Ammonium Scanning in an Enzyme Active Site

THE CHIRAL SPECIFICITY OF ASPARTYL-tRNA SYNTHETASE

Received for publication, June 11, 2007, and in revised form, August 8, 2007 Published, JBC Papers in Press, August 9, 2007, DOI 10.1074/jbc.M704788200

Damien Thompson‡§1, Christine Lazennec1, Pierre Plateau1, and Thomas Simonson1,2

From the 1Laboratoire de Biochimie (CNRS, UMR7654), Department of Biology, Ecole Polytechnique, 91128 Palaiseau, France and the 2Tyndall National Institute, Lee Maltings, Prospect Row, Cork, Ireland

D-Amino acids are largely excluded from protein synthesis, yet they are of great interest in biotechnology. Aspartyl-tRNA synthetase (AspRS) can misacylate tRNAAsp with d-aspartate instead of its usual substrate, l-Asp. We investigate how the preference for l-Asp arises, using molecular dynamics simulations. Asp presents a special problem, having pseudosymmetry broken only by its ammonium group, and AspRS must protect not only against D-Asp, but against an “inverted” orientation where the two substrate carboxylates are swapped. We compare l-Asp and D-Asp, in both orientations, and succinate, where the ammonium group is removed and the ligand has an additional negative charge. All possible ammonium positions on the ligand are thus scanned, providing information on electrostatic interactions. As controls, we simulate a Q199E mutation, obtaining a reduction in binding free energy in agreement with experiment, and we simulate TyrRS, which can misacylate tRNA194 with D-Tyr. For both TyrRS and AspRS, we obtain a moderate binding free energy difference ΔG between the l- and d-amino acids, in agreement with their known ability to misacylate their tRNAs. In contrast, we predict that AspRS is strongly protected against inverted l-Asp binding. For succinate, kinetic measurements reveal a ΔG of over 5 kcal/mol, favoring l-Asp. The simulations show how chiral discriminations arises from the structures, with two AspRS conformations acting in different ways and proton uptake by nearby histidines playing a role. A complex network of charges protects AspRS against most binding errors, making the engineering of its specificity a difficult challenge.

The aminoacyl-tRNA synthetases (aaRSs)3 play a crucial role in preserving the accuracy of genetic code translation, linking each amino acid to a cognate tRNA, which carries the corresponding anticodon. Each of the 20 aaRSs must bind its target amino acid substrate with a high specificity (1–8). In the “class IIb” subgroup, formed of AspRS (9–16, 21–26), AsnRS (17, 24), and LysRS (18–20, 25), a network of electrostatic interactions ensures binding specificity for the cognate amino acid side chain in each case. Experimental and theoretical studies show, for example, that the preference of AspRS for Asp over its neutral analogue, Asn, involves a complex mechanism (24, 25). Proton binding, ATP binding, and long-range electrostatic interactions play important roles, and net positive electrostatic potential in the active site permits AspRS to bind l-Asp much more strongly than l-Asn.

The translation apparatus must also preserve the homochirality of proteins, and aaRSs should be specific for l-amino acids. Kinetic experiments show, however, that AspRS does not strongly distinguish between the “left-handed” l-Asp and “right-handed” d-Asp stereoisomers. Indeed, d-Asp-tRNAAsp is produced at detectable levels by Escherichia coli AspRS (26), at a rate only 4,000 times lower than l-Asp-tRNAAsp. To preserve homochirality in vivo (27), editing (28) of d-Asp-tRNAAsp by a d-aminoacyl-tRNA deacylase is performed (26). For biotechnology applications, it is not always desirable to preserve homochirality, and d-amino acids are of great potential interest in protein engineering and design. For example, mixed l- and d-amino acid proteins are used in the development of synthetic vaccines (29, 30). It is hence of interest, both from a fundamental and an engineering perspective, to better understand how aaRSs distinguish between l- and d-amino acids.

In the present work, we use simulations and experiments to investigate the origins of chiral specificity of E. coli AspRS for its l-aspartate substrate (l-Asp). Asp has a pseudosymmetry, broken only by its ammonium group, and so the enzyme must protect not only against d-Asp, but against an alternate, “inverted” orientation where the two substrate carboxylates are swapped (Fig. 1A). The inverted orientation is seen in the crystal structure of a homologous enzyme, asparagine synthetase (31). We compare both l-Asp and D-Asp, in regular and inverted orientations, and the metabolite succinate, where the ammonium group is removed altogether and the ligand has an additional negative charge. This is done with a state of the art molecular dynamics free energy (MDFE) technique (21–25, 32–39), by computing the change in binding free energy when an ammonium group is inserted at either possible position on either succinate methylene group.

