ADP-Hep binding domain. The C-terminal domains of HepI (residues 179-322) and HepII (residues 182-348) were aligned to aid placement of the ADP-Hep. A deacylated analog of the Hep-Kdo_2-Lipid A acceptor substrate was docking into the center of the N-terminal cavity with a 30 Å x 30 Å x 30 Å grid box via Glide.[43] The final docked poses were constrained to maintain contact with one of four residues (Trp10, Asp13, Lys92, Lys195) via hydrogen bonds. These residues were chosen due to their high level of sequence conservation amongst HepII homologues and their structural superposition with HepI residues that were previously determined to be in contact with its sugar acceptor. The top 5 poses according to their docking score were subjected to MMGBSA[44] and the top two poses were selected for molecular dynamics simulations. pK_as for ionizable sidechains were calculated using PROPKA with 11 frames equally spaced from the second half of the trajectory. [45-46] Simulations were performed with the GROMACS-2021.1 simulation package and the Amber99SB forcefield.[47-49] Ligands were prepared as previously

described.[42] Briefly, Atom types were assigned from the second generation Generalized Amber Forcefield (GAFF2) and charges were assigned with the AM1-BCC model.[50-51] PROPKA was utilized in assigning ionization states of titratable sidechains of HepII.[45-46] The system was placed into a dodecahedron periodic boundary condition with a 10 Å buffer region and solvated with TIP3P[52] water model. Furthermore, the system was neutralized with the addition of 0.150 M sodium and chloride counterions. Energy minimization was performed for 50000 steps with the steepest descent algorithm. The system was equilibrated with subsequent isothermal-isochoric (NVT) and isothermal-isobaric (NPT) ensembles for a total of 12 ns with a 2 fs timestep. During equilibration, restraints (1000 kJ/mol/nm^2) were applied to all atoms and progressively removed from the sidechains and then the backbone. Ligands were restrained during equilibration to allow the rearrangement of the active site to accommodate the ligands prior to the production simulations. Temperature and pressure were maintained with the v-rescale thermostat and Berendsen barostat with coupling of 0.1 and 2.0 ps, respectively.[48, 53] Production simulations were performed for a total of 500 ns at 300 K and 1 atm (NPT ensemble) and the temperature maintained with the v-rescale thermostat and the pressure maintained with the Parrinello-Rahman barostat with isotropic coupling.[54] Long range electrostatics were calculated with the particle-mesh-ewald (PME) with a fourth order cubic interpolation and 1.6 Å grid spacing.[55] Short range nonbonded interactions were calculated with a 10.0 Å cutoff. Bonds to hydrogen atoms were constrained with the LINCS algorithm.[56] Ligand interaction diagrams were made with Maestro, molecular models were made in PyMOL and reactions were drawn in ChemDraw.[41, 57]

Expression and Purification of HepII:

A pET28a plasmid containing the *E. coli* HepII (rfaF) was graciously provided by the New England Structural Genomics Consortium. The plasmid was transformed into *E. coli* BL21-AI cells (Invitrogen) and plated on Luria-Bertani (LB) agar with kanamycin (50 µg mL^{-1}), with incubation for 18 hours at 37°C. One isolated colony was transferred from the plate to 10 mL of LB kanamycin (Kan) solution and incubated overnight at 37°C with shaking at 220 rpm. One liter of LB kanamycin was inoculated with 10 mL of the overnight seed culture and grown at 37°C with shaking at 220 rpm until an OD_{600} of 0.55 was reached. Protein expression was induced with a combination of isopropyl ß-D-1-thiogalactopyranoside (IPTG) and arabinose at final concentrations of 1 mM and 0.0002%, respectively. Upon induction, the temperature was reduced to 30°C and the cells were allowed to continue growing for an additional 24 hours. Cells were collected from the media by centrifugation at 5,400 x g for 10 minutes and resuspended in 20 mL of a binding buffer (20 mM HEPES, 1 mM imidazole, 0.5 M NaCl, pH 7.5) containing Lysozyme (1.25 mg mL^{-1}), pepstatin A (1 µg mL^{-1}) and aprotinin (1 µg mL^{-1}). Cells were lysed with an Emulsiflex –C5 homogenizer at 13,000 psi and the lysate was subsequently clarified by centrifugation at 21,000 x g for 30 minutes.

A 10 mL column of Toyopearl AF Chelate 650M resin was charged with 3 column volumes (CV) of 10 mM cobalt sulfate and equilibrated with 3 CV of binding buffer (see above). Lysate was loaded at 1.0 mL/min and weakly bound proteins were eluted with an additional 3 CV of bind buffer. HepII was eluted with an imidazole gradient from a low imidazole buffer (20 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 40 mM imidazole, 0.5 M NaCl, pH 7.5) to a high imidazole buffer (20 mM HEPES, 1M imidazole, 0.5 M NaCl, pH 7.5). Fractions containing HepII were pooled, concentrated via centrifugal ultrafiltration (30 KDa MWCO

Vivaspin 20) and buffer exchanged via Biorad P6 Desalting cartridge into a storage buffer (10 mM tris-HCl, 150 mM $NaSO_4$, 150 mM arginine (Arg), 40% glycerol, pH 7.5). The protein was further concentrated by centrifugal ultrafiltration (30 KDa MWCO Vivaspin 20) to 10 mg/mL, flash-frozen with liquid nitrogen and stored at -80°C. HepII can be stored in this condition for 6 months without significant loss in activity.

Sugar Acceptor Isolation and Deacylation:

E. coli deletion strains of HepII (ΔrfaF::Kan) and the subsequent kinase (ΔrfaP::Kan) were purchased from the Yale Keio collection.[58] Sugar acceptor substrates were isolated from their respective knockout cell lines and prepared, as previously described.[59-60] Briefly, two liters of cell were grown to an OD_{600} of 1 and pelleted at 5,400 x g for 10 minutes. Cells were washed with 40 mL of water, 40mL of ethanol, twice with 40 mL of acetone and once with diethyl ether. After each solvent wash, cells were pelleted at 5,400 x g and the supernatant was discarded. Cells were air dried at room temperature for one hour and pulverized with mortar and pestle prior to lipid extraction. One gram of dried cells was combined with 20 mL of a 2:5:8 phenol/chloroform/petroleum ether solution for 30 minutes while rocking. Cellular debris was pelleted at 5,400 x g for 10 minutes and supernatant was poured into a 100 mL round bottom flask. Extraction was repeated once more with the cellular debris and the resulting supernatant was post centrifugation was combined with the previously extracted supernatant in the 100 mL round bottom flask. Chloroform and petroleum ether were removed under reduced pressure. 75 mL of acetone, 15 mL of diethyl ether and 3 drops of water were added to the resulting solution to precipitate the Hep-Kdo_2 Lipid A. The precipitant was pelleted at 20,000 x g for 10 minutes and washed with 2 mL of 80% phenol and 2 mL of diethyl ether, with the precipitate

being pelleted at 20,000 x g for 10 minutes between each wash, with the supernatant being discarded. The washed precipitate was dissolved in 20 mL of water containing 0.5% triethylamine, frozen, and then lyophilized to dryness.

Hep-Kdo_2 Lipid A was O-deacylated with hydrazine, with 10 mg of lipid reacting with 1 mL of hydrazine with stirring at 37°C for one hour. The reaction was quenched and the lipid was precipitated with 10 mL of ice cold acetone. The precipitate was pelleted at 20,000 x g for 30 minutes and the supernatant was discarded. The precipitate was washed separately with cold acetone and diethyl ether with centrifugation at 20,000 x g for 30 minutes between each wash step, with the supernatant being discarded. The washed precipitate was dissolved in 20 mL of water, frozen, and lyophilized. The formation of the O-deacylated product was confirmed via ESI mass spectrometry.

Sugar Donor Isolation/Biosynthesis:

The native sugar donor (heptosylated adenosine diphosphate, ADP-Hep) was extracted and purified from *E. coli* WBB06 cells as previously described.[29, 60] Briefly, two liters of WBB06 cells were grown in LB tetracycline (10 µg mL^{-1}) to an OD_{600} of 1. Cells were collected by centrifugation at 5,400 x g for 10 minutes and the supernatant was discarded. Pelleted cells were resuspended in 40 mL of 50% ethanol (1:1 ethanol:H_2O) for 30 minutes on ice, with stirring. The cellular suspension was centrifuged at 3,000 x g for 10 minutes to pellet cellular debris and the supernatant was kept for further processing while the pellet was discarded. Ethanol was removed under reduced pressure with a vacuum centrifuge until the volume of the sample was reduced to 20 mL. The sample was flown through an Amicon 3000 MWCO centrifugal filter via centrifugation at 3,000 x g to remove proteins that would interfere with

downstream isolation and characterization of the sugar donor (ADP-Hep). The flowthrough was loaded onto a 50 mL DEAE (diethylethanolamine) column and ADP-Hep was eluted with a 0-1 M triethylammonium bicarbonate buffer (pH 8.0) gradient. Fractions containing ADP-Hep were identified via ESI (electrospray ionization) mass spectrometry. Fractions of ADP-Hep were pooled, lyophilized, and stored at -80°C. NMR data is consistent with previously published data.

The ADP-Hep analogue sugar donor ADP-mannose (ADP-Man) was biosynthesized with slight modifications as previously described.[61] Briefly, 30 mg of mannose was combined with, 184 mg of ATP and 9.0 mg of HldE in 6 mL of a reaction buffer consisting of 500 mM tris-HCl (pH 8.0) and 5 mM $MgCl_2$. The reaction was allowed to proceed overnight with monitoring via ^{31}P-NMR spectroscopy. 10 µU of calf intestinal alkaline phosphatase (CIP) was added to dephosphorylate remaining ATP/ADP/AMP. The reaction mixture was then flowed over a Bio-Rad P6 desalting cartridge and lyophilized. Isolation of product was confirmed via NMR spectroscopy (^{1}H, ^{13}C, ^{31}P).