By scanning all possible ammonium positions on the ligand, we can also address an interesting, secondary question: the strength of electrostatic interactions in the active site. Simulations have played a significant role in establishing the impor-
Chiral Specificity

FIGURE 1. A, schematic of the AspRS active site with bound Asp. Dashed lines on Asp indicate alternative NH$_3^+$ positions, indicated as follows: L, L-Asp; D, D-Asp; invL, inverted L-Asp; and invD, inverted D-Asp. All these ligands and orientations differ only by the placement of the backbone ammonium group (or its removal, in the case of succinate). Motif 2 designates a conserved sequence motif, characteristic of class II aaRSs. B, the active site with an open or closed flipping loop. Overlay of PDB crystal structures 1IL2 (15) (closed loop) and 1COA (13) (open loop). An AspAMP ligand is shown in ball-and-stick representation; the Arg$^{498}$, Lys$^{198}$, and Arg$^{217}$ side chains are shown as sticks; the flipping loop is shown as a tube, with the Glu$^{171}$ side chain shown as sticks. PDB code 1IL2 is colored black and PDB code 1COA is colored gray.

The MD simulations technique has been extensively validated in recent years (see for example, our recent work in Refs. 24 and 44–46; see Refs. 39 and 47 for reviews). We test it further by comparing our results to published measurements of l-/d-amino acid tRNA$^{Asp}_{tRNA}$ acylation rates (26). The simulations predict that D-Asp does not bind much less strongly than L-Asp, consistent with the observed enzymatic activity toward D-Asp. We also perform two sets of control simulations. We simulate the Q199E mutation of AspRS, obtaining a reduction in l-Asp binding free energy in good agreement with published experiments (10). We also compare l-Tyr and d-Tyr binding to E. coli TyrRS, which is known to have d-Tyr activity (26). The computed l-Tyr/d-Tyr binding free energy difference $\Delta \Delta G$ is also in good agreement with experimental estimates (26). Finally, we perform additional experiments in this work, measuring the kinetics of ATP/pyrophosphate exchange in the presence of L-Asp, catalyzed by AspRS and inhibited by succinate. This should reveal the binding free energy difference between l-Asp and succinate (24, 26). Succinate binding is weak, so that only a lower bound is obtained in practice, $\Delta \Delta G > 5$ kcal/mol favoring l-Asp. The MD simulations are consistent with this bound.

Thus validated, the simulations provide a detailed model for the sources of chiral specificity in AspRS. We mention four of the salient features here. (i) Inverted d-Asp binding (Fig. 1A) gives a free energy just 4.7 kcal/mol higher than l-Asp. This moderate value appears sensible: if we assume that the lower rate of tRNA charging with D-Asp is entirely due to weaker D-Asp binding, then the experimental rate reduction of 4000 corresponds to a binding free energy difference of 4.9 kcal/mol, very close to the MD estimate. Thus, discrimination against D-Asp can be explained by unfavorable binding. (ii) Inverted D-Asp binding is slightly better than D-Asp binding in the regular orientation. This raises the intriguing possibility that D-Asp may acylate tRNA with its side chain rather than its backbone carboxylate (at least partly), a question that is not experimentally resolved. (iii) Inverted l-Asp binding leads to poorer interactions and a binding free energy well below that of l-Asp in the regular, experimental orientation. Succinate is the most complicated case. Its binding depends on the conformation of a flexible, “flipping loop” next to the active site (Fig. 1B). This loop closes over the cognate l-Asp ligand when it binds, switching from an “open” to a “closed” conformation (11, 12). Proton uptake by a nearby histidine assists the formation of a closed loop enzyme-succinate complex. Taking the experiments and simulations together, we infer (iv) that the closed loop is strongly preferred by l-Asp, and this leads to a much weaker binding of succinate, which requires the open loop. Overall, the chiral specificity of AspRS arises from a complex network of charges, whose collective behavior leads to effective protection against most binding errors.

EXPERIMENTAL PROCEDURES

Succinate Force Field Parameters—Succinate was taken from the PDB (entry 1HQJ, a 1.2-Å resolution crystal structure (48)), and its electronic structure computed, using Gaussian03 (49) with the B3LYP (Becke three-parameter hybrid functional combined with Lee-Yang-Parr correlation functional), hybrid Hartree-Fock density functional theory (50), and the 6-311++G** basis set. Charges were computed by a Natural Population Analysis (51). The same procedure was applied to l-Asp, as a consistency check, giving charges in good agreement with the well established, CHARMM22 force field (52) used below. Other parameters for succinate were taken from the CHARMM22 l-Asp model. For more details, see supplementary materials.