Circular Dichroism Spectroscopic Analysis:

Protein secondary structure and thermostability were determined via Circular Dichroism as previously described for HepI.[23] Spectra were recorded on a Jasco J-810 Spectropolarimeter and scans were obtained at varying temperatures (10°C - 95°C) between 190 nm to 250 nm in quartz cuvettes (0.2 x 0.1 cm). All conditions were performed in triplicate with 5 µM HepII in a buffer consisting of 10 mM tris-HCl (pH 7.5) and 100 mM KCl. In conditions where substrates/products (donor/acceptor) were present, 100 µM of these compounds were used to ensure complete binding of substrate to enzyme. Resulting spectra were analyzed via Matlab

R2018a and melting temperature (T_M) was calculated by fitting data to a sigmoid curve and determining the temperature at which the protein was 50% unfolded.

Tryptophan Fluorescence Binding Studies:

Emission spectra were collected on a Horiba Fluoromax-4 spectrofluoremeter. Samples were excited at 295 nm with a 6 nm slit width and resulting emission spectra was collected from 350 nm to 410 nm with a 8 nm slit width. Excitation and emission polarizers were set to 0° and 57.3°, respectively. HepII was maintained at 1 µM concentration and ligands were varied in concentration from 1 nM to 10 µM. Independent samples were made for each ligand concentration to avoid photobleaching and each concentration was repeated in triplicate. Resulting emission spectra were fit to a log normal distribution to determine the intensity at λ_{max}. The intensities were then fit to a binding curve to determine the binding affinity (K_D) with the equation 1, where [E·S] is presumed to be the fluorescence intensity that is a function of the total enzyme concentration ($[E]_{total}$) and substrate concentration ($[S]_{total}$).[62-63] Resulting errors are reported as standard errors.

$$[E \cdot S] = \frac{([E]_{total} + [S]_{total} + K_D) - \sqrt{([E]_{total} + [S]_{total} + K_D)^2 - 4 \cdot [E]_{total} \cdot [S]_{total}}}{2 \cdot [E]_{total}} \quad (1)$$

HepII Enzyme Kinetics:

Enzymatic rates were determined via pyruvate kinase/lactate dehydrogenase (PK/LDH) coupled assay as previously described.[23-24] Each reaction mixture contained 10 nM HepII, 50 U of PK, 50 U of LDH, 100 uM phosphoenol pyruvate, 100 µM NADH and 100 µM dithioerythritol in a reaction buffer (50 mM HEPES, 50 mM KCl, 10 mM $MgCl_2$, pH 7.5). When varying one substrate, the other substrate was introduced at 10,000 fold excess (100 uM). The sugar

donor/acceptor were varied between 100 nM to 20 µM. Reduction of NADH was monitored at 340 nm with a constant temperature of 37° C on a Cary UV-Vis spectrophotometer. Reactions were initiated by the addition of enzyme. Reactions involving the lowest concentration of substrate (100 nM) approached the sensitivity limit of the instrument, therefore, initial rates were determined by fitting the full progress curve rather than the first 10% with equation 1.[64] The Michaelis constant (K_M) and turnover number (k_{cat}) were calculated by fitting initial rates as a function of substrate concentration to the Michaelis-Menten equation.

$$[S](t) = K_M \cdot W\left\{\frac{[S]_0}{K_M} \cdot e^{\frac{[S]_0 - V_M \cdot t}{K_M}}\right\} \quad (2)$$

Results

Multiple Sequence Alignment and Sequence Conservation Analysis:

HepII catalyzes the transfer of a heptose onto the growing LPS inner core region (Figures 1-2). This enzymatic reaction is directly preceded by HepI that similarly transfers a heptose residue onto the Kdo_2-Lipid A, producing the substrate for HepII (Figure 1). Based upon previous molecular biology experiments, both heptosyltransferase enzymes use an ADP-L-*glycero*-D-*manno*-heptose (ADP-Hep) as their sugar donor substrate. A sequence similarity comparison between *E. coli* HepI and HepII reveals a 32% sequence identity, with common residues located mainly within the C-terminal Rossmann-domain, where the ADP-Hep is expected to bind (Figure 3, S1). Furthermore, as previously observed in a multiple sequence alignment of HepI-III, they share a highly conserved Asp13 that is believed to facilitate the transfer reaction by acting as a

base during the sugar transfer reaction (Figure 2, S1-S2).[25, 65] To elucidate the residues potentially involved in HepII substrate interactions, we performed a sequence alignment of HepI with HepII (HepI PDB: 6DFE and HepII PDB:1PSW; Figure S1), in addition to generating representative multiple sequence alignments (MSAs) of HepI (Figure S2) and HepII (Figure 4) homologues. Previously identified residues in HepI that are involved in substrate binding are indicated[21, 25, 42] and were used as a guide to identification of residues in a HepII that might be important for catalysis, in addition to examining residues that are highly conserved in a multiple sequence alignment (MSA) of HepII. Based upon these analyses of HepII, we hypothesized FDHLA binding residues to be Pro8, Trp10, His61, Ser90, Lys92, Gly192, Lys195, and Asp269 (Table 1A, S1). For the native sugar donor, HepII potentially important residues involved in ADP-Hep binding include Trp10, Glu190, Gly192, Lys195, Thr250, Leu252, Ala255, Asp269, Ser270, Gly271, Leu272, and His274 (Table 1B, S1). In our multiple sequence alignment of HepII (Figure 4), a high degree of conservation of Asp13 in HepII homologues from a variety of species was observed, which is consistent with observations that Asp is a conserved catalytic base for inverting GT-B enzymes.[66] Furthermore, each of the HepII residues listed above maintain at least 90% conservation in a representative MSA generated from 150 HepII homologues spanning 35-95% identity to the HepII from *E. coli*.

Ternary Model Development and Refinement:

Both HepI and HepII are part of the GT-B structural class, and a structural alignment yields a superposition with a C_α root mean square deviation of their structures of 3.8. Due to the high sequence similarity of the C terminal domains of HepI and HepII, placement of the sugar donor (ADP-Hep) in HepII was modeled (Figure 3) through a superposition of previously

observed crystal structure orientation of the complex ADP-Hep with HepI (PDB 2H1H). For the sugar acceptor substrate (FDHLA), we docked this ligand in the N-terminal domain of HepII, so as to maintain a minimum of one hydrogen bond contact with each highly conserved residues identified from the MSA analysis predicted to be involved with FDHLA binding, including residues Trp10, Asp13, Lys92, Lys195 (Table 1A). In pose2, there is electrostatic complementarity in the active site with arginines and lysines in the active site (Figure 3). The top two poses were subjected to further structural minimization via 500 ns molecular dynamics (MD) simulation. The apo form was also simulated to provide a reference for the unbound complex.

The backbone root mean square deviation (RMSD) across the second half of the trajectory for apo, pose1, pose2 complexes are 2.08 ± 0.16 Å, 2.12 ± 0.18 Å, 1.88 ± 0.10 Å, respectively (Figure 5A, S11A, Table S3). The C^{α} root mean square fluctuations (C^{α}RMSF) for apo and pose1 reveals residues with fluctuations greater than 1.5 Å in the 60s and a majority of the C-terminal domain (Figure S11B, Table S3). Whereas pose2 only has a 15 residues with fluctuations greater than 1.5 Å in the 230s and 300s (Figure 5B, S11B, Table S3). To assess the merits of the predicted ligand binding residues from the MSA examination (above), the HepII•ADP-Hep•FDHLA ternary complex simulation was used to identify amino acids adjacent to the ligands and ligand interaction maps were generated (Figure 6). In the simulation of the ternary complex (pose2), the acceptor (FDHLA) is held in place by electrostatic interactions between Arg69, Lys92, Lys316 and the phosphate groups on the acceptor (Figure 6A, S12A). The acceptor is further stabilized by hydrogen bonding interactions with other residues including Ser9, Trp10, and Asn89. Interactions between the acceptor and HepII in pose1 only differs by an

electrostatic interaction, where the Lys316 of pose2 is replaced by Lys125 in pose1 (Figure S13A, S14A). The donor (ADP-Hep) is stabilized by electrostatic contact between the alpha phosphate and Lys195 (Figure 6B, S12B). Furthermore, a hydrogen bonding interaction between the primary amine of the adenosine ring and Thr250 sidechain aids in holding the donor in place. There is also a T shaped pi-pi stacking interaction between Trp40 and the indole ring of the adenine. Lastly, there are additional hydrogen bonding interactions between ADP-Hep and residues Glu190, Gly271, Asp269, His 274 and Lys313. In pose1, very similar interactions are observed and only differing by the lack of pi-pi or hydrophobic interactions between the acceptor and Trp40 (Figure S13B, S14B).

Lastly, the pK_a of ionizable sidechains were calculated to probe whether the catalytic residue has a perturbed pK_a in-line with the hypothesized mechanism of the transfer reaction (Figure 2). The pK_a of Asp13 in HepII apo, ternary complex pose1 and pose2 are 5.31 ± 0.46, 5.26 ± 0.71, and 6.17 ± 0.32 (Table S4). A similar elevation of the pK_a of Asp13 in HepI was previously observed.[42]

Experimental Secondary Structural Characterization of HepII:

HepII from *E. coli* was successfully purified to greater than 90% purity (Figure S3). The secondary structural characteristics of HepII, as determined by CD spectral analysis, suggest a protein with a mixed α/β structure in the apo state (Figure 7A). In the presence of ODHLA, HepII gains a more pronounced second minima, suggestive of an increase in alpha helicity. The melt temperature of HepII apo is 60^{o} C (Figure 7B, S4A). In the presence of the sugar acceptor (ODHLA), HepII maintains secondary structure at higher temperatures (95^{o} C) with an increase in the overall signal intensity at the same protein concentration (Figure 7B, S4A-B). The sugar

acceptor product (ODH_2LA) has similar stabilizing affect and induces the same secondary structural changes as ODHLA, but with a decrease in the signal intensity at the same protein concentration (Figure 7B, S7E). ADP and ADP-Hep have similar melt temperatures and do not exhibit any substrate induced stabilizing affect or secondary structure change as with ODHLA or ODH_2LA (Figure 7B, S7B-D).

HepII Substrate Binding and Kinetics:

HepII has seven tryptophan residues that can act as potential reporters for binding. As mentioned above, Trp10 is in the active site and sits in the space between both ligands and is highly conserved (Table 1). In the presence of the native sugar donor (ADP-Hep) or acceptor (ODHLA), HepII's tryptophan fluorescence emission maximum exhibits a decrease in signal intensity as a function of concentration as compared to that of the apo protein (Figure 8A, S5). Using this change in fluorescence intensity upon ligand binding, binding affinities of native substrates to HepII were determined. The native sugar donor (ADP-Hep) has a K_D of 0.66 ± 0.46 µM (Figure 8B, S5, Table 2). For the sugar acceptor substrate analog (O-deacylated-Hep-Kdo_2 Lipid A) has a K_D of 0.48 ± 0.14 µM (Table 2, Figure S5).

To gain further insight into HepII catalysis, the turnover rate and Michaelis equilibrium constant were determined for the native/non-native sugar donors and acceptors. For the native sugar acceptor, the kinetics were determined in the presence and absence of Triton X-100 to ensure micelle formation of the acceptor was not inhibiting the reaction (Table 2, S1, Figure S6). For HLA with (without) triton X-100, the k_{cat} and K_M are 0.57 ± 0.06 s^{-1} (0.33 ± 0.03 s^{-1}) and 0.49 ± 0.13 µM (0.41 ± 0.09 µM), respectively. The O-deacylated-Heptosyl-Kdo_2 Lipid A has a k_{cat} of 0.423 ± 0.007 s^{-1} and K_M of 0.58 ± 0.08 µM. When varying sugar donor concentrations, the ADP-

Hep (ADP-Man) k_{cat} and K_M are 0.47 ± 0.02 s^{-1} (0.54 ± 0.02 s^{-1}) and 0.57 ± 0.09 µM (0.77 ± 0.08 µM), respectively.

Discussion

Conserved Binding Site Residues:

HepI and HepII facilitate consecutive transferase reactions with an identical sugar donor (ADP-Hep) and a sugar acceptor where the product of the HepI reaction is the substrate of the HepII reaction (Figure 1-2). It is therefore not surprising that most of the sequence conservation between these two enzymes occurs in the C-terminal domain (Figure S1). More surprising is the lack of sequence conservation in the N-terminal domain to facilitate binding of HLA which is a common product/substrate between the two enzymes. Examination of conserved active site residues reveals maintenance of the catalytic base and three of nine positively charged arginine (Arg) and lysineslysine (Lys) residues that in HepI are responsible for anchoring the negatively charged phosphate groups on the acceptor via electrostatic interactions. Only the homologues of HepI residues Lys98, Arg 143 and Lys192, are maintained and conserved HepII (corresponding to residues Lys92, Arg146 and Lys195 in HepI). Arg146 is greater than 20 Å from the nearest phosphate on the ligand, therefore, it may not play a major role in substrate binding and stabilization. The Lys195 stabilizes the donor and is unavailable to bind to the acceptor in the C-terminal domain (Figure 6). Charged residues that bind to the acceptor, suggested by our simulations include Arg69, Lys125, Lys313, Arg315, and Lys316. Residues Lys313, Arg315 and Lys316 are conserved 17%, 24% and 57% of the time, respectively. Therefore, they are likely to play a minor role in substrate binding. Arg69 is conserved 98% of

the time, therefore, it may coordinate to the same phosphate as Lys92 (Figure 6A). Lys125 is conserved 73% of the time and is a better candidate for anchoring the acceptor substrate in the active site. We hypothesize that these conserved charged residues may facilitate binding of the acceptor substrate and subsequent promotion of conformational changes that are necessary for catalysis by bringing the nucleophilic hydroxyl of the acceptor close to the anomeric carbon of the donor (Figure 6A, S12A, S14A), as has been previously observed in HepI.[42] Both enzymes have a highly conserved Asp13 that potentially acts as the catalytic base that removes the proton from the hydroxyl group of the acceptor to promote a nucleophilic attack on the anomeric carbon (Figure 2). This is further suggested by the perturbed pK_as of Asp13 calculated for HepII and previously calculated for HepI (Table S4).[42] HepI has a serine at the 10th position, whereas, HepII has a tryptophan (Trp), so no other potentially conserved catalytic residues have been identified. Due to the added sugar on the HepII substrate, the stabilizing hydrogen bond from the serine in HepI may have been replaced by a tryptophan to form hydrophobic interactions with the hydrophobic face of this added sugar. Since Trp10 also is predicted to be adjacent to the bound ADP-Hep, this residue could be important for form anion-π interactions[67] with the phosphates of ADP-Hep. Alternatively, as suggested by our modeling, the highly conserved Trp40 (92%) may be interacting with the indole ring of the donor and contributes to the fluorescence quenching that is observed experimentally for the donor, while the substrate induced quenching of the acceptor is due to hydrophobic interactions with the acceptor (Figure 6). In addition to the Trp10, HepII has several other HepI equivalent residues that similarly stabilize ADP-Hep binding. The backbone of Glu190 amine forms hydrogen bonds with the phosphates of ADP-Hep, similar to the HepI equivalent Thr187. The HepI residue Met 242 which

is strictly conserved and responsible for binding of the NH_2 group of the adenine base through a backbone carbonyl, is now replaced by Thr250, which is conserved 98% of HepII homologues analyzed. Furthermore, HepII seems to have a highly conserved heptose "sensing" residue Asp269 that hydrogen bonds to the hydroxyl groups of the heptose moiety, analogous to Asp261 in HepI. Overall, examination of the sequence and structural superpositions of HepI and HepII have allowed identification of a collection of amino acid residues that warrant further investigation for their involvement in ligand binding and catalysis in subsequent studies.

Binding Induced Structural Changes:

In the crystal structure of HepII, there are two stretch of residues that are disordered (58-64, 311-320). The flexibility of these two regions are evident in the RMSF including downstream residues 65-77 (Figure 5B, S11B). Residues 58-64 are analogous to residues 58-70 in HepI which have been previously observed to adopt a more alpha-helical orientation upon ligand binding and has been observed via CD and through crystal structures (Apo: PDB 2GT1, Psuedoternary: PDB 6DFE).[21, 23, 25, 42] In our HepII simulations, residues 58-64 become less dynamic upon presence of the acceptor in both poses (Figure 5B, S11B) which could promote formation of an ordered secondary structure. Additionally, the disordered 60s loop adds a turn to the downstream helix and the helix gains a turn on either side (Figure S15). This *in silico* observation of a ligand induced secondary structural rearrangement is also supported by observed changes in the CD spectral comparisons of HepII apo and liganded forms (Figure 7A). These combined observations suggest that HepII experiences a net increase in α helicity in the presence of the sugar acceptor and this could be attributed to the stabilization of the 60s α helix downstream of the disordered loop (58-64) (Figure 7A, S15). Furthermore, residues Lys195

in the C- terminal domain also becomes ordered by the presence of the sugar acceptor, whereas the acceptor binds in the N-terminal domain. Engagement of both domains by the acceptor may explain why only in the presence of the acceptor does HepII experience significant thermal stabilization.

Based on our ternary structural model of HepII, we hypothesized that binding of either substrate would have an impact on the tryptophan fluorescence emission spectra because of the proximity of Trp10 or Trp40 to both ligands (Figure 8A). Indeed, we observe a concentration dependent quenching of tryptophan fluorescence upon binding of both the sugar donor and the sugar acceptor. We hypothesize that this tryptophan may be there to assist in the exclusion of water from the active site by acting as part of a cage surrounding the Asp13, the acceptor hydroxyl and the anomeric donor. This tryptophan may also be important in intercalating into the membrane to enhance acceptor substrate binding of HepII to enable catalysis. Our modeling also suggests a possible role for Trp40 aiding in substrate binding of the donor. Further studies are needed to definitively understand the mechanistic role of Trp10 and Trp40, including studies of HepII binding to a membrane anchored acceptor substrate embedded in a micelle, to elucidate its role (if any) in catalysis.

Substrate Selectivity Evaluation:

While the enzyme that catalyzes the Kdo-transfer reaction early in the inner-core LPS biosynthesis is sensitive to the number of acyl chains on the acceptor[68] the same is not true of HepI or HepII. The deacylation of the HepI and HepII sugar acceptors provide tremendous practical advantages for experiments by increasing the solubility of the substrate and eliminating the need for detergents. As previously demonstrated, HepI experiences a 2-5 fold

increase in the catalytic efficiency upon deacylation of the acceptor mainly due to a decrease in the K_M.[24] On the contrary, the catalytic efficiency of HepII is unaffected by the presence or absence of acyl groups on the acceptor where an increase in the turnover rate (k_{cat}) is counterbalanced by an increase in the K_M (Table 2, Table S1). Furthermore, as previously demonstrated *in vivo*[26] and quantified in our work, HepII has an approximately 2.3 fold increase in its catalytic efficiency due to a greater affinity for the acceptor substrate as compared by the differences in the K_M. Additionally, HepII can also utilize a hexose (ADP-Man) sugar donor, as previously observed in HepI. The K_M for the hexose is approximately 1.4 fold lower than the native heptose, again being less impacted than HepI which exhibits a nearly 10 fold decrease in the Michaelis constant for the same reduction in susbtrate size.[69] The binding affinities determined by tryptophan fluorescence quenching yield similar affinities for O-deacylated acceptor and native donor. Overall, the substrate utilization profiles for the two enzymes are quite similar and the ability of HepII to also utilize a sugar acceptor substrate with fewer acyl chains has the potential to allow for future work to obtain liganded crystal structures of HepII to validate the predicted ligand-binding interactions described here.