MD Simulations—We showed previously (24, 25) that co-binding of the second substrate, ATP with three divalent cations (12), assists amino acid discrimination in AspRS. Hence,
ATP:Mg\(^{2+}\) is always present in the current calculations, and “AspRS-ligand” is used as a shorthand for ATP-Mg\(^{2+}\)-AspRS-ligand complexes. In line with the previous simulations, we model the AspRS binding pocket in its closed state (12, 15) unless otherwise mentioned. Closing the flipping loop (Fig. 1B; also shown in green in the panels of Fig. 2) is thought to aid amino acid binding by bringing a negative glutamate (Glu\(^{171}\), Fig. 1B) into the coordination sphere of the backbone ammoxonium. Succinate lacks the ammoxonium group, and so we also did AspRS-succinate simulations with a more open binding pocket geometry (11, 14, 21–23).

The protonation states of several active site histidines play a role in AspRS specificity. The simulations correspond to pH 7. Poisson-Boltzmann/Linear Response Approximation calculations compare the free energy to protonate each histidine in the active site and a side chain analogue, methylimidazole, in solution (37, 53) and have been shown to provide good agreement with measured ionization states, where available (53). The corresponding double free energy difference \(\Delta G\) is related to the \(pK_a\) values of the two groups: \(pK_{a,\text{prot}} = pK_{a,\text{model}} + (1/2.303 \ RT) \Delta G\). From the known \(pK_{a,\text{model}}\) of 6.5 we can then deduce \(pK_{a,\text{prot}}\). We showed previously (24) that at pH 7, the closed loop active site has a doubly protonated His\(^{448}\) when L-Asp is bound, and kinetic measurements of Asp/Asn competitive binding to AspRS as a function of ATP concentration supported the doubly protonated His\(^{448}\) state. We use this protonation state below, with the following exception. His\(^{448}\) remains uncharged in the open active site when either L-Asp (in agreement with previous simulations (21–23)) or succinate are bound. With the closed loop, an adjacent histidine, His\(^{445}\), is uncharged with L-Asp, but charged with succinate (Table 1).

We performed 1 ns of free dynamics for each system, following thermalization and equilibration. The constant PT ensemble and the Particle Mesh Ewald cell used have been described in detail elsewhere (24, 25). Briefly, we consider a 24-Å sphere of protein centered on the ligand, immersed in a 73-Å water box with four sodium counterions. We also performed 500 ps of dynamics for D-Asp and succinate solvated in large water boxes; these, together with the solvated L-Asp system produced earlier (24), provide reference systems for the alchemical free energy simulations. We used the CHARMM program (54) version c30b1 for all calculations. The sensitivity of the calculations to various model details has been probed in a series of previous studies; for example, model radii of 20, 24, and 28 Å have been compared (21–25).

| Loop conformation | Ligand  | \(\Delta G_{\text{prot}}\) | \(\Delta G_{\text{prot}}\) | \(\Delta G_{\text{model}}\) | \(pK_a\) shift | His\(^{448}\) state |
|-------------------|--------|----------------|----------------|----------------|----------------|----------------|
| Open              | L-Asp  | −10.6          | +11.1          | +21.7          | +16.0          | Neutral        |
| Open              | Succinate | −10.6          | +8.4           | +2.2           | +1.6           | Neutral        |
| Closed            | L-Asp  | −10.6          | +1.7           | +12.3          | +8.9           | Neutral        |
| Closed            | Succinate | −10.6          | −20.9          | −10.3          | −7.5           | Charged        |

For the L-Tyr versus D-Tyr binding simulations in \textit{E. coli}\, TyrRS (55), we used a less expensive, spherical, continuum reaction field model, as described in previous studies (21–23, 25). Briefly, it included a 24-Å sphere of protein centered on the ligand, along with the water molecules inside the 24-Å sphere (about 150 crystal waters and 750 additional waters). Water and protein outside the 24-Å sphere were treated as a single, homogeneous dielectric medium with a dielectric constant of 80 (56, 57). Electrostatic interactions between atoms within the sphere were computed without any cutoff, using an efficient multipole approximation for distant groups (58). A multipolar expansion with 20 terms was used to approximate the reaction field due to the surrounding continuum (56, 57). Newtonian dynamics were used for the inner 20 Å of the sphere, and Langevin dynamics for the outer region (20–24 Å), with a bath temperature of 293 K. We sampled both L-Tyr-TyrRS and D-Tyr-TyrRS complexes for 5 ns, following 300 ps of thermalization and equilibration. We also performed 1 ns of dynamics for solvated L-Tyr and D-Tyr, providing reference systems for the alchemical free energy simulations.