Conclusion

Our work here has provided the first quantitative determination of HepII functionality and its similarities and differences with HepI. In light of the low sequence similarity between HepI and HepII, their structural homology and multiple sequence alignments have allowed us to infer the conserved residues that may contribute to substrate binding and catalysis. This information has also allowed us to create a working model for the ternary complex HepII, as

previously done for HepI.[42] This working model inspired the use of tryptophan fluorescence to determine the binding affinities of native and non-native donor and acceptor ligands, which also helps substantiate the validity of the ternary structural model of HepII. HepII has a higher catalytic efficiency and a submicromolar affinity for its substrates, which is distinct from prior observations for HepI. Further studies will need to be performed to fully understand the importance of these differences in kinetic parameters. HepII's insensitivity to the number of acyl chains on the acceptor may be a result of the longer sugar polymer region and suggests that HepII may not need to rely on the lipids as another source of ligand contacts. This result coupled with the smaller K_M for HepII binding of the sugar acceptor substrate could enable HepII to be a better target for inhibitor design studies, but further research is needed to determine if tight-binding inhibitors can be developed for this enzyme. Additionally, changes in secondary structure and thermal stability in the presence of the acceptor is akin to HepI and may be a conserved behavior for heptosyltransferase enzymes, or perhaps GT-B enzymes more broadly. This work presents new insights and confirms previously determined *in vitro* characteristics of HepII that paves the way for the future design of novel therapeutics against this family of enzymes.

Accession codes

More information available about Heptosyltransferase II from *E. coli* can be found in the PDB (PDB ID: 1PSW) and Uniprot databases (ID: P37692).

Acknowledgements

Daniel J. Czyzyk and Noreen K. Nkosana for assistance in determining the conditions for HepII protein expression, the New York Structural Genomics Research Consortium for a plasmid containing the subcloned *E. coli* HepII gene and the Minnesota Super Computing Institute for the Schrodinger Suite.

References

1. Pérez-Rodríguez, F.; Mercanoglu Taban, B., A State-of-Art Review on Multi-Drug Resistant Pathogens in Foods of Animal Origin: Risk Factors and Mitigation Strategies. *Frontiers in Microbiology* **2019,** *10*, 2091.
2. Tanwar, J.; Das, S.; Fatima, Z.; Hameed, S., Multidrug resistance: an emerging crisis. *Interdiscip Perspect Infect Dis* **2014,** *2014*, 541340-541340.
3. Bassetti, M.; Peghin, M.; Vena, A.; Giacobbe, D. R., Treatment of Infections Due to MDR Gram-Negative Bacteria. *Frontiers in Medicine* **2019,** *6*, 74.
4. Zhang, G.; Meredith, T. C.; Kahne, D., On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr Opin Microbiol* **2013,** *16* (6), 779-785.
5. Raetz, C. R. H.; Whitfield, C., Lipopolysaccharide endotoxins. *Annu Rev Biochem* **2002,** *71*, 635-700.
6. Baldrick, P.; Richardson, D.; Elliott, G.; Wheeler, A. W., Safety Evaluation of Monophosphoryl Lipid A (MPL): An Immunostimulatory Adjuvant. *Regulatory Toxicology and Pharmacology* **2002,** *35* (3), 398-413.
7. Muhammad, M. H.; Idris, A. L.; Fan, X.; Guo, Y.; Yu, Y.; Jin, X.; Qiu, J.; Guan, X.; Huang, T., Beyond Risk: Bacterial Biofilms and Their Regulating Approaches. *Frontiers in Microbiology* **2020,** *11*.
8. De Leon, G. P.; Elowe, N. H.; Koteva, K. P.; Valvano, M. A.; Wright, G. D., An In Vitro Screen of Bacterial Lipopolysaccharide Biosynthetic Enzymes Identifies an Inhibitor of ADP-Heptose Biosynthesis. *Chemistry & Biology* **2006,** *13* (4), 437-441.
9. Erwin, A. L., Antibacterial Drug Discovery Targeting the Lipopolysaccharide Biosynthetic Enzyme LpxC. *Cold Spring Harb Perspect Med* **2016,** *6* (7), a025304.
10. Zhang, G.; Baidin, V.; Pahil, K. S.; Moison, E.; Tomasek, D.; Ramadoss, N. S.; Chatterjee, A. K.; McNamara, C. W.; Young, T. S.; Schultz, P. G.; Meredith, T. C.; Kahne, D., Cell-based screen for discovering lipopolysaccharide biogenesis inhibitors. *Proceedings of the National Academy of Sciences* **2018,** *115* (26), 6834.
11. Han, W.; Ma, X.; Balibar, C. J.; Baxter Rath, C. M.; Benton, B.; Bermingham, A.; Casey, F.; Chie-Leon, B.; Cho, M.-K.; Frank, A. O.; Frommlet, A.; Ho, C.-M.; Lee, P. S.; Li, M.; Lingel, A.; Ma, S.; Merritt, H.; Ornelas, E.; De Pascale, G.; Prathapam, R.; Prosen, K. R.; Rasper, D.; Ruzin, A.; Sawyer, W. S.; Shaul, J.; Shen, X.; Shia, S.; Steffek, M.; Subramanian, S.; Vo, J.; Wang, F.; Wartchow, C.; Uehara, T., Two Distinct Mechanisms of Inhibition of LpxA Acyltransferase Essential for Lipopolysaccharide Biosynthesis. *Journal of the American Chemical Society* **2020,** *142* (9), 4445-4455.
12. Cho, J.; Lee, M.; Cochrane, C. S.; Webster, C. G.; Fenton, B. A.; Zhao, J.; Hong, J.; Zhou, P., Structural basis of the UDP-diacylglucosamine pyrophosphohydrolase LpxH inhibition by sulfonyl piperazine antibiotics. *Proceedings of the National Academy of Sciences* **2020,** *117* (8), 4109.

13. Amor, K.; Heinrichs, D. E.; Frirdich, E.; Ziebell, K.; Johnson, R. P.; Whitfield, C., Distribution of core oligosaccharide types in lipopolysaccharides from Escherichia coli. *Infect Immun* **2000,** *68* (3), 1116-1124.
14. Knirel, Y. A., Structure of O-Antigens. In *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Knirel, Y. A.; Valvano, M. A., Eds. Springer Vienna: Vienna, 2011; pp 41-115.
15. Chng, S.-S.; Gronenberg, L. S.; Kahne, D., Proteins Required for Lipopolysaccharide Assembly in Escherichia coli Form a Transenvelope Complex. *Biochemistry* **2010,** *49* (22), 4565-4567.
16. Mi, W.; Li, Y.; Yoon, S. H.; Ernst, R. K.; Walz, T.; Liao, M., Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature* **2017,** *549* (7671), 233-237.
17. Le Brun, A. P.; Clifton, L. A.; Halbert, C. E.; Lin, B.; Meron, M.; Holden, P. J.; Lakey, J. H.; Holt, S. A., Structural Characterization of a Model Gram-Negative Bacterial Surface Using Lipopolysaccharides from Rough Strains of Escherichia coli. *Biomacromolecules* **2013,** *14* (6), 2014-2022.
18. Coleman, W. G., Jr.; Leive, L., Two mutations which affect the barrier function of the Escherichia coli K-12 outer membrane. *Journal of bacteriology* **1979,** *139* (3), 899-910.
19. Holst, O., Structure of the Lipopolysaccharide Core Region. In *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Knirel, Y. A.; Valvano, M. A., Eds. Springer Vienna: Vienna, 2011; pp 21-39.
20. Brabetz, W.; Muller-Loennies, S.; Holst, O.; Brade, H., Deletion of the Heptosyltransferase Genes rfaC and rfaF in Escherichia Coli K-12 Results in an Re-Type Lipopolysaccharide with a High Degree of 2-Aminoethanol Phosphate Substitution. *European Journal of Biochemistry* **1997,** *247* (2), 716-724.
21. Blaukopf, M.; Worrall, L.; Kosma, P.; Strynadka, N. C. J.; Withers, S. G., Insights into Heptosyltransferase I Catalysis and Inhibition through the Structure of Its Ternary Complex. *Structure* **2018,** *26* (10), 1399-1407 e5.
22. Cote, J. M.; Hecht, C. J. S.; Patel, K. R.; Ramirez-Mondragon, C. A.; Sham, Y. Y.; Taylor, E. A., Opposites Attract: Escherichia coli Heptosyltransferase I Conformational Changes Induced by Interactions between the Substrate and Positively Charged Residues. *Biochemistry* **2020,** *59* (34), 3135-3147.
23. Cote, J. M.; Ramirez-Mondragon, C. A.; Siegel, Z. S.; Czyzyk, D. J.; Gao, J.; Sham, Y. Y.; Mukerji, I.; Taylor, E. A., The Stories Tryptophans Tell: Exploring Protein Dynamics of Heptosyltransferase I from Escherichia coli. *Biochemistry* **2017,** *56* (6), 886-895.
24. Czyzyk, D. J.; Liu, C.; Taylor, E. A., Lipopolysaccharide Biosynthesis without the Lipids: Recognition Promiscuity of Escherichia coli Heptosyltransferase I. *Biochemistry* **2011,** *50* (49), 10570-10572.
25. Grizot, S.; Salem, M.; Vongsouthi, V.; Durand, L.; Moreau, F.; Dohi, H.; Vincent, S.; Escaich, S.; Ducruix, A., Structure of the Escherichia coli heptosyltransferase WaaC: binary complexes with ADP and ADP-2-deoxy-2-fluoro heptose. *J Mol Biol* **2006,** *363* (2), 383-94.
26. Gronow, S.; Brabetz, W.; Brade, H., Comparative functional characterization in vitro of heptosyltransferase I (WaaC) and II (WaaF) from Escherichia coli. *European Journal of Biochemistry* **2000,** *267* (22), 6602-6611.
27. Ramirez-Mondragon, C. A.; Nguyen, M. E.; Milicaj, J.; Hassan, B. A.; Tucci, F. J.; Muthyala, R.; Gao, J.; Taylor, E. A.; Sham, Y. Y., Conserved Conformational Hierarchy across Functionally Divergent Glycosyltransferases of the GT-B Structural Superfamily as Determined from Microsecond Molecular Dynamics. *International Journal of Molecular Sciences* **2021,** *22* (9).
28. Nkosana, N. K.; Czyzyk, D. J.; Siegel, Z. S.; Cote, J. M.; Taylor, E. A., Synthesis, kinetics and inhibition of Escherichia coli Heptosyltransferase I by monosaccharide analogues of Lipid A. *Bioorg Med Chem Lett* **2018,** *28* (4), 594-600.

29. Kadrmas, J. L.; Raetz, C. R., Enzymatic synthesis of lipopolysaccharide in Escherichia coli. Purification and properties of heptosyltransferase i. *J Biol Chem* **1998,** *273* (5), 2799-807.
30. Sirisena, D. M.; Brozek, K. A.; MacLachlan, P. R.; Sanderson, K. E.; Raetz, C. R., The rfaC gene of Salmonella typhimurium. Cloning, sequencing, and enzymatic function in heptose transfer to lipopolysaccharide. *Journal of Biological Chemistry* **1992,** *267* (26), 18874-18884.
31. Kadrmas, J. L.; Brozek, K. A.; Raetz, C. H., Lipopolysaccharide Core Glycosylation in Rhizobium leguminosarum: AN UNUSUAL MANNOSYL TRANSFERASE RESEMBLING THE HEPTOSYL TRANSFERASE I OF ESCHERICHIA COLI*. *Journal of Biological Chemistry* **1996,** *271* (50), 32119-32125.
32. Sirisena, D. M.; MacLachlan, P. R.; Liu, S. L.; Hessel, A.; Sanderson, K. E., Molecular analysis of the rfaD gene, for heptose synthesis, and the rfaF gene, for heptose transfer, in lipopolysaccharide synthesis in Salmonella typhimurium. *Journal of Bacteriology* **1994,** *176* (8), 2379-2385.
33. Oldfield Neil, J.; Moran Anthony, P.; Millar Lorna, A.; Prendergast Martina, M.; Ketley Julian, M., Characterization of the Campylobacter jejuni Heptosyltransferase II Gene, waaF, Provides Genetic Evidence that Extracellular Polysaccharide Is Lipid A Core Independent. *Journal of Bacteriology* **2002,** *184* (8), 2100-2107.
34. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N., ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* **2016,** *44* (W1), W344-W350.
35. Celniker, G.; Nimrod, G.; Ashkenazy, H.; Glaser, F.; Martz, E.; Mayrose, I.; Pupko, T.; Ben-Tal, N., ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Israel Journal of Chemistry* **2013,** *53* (3-4), 199-206.
36. Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research* **2010,** *38* (suppl_2), W529-W533.
37. Berezin, C.; Glaser, F.; Rosenberg, J.; Paz, I.; Pupko, T.; Fariselli, P.; Casadio, R.; Ben-Tal, N., ConSeq: the identification of functionally and structurally important residues in protein sequences. *Bioinformatics* **2004,** *20* (8), 1322-1324.
38. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **2011,** *7* (1), 539.
39. Robert, X.; Gouet, P., Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **2014,** *42* (W1), W320-W324.
40. Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A., A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics* **2004,** *55* (2), 351-367.
41. Bell, J.; Cao, Y.; Gunn, J.; Day, T.; Gallicchio, E.; Zhou, Z.; Levy, R.; Farid, R., PrimeX and the Schrödinger computational chemistry suite of programs. 2012; pp 534-538.
42. Hassan, B. A.; Milicaj, J.; Ramirez-Mondragon, C. A.; Sham, Y. Y.; Taylor, E. A., Ligand-Induced Conformational and Dynamical Changes in a GT-B Glycosyltransferase: Molecular Dynamics Simulations of Heptosyltransferase I Complexes. *Journal of Chemical Information and Modeling* **2022,** *62* (2), 324-339.
43. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra Precision Glide:□ Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *Journal of Medicinal Chemistry* **2006,** *49* (21), 6177-6196.
44. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T., End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews* **2019,** *119* (16), 9478-9508.

45. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation* **2011,** *7* (2), 525-537.
46. Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H., Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *Journal of Chemical Theory and Computation* **2011,** *7* (7), 2284-2295.
47. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006,** *65* (3), 712-25.
48. Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. v.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984,** *81* (8), 3684-3690.
49. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J., GROMACS: fast, flexible, and free. *J Comput Chem* **2005,** *26* (16), 1701-18.
50. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **2002,** *23* (16), 1623-1641.
51. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157-1174.
52. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.
53. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007,** *126* (1), 014101.
54. Parrinello, M.; Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981,** *52* (12), 7182-7190.
55. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995,** *103* (19), 8577-8593.
56. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997,** *18* (12), 1463-1472.
57. Cousins, K. R., ChemDraw Ultra 9.0. CambridgeSoft, 100 CambridgePark Drive, Cambridge, MA 02140. www. cambridgesoft.com. See Web site for pricing options. *Journal of the American Chemical Society* **2005,** *127* (11), 4115-4116.
58. Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H., Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2006,** *2*, 2006.0008-2006.0008.
59. Galanos, C.; Lüderitz, O.; Westphal, O., A New Method for the Extraction of R Lipopolysaccharides. *European Journal of Biochemistry* **1969,** *9* (2), 245-249.
60. Milicaj, J.; Castro, C. D.; Jaunbocus, N.; Taylor, E. A., Extraction of ADP-Heptose and Kdo2-Lipid A from E. coli Deficient in the Heptosyltransferase I Gene. *Applied Sciences* **2021,** *11* (18).
61. Morrison, J. P.; Tanner, M. E., A Two-Base Mechanism for Escherichia coli ADP-l-glycero-d-manno-Heptose 6-Epimerase. *Biochemistry* **2007,** *46* (12), 3916-3924.
62. Jarmoskaite, I.; AlSadhan, I.; Vaidyanathan, P. P.; Herschlag, D., How to measure and evaluate binding affinities. *eLife* **2020,** *9*, e57264.
63. Pollard, T. D., A guide to simple and informative binding assays. *Mol Biol Cell* **2010,** *21* (23), 4061-4067.
64. Goličnik, M., Explicit reformulations of time-dependent solution for a Michaelis–Menten enzyme reaction model. *Analytical Biochemistry* **2010,** *406* (1), 94-96.

65. Cote, J. M.; Taylor, E. A., The Glycosyltransferases of LPS Core: A Review of Four Heptosyltransferase Enzymes in Context. *International Journal of Molecular Sciences* **2017,** *18* (11).
66. Breton, C.; Šnajdrová, L.; Jeanneau, C.; Koča, J.; Imberty, A., Structures and mechanisms of glycosyltransferases. *Glycobiology* **2006,** *16* (2), 29R-37R.
67. Lucas, X.; Bauzá, A.; Frontera, A.; Quiñonero, D., A thorough anion–π interaction study in biomolecules: on the importance of cooperativity effects. *Chemical Science* **2016,** *7* (2), 1038-1050.
68. Belunis, C. J.; Raetz, C. R., Biosynthesis of endotoxins. Purification and catalytic properties of 3-deoxy-D-manno-octulosonic acid transferase from Escherichia coli. *Journal of Biological Chemistry* **1992,** *267* (14), 9988-9997.
69. Li, T.; Tikad, A.; Fu, H.; Milicaj, J.; Castro, C. D.; Lacritick, M.; Pan, W.; Taylor, E. A.; Vincent, S. P., A General Strategy to Synthesize ADP-7-Azido-heptose and ADP-Azido-mannoses and Their Heptosyltransferase Binding Properties. *Organic Letters* **2021,** *23* (5), 1638-1642.

Table 1. Conserved binding residues in HepI and equivalent residues in HepII. HepII residues equivalent to HepI binding residues based on multiple sequence alignment for **(A)** FDLA/FDHLA and **(B)** ADP-Hep/ADP. HepI residues are color coded by contact with substrate (red), product (cyan), or both (green).

A)

Residue Position HepI	% Conservation	Residue Position HepII	% Conservation
Lys7	84	Gly7	68
Thr8	75	Pro8	98
Ser9	95	Ser9	67
Ser10	99	Trp10	100
Asp13	100	Asp13	100
Arg60	98	Gly60	59
Arg61	90	His61	96
Arg63	93	X	
Lys64	82	X	
Leu96	96	Ser90	94
Lys98	96	Lys92	100
Arg120	80	Met112	37
Glu121	97	Asp124	56
Arg143	98	Arg146	23
Arg189	68	Gly192	99
Lys192	100	Lys195	100

B)

Asp261	100	Asp269	100
Residue Position HepI	**% Conservation**	**Residue Position HepII**	**% Conservation**
Ser10	99	Trp10	100
Met11	47	Val11	69
Thr187	88	Glu190	96
Thr188	48	Phe191	57
Arg189	68	Gly192	99
Lys192	100	Lys195	100
Gly218	92	Ser221	87
Glu222	98	His225	12
Arg225	78	Gly228	38
Lys241	44	Glu249	11
Met242	29	Thr250	98
Ser243	42	Gln251	8
Leu244	89	Leu252	96
Val247	38	Ala255	91
Asp261	100	Asp269	100
Thr262	88	Ser270	100
Gly263	100	Gly271	100
Leu264	92	Leu272	100
His266	100	His274	100
Ile287	31	Ile308	18

Table 2. Michaelis constant, turnover number, catalytic efficiency, and binding affinities of HepII substrates. Kinetic parameters determined for HepII for fully acylated (H-Kdo2-Lipid A) and partially deacylated (O-deacylated-H-Kdo2-Lipid A) donor and acceptors (ADP-Hep, ADP-Man) and binding affinities as determined via tryptophan fluorescence quenching.

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1}) x 10^6	K_D (μM)
HLA w/ Triton X 100	0.57 ± 0.06	0.49 ± 0.13	1.2 ± 0.4	
HLA	0.33 ± 0.03	0.41 ± 0.09	0.8 ± 0.2	
ODHLA	0.423 ± 0.007	0.58 ± 0.08	0.7 ± 0.1	0.48 ± 0.14
ADP-Hep	0.47 ± 0.02	0.57 ± 0.09	0.8 ± 0.2	0.66 ± 0.46
ADP-Man	0.54 ± 0.02	0.77 ± 0.08	0.7 ± 0.1	

Figure 1. Lipopolysaccharide components and inner core biosynthesis. (A) Regions of the full lipopolysaccharide (LPS) and their constituent saccharides. **(B)** Cartoon representation of the role of Heptosyltransferase I and II and their association with the inner membrane of the intracellular side to catalyze the transfer of heptose moiety onto the elongating glycopolymer (bottom left) prior to transportation (bottom/top right) and embedding of the LPS into the outer membrane (top left).