Free Energy Calculations—Alchemical free energy calculations (32, 33) (MDFE) and Poisson-Boltzmann free energy calculations (23) (PBFE) were done to compute binding free energy differences between ligands, with the methodology described earlier (21–23, 38, 39). Briefly, the MDFE method transforms one ligand (e.g. L-Asp) into another (e.g. D-Asp) by gradually deleting an ammonium group in one position and introducing it in another. To do this, the terms in the energy function associated with the two ammonium groups are scaled, respectively, by 1-\(\lambda\) and \(\lambda\), where \(\lambda\) is a “coupling parameter” that varies gradually from zero to one. The free energy derivative with respect to \(\lambda\) has the form \(dG/d\lambda = \langle dU/d\lambda \rangle\), where \(U\) is the energy function and the brackets indicate an average over an MD simulation performed with a particular value of \(\lambda\). The derivative is integrated numerically to obtain \(\Delta G\). The procedure is applied to the protein-ligand complex and to the ligand alone in solution; subtracting the two free energy changes gives \(\Delta G\), the binding free energy difference between the two ligands (L-Asp and D-Asp). With this procedure, the contribution \(\Delta G_i\) of an individual amino acid \(i\) to \(\Delta G\) can be defined and computed by isolating the contribution of \(dU_i/d\lambda\) to \(dG/d\lambda\), where \(U_i\) represents all the energy terms involving amino acid \(i\). \(\Delta G_i\) is known as a free energy component (38, 39). For AspRS, we performed two MDFE runs in each direction for each pair of native/competitor ligand systems, initiating runs from starting points 0.3 and 0.5 ns on the MD trajectories. For TyrRS, we performed 10 MDFE runs in each direction, taking a starting

\(^{4}\text{D. Thompson, C. Lazennec, P. Plateau, and T. Simonson, unpublished data.}\)
TABLE 2
MDFE alchemical free energy changes in solution and in AspRS for L-Asp versus competitor ligand binding

| Competitor ligand | ΔG in solution | ΔG in AspRS | ΔΔG | ΔΔGelectrostatic | ΔΔGvand der Waals |
|-------------------|----------------|-------------|------|------------------|------------------|
| L-Asp inverted    | 0.0/0.0        | 7.2, 16.5/11.1, 6.9 | 10.4 | 7.2              | 3.2              |
| D-Asp inverted    | 0.0/0.0        | 8.7, 7.9/3.5, 3.3  | 5.9  | 5.1              | 0.8              |
| Succinate (open loop) | −40.1/−41.7  | −42.3, −40.0/−46.7, −44.4 | −2.4 | −3.9             | 1.5              |
| Succinate (closed loop) | −40.1/−41.7  | −37.4, −38.6/−44.9, −41.2 | 0.4 (+12.7) | −1.8            | 2.2              |

point every 0.5 ns along the 5-ns L-Tyr·TyrRS and D-Tyr·TyrRS MD trajectories.

Additional MDFE simulations were used to compute the change in Asp binding upon introducing the Q199E mutation. As shown in the supplementary materials (supplemental materials Table S1), the Q199E mutation induces His512 protonation, so the Glu499 transformation was performed simultaneously with His512 charging. We used four 5-ns simulations to perform the transformation using the thermodynamic cycle shown in supplemental materials Fig. SM3 and the charge perturbation scheme shown in supplemental materials Fig. SM2.

The PBFE method computes the electrostatic contribution to the ligand binding free energy by solving the Poisson-Boltzmann equation numerically to obtain the electrostatic free energy. The free energy is computed for the protein-ligand complex and for the separate protein and ligand. The solvent is treated as a homogeneous dielectric medium with the dielectric constant of bulk water, 80. The protein and ligand are treated as another dielectric medium, with a lower dielectric constant ε, which is empirically adjusted. Calculations are performed for an ensemble of ~250 structures drawn from the MD simulations of the protein-ligand complex (above); the computed binding free energies are averaged over the ensemble. The structures of the separate protein and ligand are obtained by simply discarding the unwanted partner.