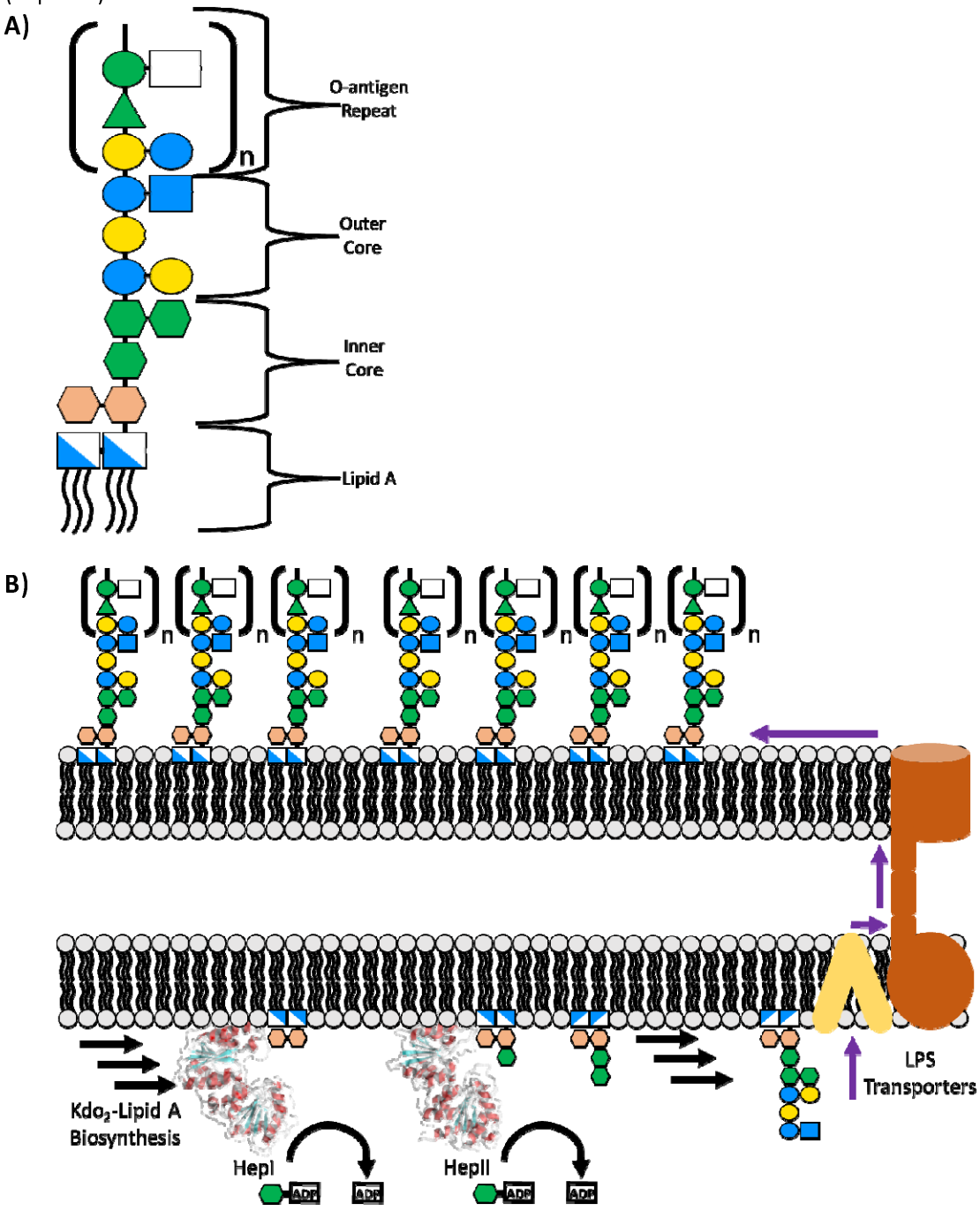
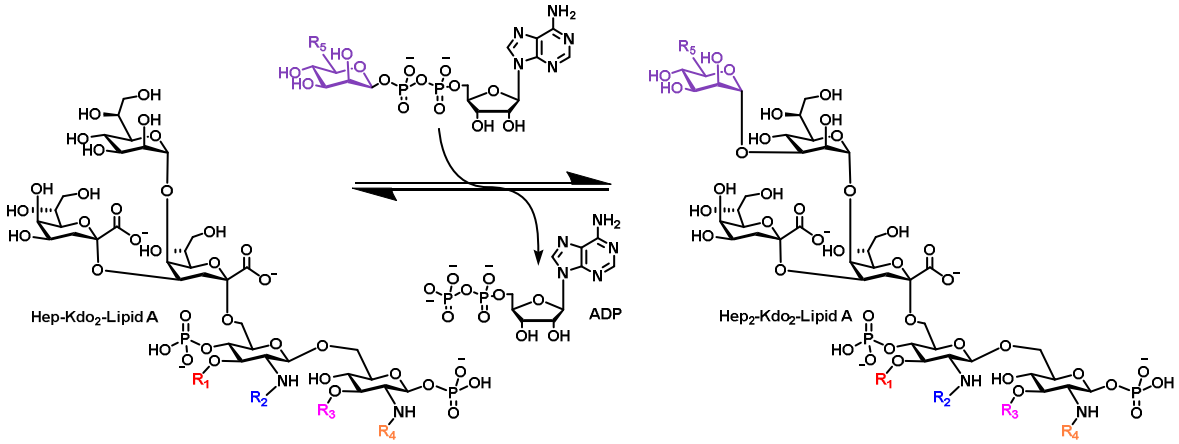


Figure 2. Reaction catalyzed by HepII with various substrates. (A) Reaction catalyzed by Heptosyltransferase II where a 6 or 7 carbon saccharide is transferred onto a Hep-Kdo_2-Lipid A with the release of ADP to form the product, Hep_2-Kdo_2-Lipid A. **(B)** The Hep-Kdo_2-Lipid A can be deacylated at the O positions (R1 and R3) to aid with solubility and both the 7 carbon native donor and a 6 carbon analogue were tested (R5).

A)



B)

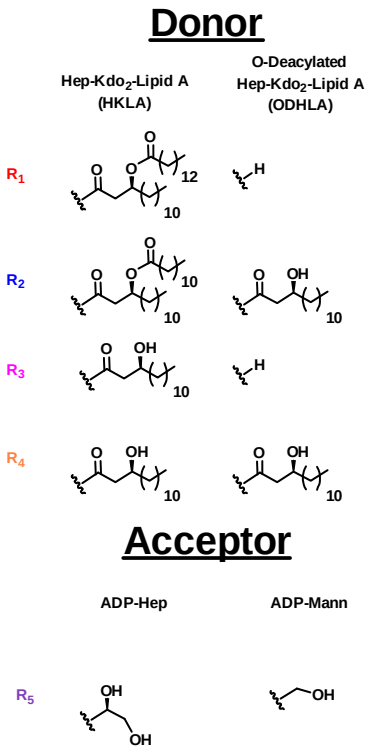


Figure 3. **Hybrid ternary complex model.** A ternary complex model of HepII was generated based on the crystal structure of the apo (PDB: 1PSW) with both ligands modeled in their respective domains (left). An electrostatic potential surface of HepII that demonstrates the positively charged surface of the active site (boxed lysine/arginine residues) to accommodate binding of the acceptor (circled phosphate groups of Hep-Kdo2-Lipid A) (right).

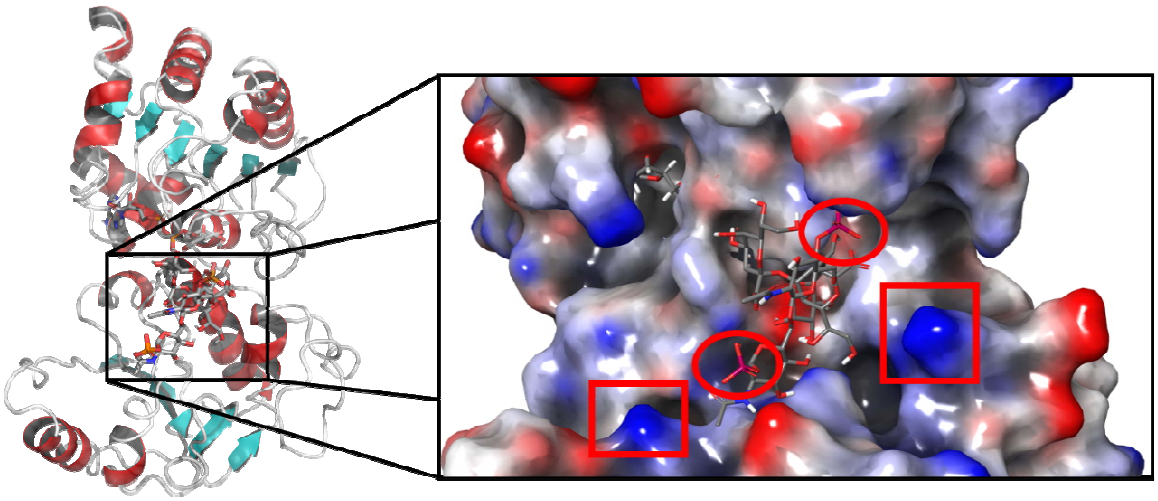


Figure 4. **Representative sequences of HepII multiple sequence alignment.** First 5 sequences of multiple sequence alignment for HepII with the query sequence (first sequence) from the crystalized E. coli HepII (PDB: 1PSW) and the secondary structural elements of the first sequence are represented immediately above the sequence. Residues in red boxes denote strictly conserved residues, and ones in bolded/boxed are greater than 70 similar.

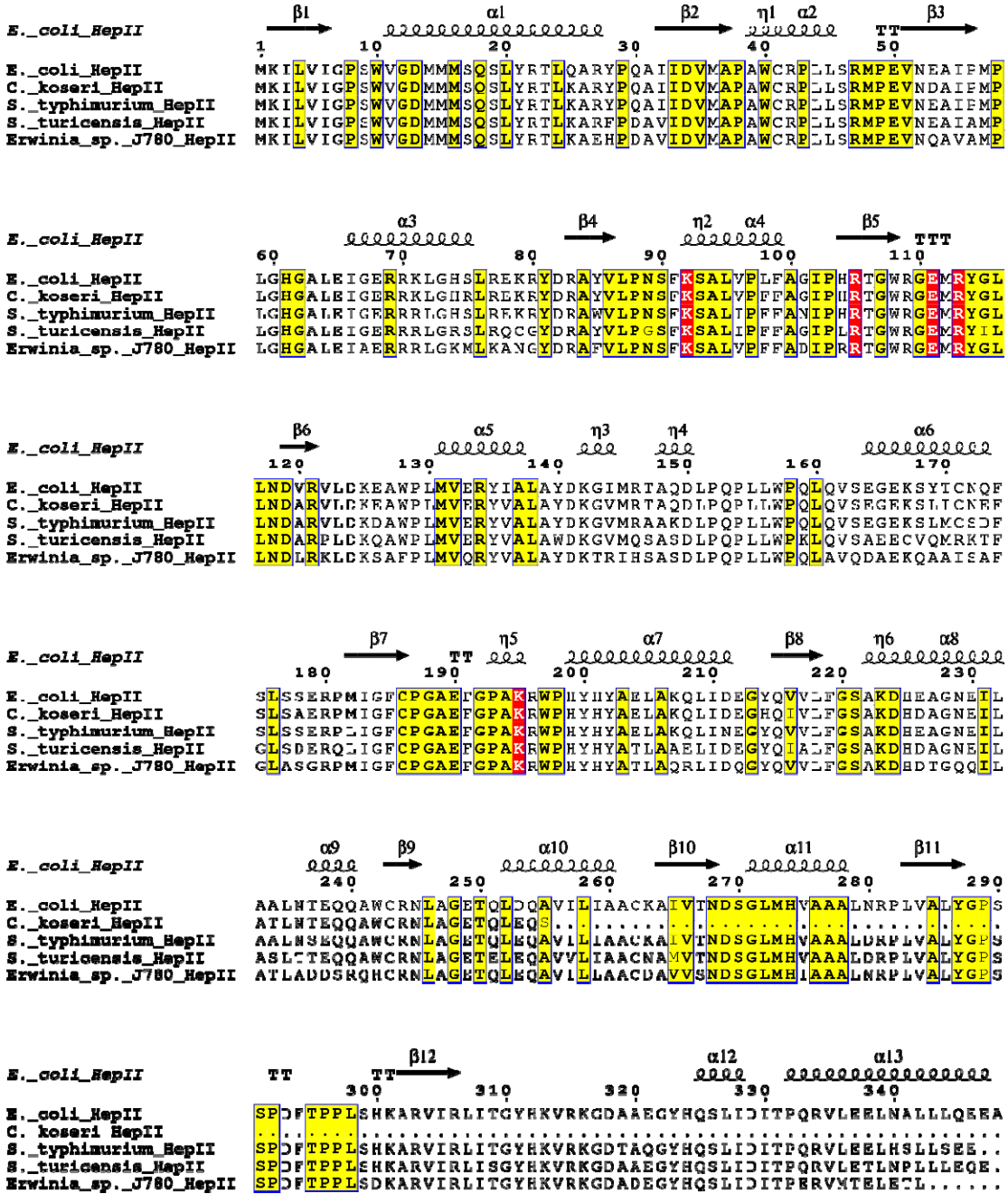
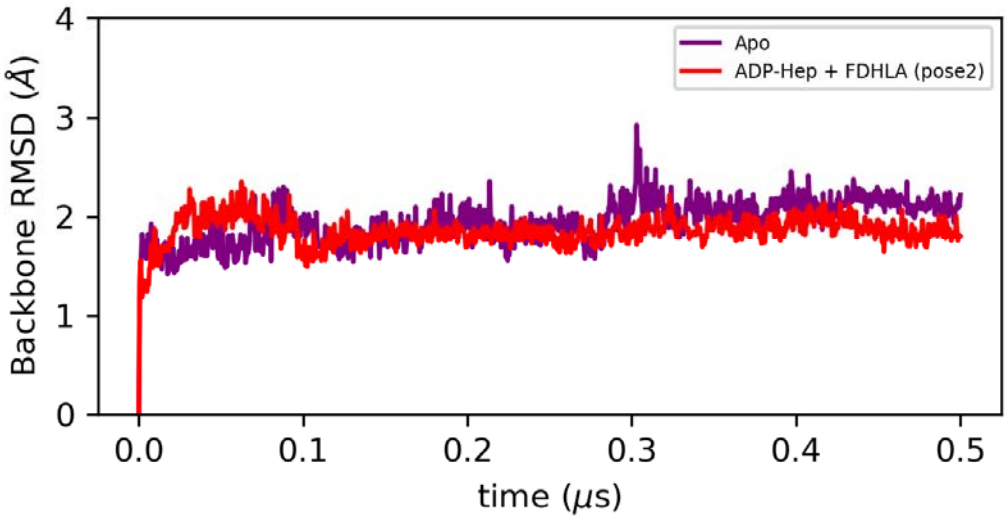


Figure 5. **RMSD and RMSF of HepII trajectory.** (A) Backbone RMSD and (B) Cα RMSF of HepII apo and pose2 of the HepII•ADP-Hep•FDHLA complex.

A)



B)

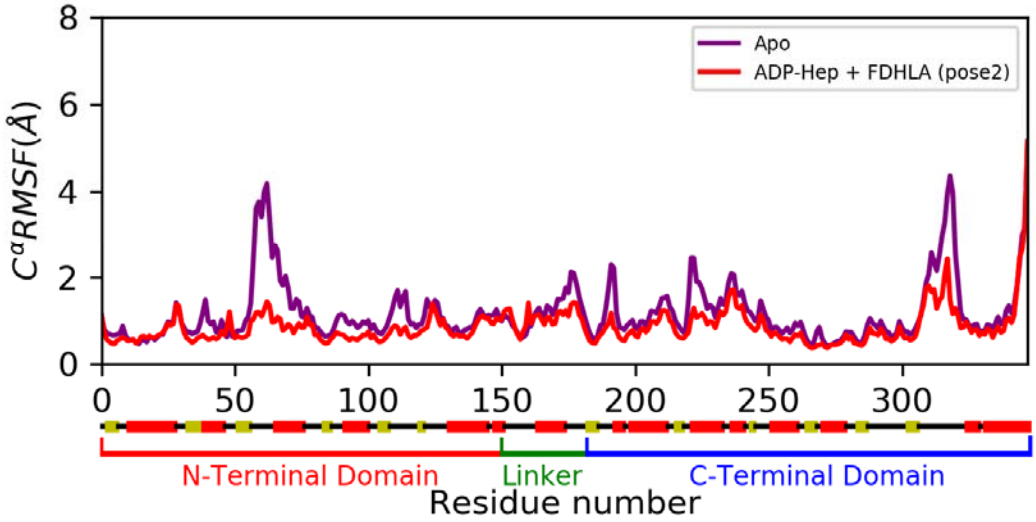
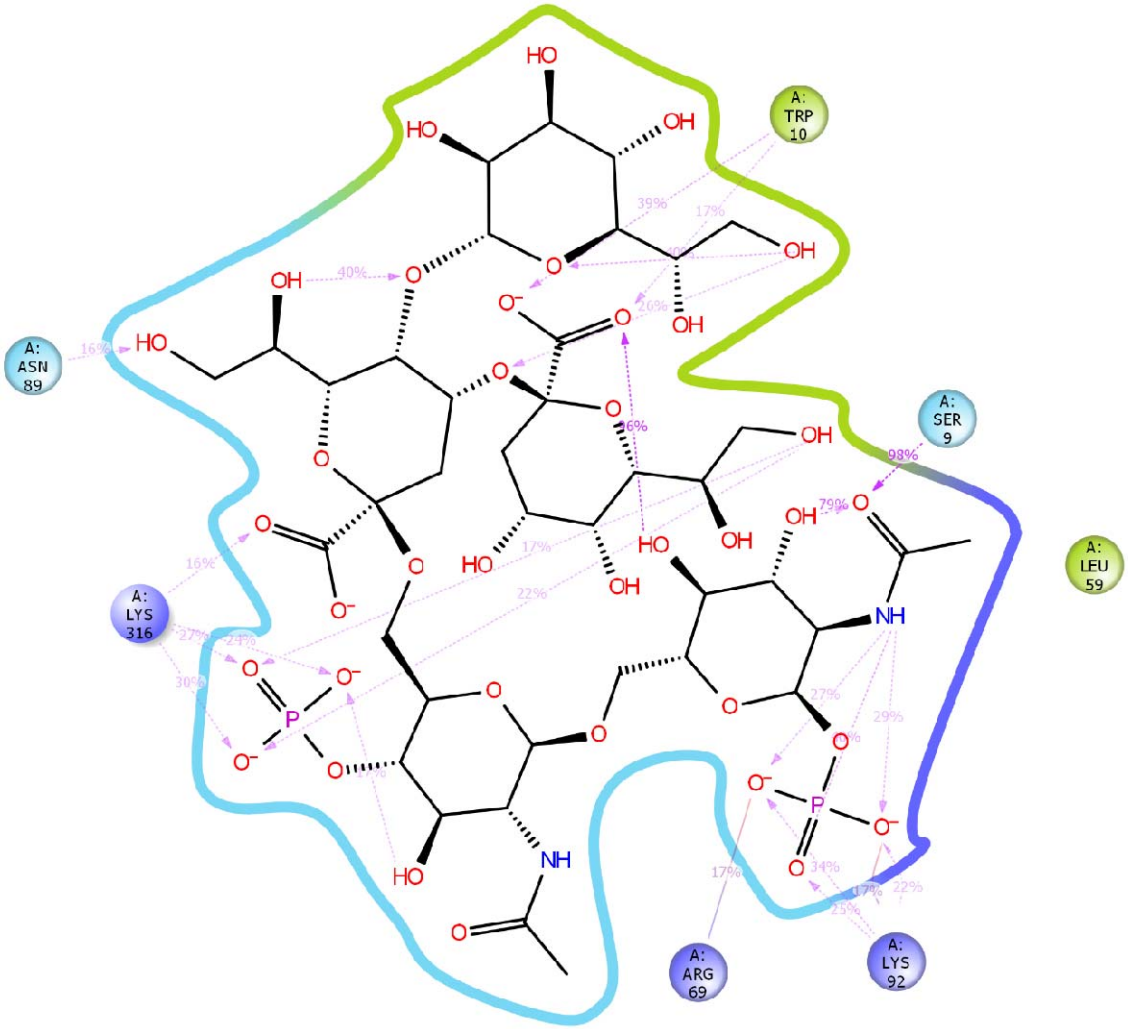


Figure 6. Ligand interaction diagram of substrates to HepII sidechains across MD trajectory.
(A) Contacts between acceptor substrate, FDHLA (pose2), and lysine/arginine of HepII that

anchor ligand in N-terminal side of active site interface. (B) Lysine and tryptophan sidechains that anchor donor, ADP-Hep, in C-terminal side of active site interface as observed from MD trajectory.