Experimental Measurement of Aspartyl-Adenylate Formation and Its Inhibition by Succinate—E. coli AspRS was purified as described previously (26). The purified enzyme was stored at −20 °C in a 20 mM Tris-HCl buffer (pH 7.8) containing EDTA (0.1 mM) and glycerol (60%). Prior to activity measurements, AspRS was diluted in a 50 mM Tris-HCl buffer (pH 7.8) containing EDTA (0.1 mM) and bovine serum albumin (200 mg/ml; from Roche Applied Science). Initial rates of ATP·PPi exchange activity were measured for 10 min at 25 °C. The 75-μl reaction mixture contained Tris·HCl (50 mM; pH 7.5), MgCl2 (7 mM), EDTA (100 μM), [32P]PPi (100 nCi, 2 mM), ATP (2 mM), and various concentrations of L-Asp and sodium succinate. The exchange reaction was started by adding catalytic amounts of AspRS (typically, 300 nM). After quenching the reaction, 32P-labeled ATP was adsorbed on charcoal, filtered, and measured by scintillation counting.

RESULTS

We first describe the simulations comparing l-Asp binding in the regular and inverted orientations. We then describe the simulations comparing l-Asp to d-Asp binding, with either d-Asp orientation. Next, we consider succinate binding: simulations are described first, with the active site flipping loop either open or closed; measurements of pyrophosphate exchange are described second; third, we use thermodynamic arguments to infer the most favorable conformation of the flipping loop. The control simulations are described last. Although discrimination against d-Asp, not succinate, is the main question of interest here, it is worth analyzing the complicated succinate case in detail, for three reasons: discrimination against succinate is biochemically important; succinate is the closest Asp analogue with two negative charges; and the succinate data give indirect evidence on the most favored flipping loop conformation when l-Asp is bound.

l-Asp Binding to AspRS: Regular Versus Inverted Orientation

Table 2 shows computed, MDFE values for the binding free energy difference ΔΔG between regular L-Asp and inverted l-Asp binding to AspRS. The uncertainty, estimated by considering pairs of forward/backward runs (runs that mutate regular L-Asp into inverted L-Asp or the reverse), is about 2 kcal/mol, similar to our previous MDFE studies of AspRS (21–25). The calculated binding specificity ΔΔG of +10.4 kcal/mol is very large, and more than sufficient to prevent binding of l-Asp in an inverted geometry.

Structurally, the inverted orientation for L-Asp gives a less ordered binding pocket. Protein atoms in the pocket have side chain r.m.s. deviations (fluctuations) of 0.9 (0.3) and 1.6 (0.4) Å, for regular L-Asp (24) and inverted l-Asp, respectively, whereas the ligands themselves have r.m.s. deviations (fluctuations) of 0.8 (0.3) and 2.4 (0.5) Å. Deviations are computed from the crystal structure; fluctuations are relative to the mean MD structure. A representative MD snapshot for inverted l-Asp is illustrated in panel C of Fig. 2 (created using Molscript (59) and rendered using Raster3D (60)). The flipping loop is less tightly closed with inverted L-Asp than with regular L-Asp (24), with Glu173 forming a weaker interaction with the (displaced) ligand ammonium. In a secondary effect, twisting of the ligand also weakens the terminal carboxylate interaction with the Glu231 side chain amide; distances shift from under 3 Å with regular L-Asp (24) to ~6 Å with inverted L-Asp.

With inverted L-Asp, the ligand ammonium group is displaced (Fig. 1A); if the rest of the structure stayed fixed, it would sterically overlap with Ala322 (Fig. 3A). Thus, one might have anticipated that binding would be penalized by unfavorable van der Waals interactions. In fact, to avoid steric overlap, both the protein and ligand shift during the simulations (Fig. 3, B and C), and the binding free energy penalty ΔΔG arises mainly from a
poorer electrostatic fit for inverted L-Asp, rather than a poorer steric fit.

Indeed, a free energy component analysis was performed, decomposing the overall $\Delta G$ into individual ligand/host electrostatic and van der Waals interactions (Table 3). Glu$^{171}$ has the highest electrostatic free energy component (21, 23, 24, 61) of the residues shown in Fig. 2C: $+16.0$ kcal/mol in favor of the regular binding orientation. Glu$^{171}$ interacts less strongly with the inverted ligand and responds by coordinating Arg$^{217}$. Arg$^{217}$ favors inverted L-Asp (by $-6.9$ kcal/mol), due to the enhanced Arg$^{217}$–Glu$^{171}$ interaction. The only other residues contributing strongly to regular/inverted specificity are the salt-bridge pair, Lys$^{198}$ and Asp$^{233}$, contributing $+10.6$ and $-10.6$ kcal/mol, respectively, and canceling each other. The inverted L-Asp side chain carboxylate binds less tightly to Lys$^{198}$, with Asp$^{233}$ increasing its coordination to Lys$^{198}$ as a result. All these residues are largely conserved in AspRSs, suggesting that the effects described are rather general. The main exception is Glu$^{171}$, which is absent in some eukaryotic AspRSs (including human and Caenorhabditis elegans).