A)



B)

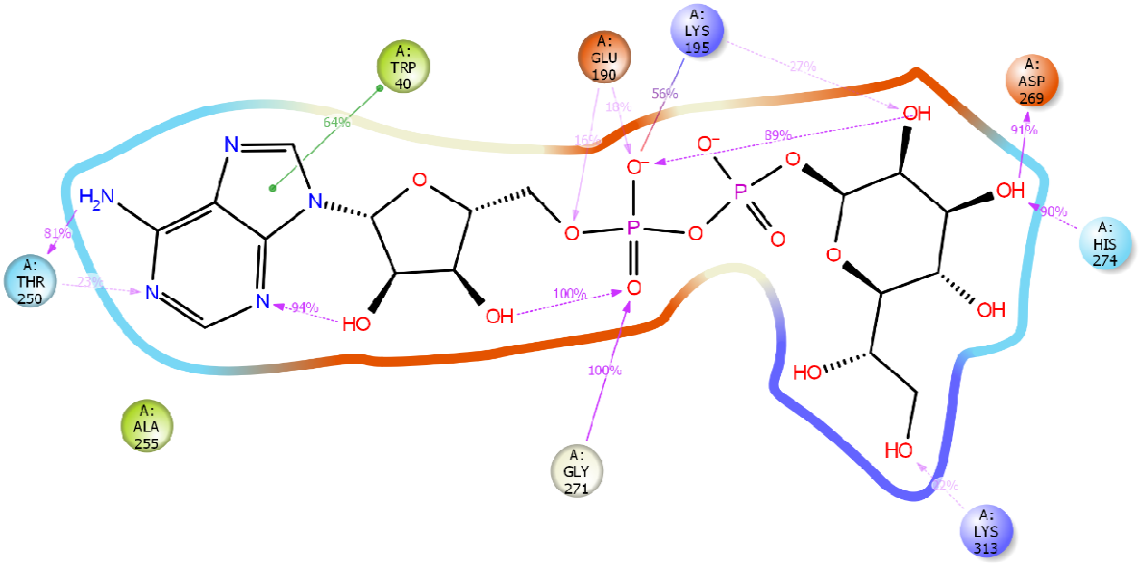
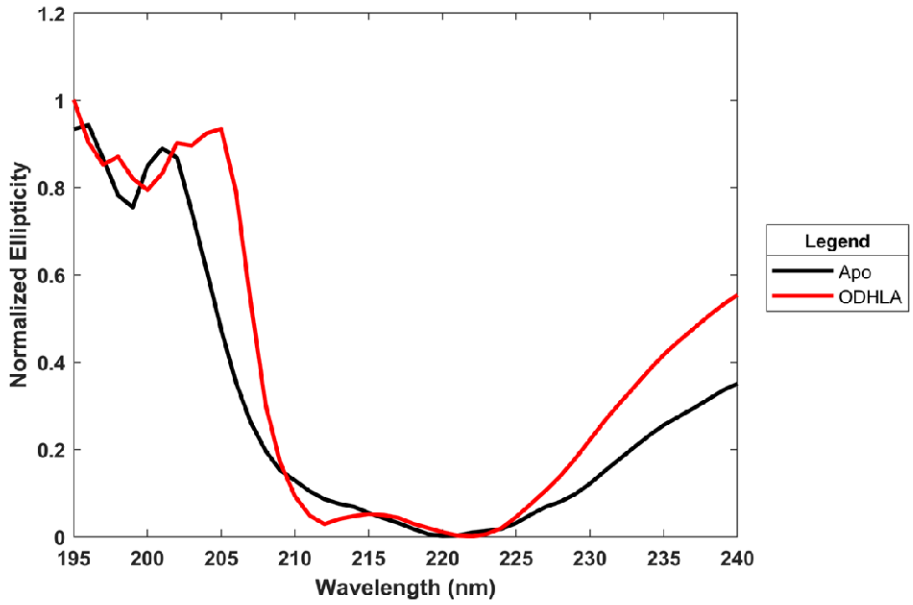


Figure 7. Circular dichroism spectroscopy of HepII with and without ligands and at various temperatures. (A) Secondary structure profile of HepII in the absence/presence (black/red) of ODHLA (acceptor) and demonstrates an increase in helicity due to a deepening of the second minima at 210 nm. **(B)** Unfolding of HepII as a function of temperature with substrates/products as determined by signal intensity at 220 nm demonstrates stabilization of HepII at high temperatures in the presence of the acceptor substrate/product (red/green) relative to apo (black).

A)



B)

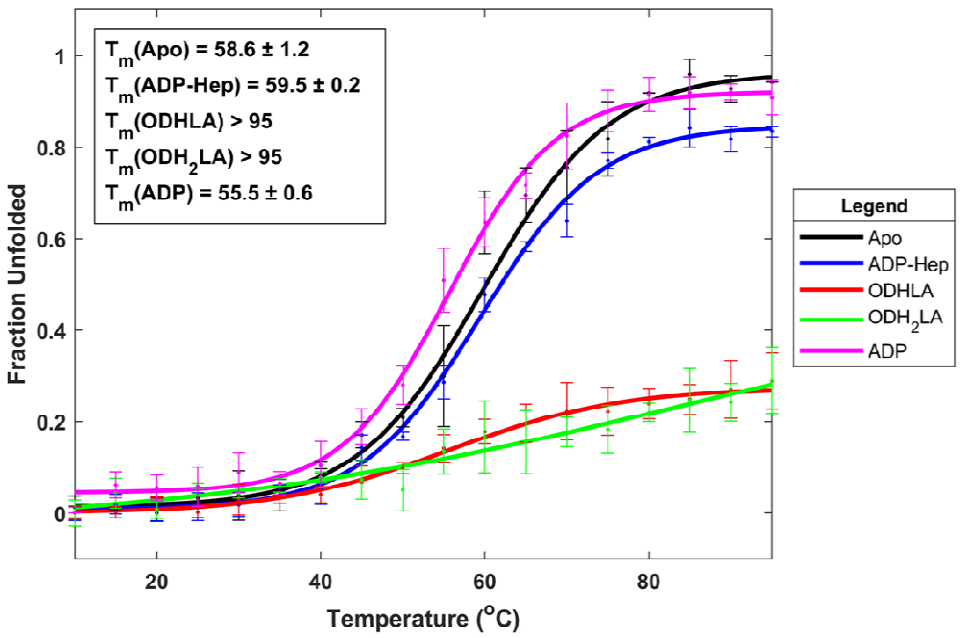
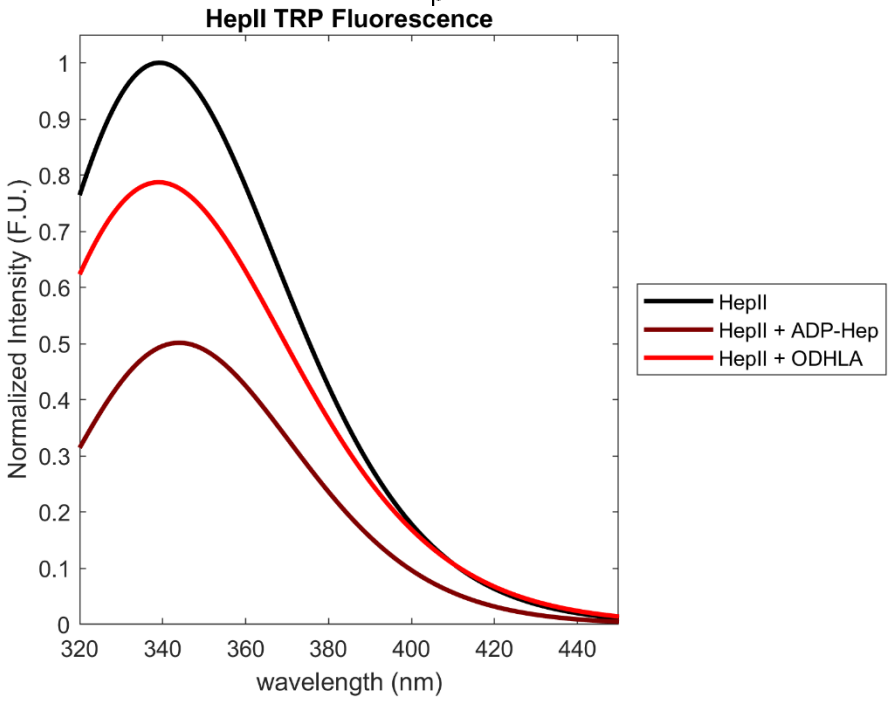


Figure 8. **Tryptophan fluorescence emission spectra and ligand binding curve. (A)** Overlay of HepII Tyrptophan fluorescence emission spectra of HepI apo (black), with ADP-Hep (brown), and ODHLA (red). **(B)** Binding curve of ADP-Hep as determined by quenching of tryptophan fluorescence as a function of APD-Hep concentration.

A)



B)