On balance, Coulomb interactions with the protein provide $+13.1$ kcal/mol in favor of the experimental L-Asp geometry, offset by $-6.2$ kcal/mol from water. van der Waals interactions provide an additional $+3.2$ kcal/mol, giving a total $\Delta G$ of $+10.4$ kcal/mol in favor of the regular L-Asp binding geometry. Thus, less favorable van der Waals interactions for inverted L-Asp contribute 31% of the overall effect, whereas Coulomb interactions contribute 69%. Co-binding of ATP and its three divalent cations does not affect binding specificity, contributing only $0.2$ kcal/mol to $\Delta G$. In summary, the weaker coordination of the displaced ligand ammonium to the flipping loop Glu$^{171}$ is the main factor preventing L-Asp from binding in an inverted geometry. Some eukaryotic AspRSs that lack Glu$^{171}$ may thus be less protected against inverted L-Asp binding.

L-Asp Versus D-Asp Binding

We compared L-Asp binding to both regular D-Asp and inverted D-Asp, completing our scan of possible ligand ammonium locations. Discrimination against D-Asp is only moderate, especially for inverted D-Asp: $\Delta G = +4.7$ kcal/mol (Table 2). This agrees with kinetic experiments (26) showing that AspRS can misacylate tRNA$^{Asp}$ with D-Asp. For D-Asp in the regular orientation, we find a slightly larger $\Delta G$ of $+5.9$ kcal/mol. Thus, the binding pocket is less protected against D-Asp than against inverted L-Asp. Displacing the ligand ammonium to produce D-Asp (in either orientation), keeping the rest of the structure fixed, does not lead to a large steric overlap with protein, in contrast to inverted L-Asp, above. The binding pocket undergoes little disruption when D-Asp binds instead of L-Asp (Fig. 2); protein and ligand r.m.s. deviations (and fluctuations) change by less than ±0.2 Å. The inverted D-Asp orientation gives a slightly less-
ordered pocket, with protein side chain and ligand r.m.s. deviations (fluctuations) rising to 1.3 (0.4) and 1.6 (0.4) Å. The flipping loop remains closed with D-Asp bound in both orientations. Compared with L-Asp (12, 15, 24), D-Asp loses its NH₃⁺ interaction with the buried water molecule W (Fig. 2B) and has a weaker interaction with Arg²₁⁷, whereas inverted D-Asp loses both the buried water molecule and Gln¹₉₈ interactions (Fig. 2A).

The moderate L-Asp/D-Asp binding specificity is reflected in the low individual free energy components in Table 3. Flipping D-Asp into the inverted orientation reverses the MDFE L-Asp/D-Asp component of most nearby residues (Table 3). A difficulty in interpreting the MDFE components is the protein/solvent compensation, so it is instructive to look at the free energy components obtained with Poisson-Boltzmann, or “PBFE” calculations (23), with the protein and solvent treated as distinct dielectric media (Table 4). The protein/ligand dielectric constant ε was set to 4 by calibrating the overall PBFE ΔΔG with respect to the MDFE results in this and earlier studies (21–25). Similar ε values have been used for other proteins (62, 63).

The PBFE free energy components (Table 4) show that only His⁴⁴⁸ and, to a lesser extent, Arg⁴⁸⁹ express a significant L-Asp preference (with components of 1–5 kcal/mol). Arg⁴⁸⁹ is completely conserved in AspRSs, whereas His⁴⁴⁸ is changed to a neutral amino acid in Archaea. The conserved Lys¹₉₈:Asp²₃₃ pair contributes almost zero. Fig. 2 shows the close proximity of the D-Asp ammonium to the His⁴⁴₈ proton (distances of 4.6 Å with L-Asp bound, versus 3.7 and 3.4 Å with regular and inverted D-Asp) and a slight shift of Arg⁴₈₉ away from the ligand (2.0 and 1.8 Å with regular and inverted D-Asp). Distances are averages over 250 MD structures, equally spaced along the trajectories. The rest of the binding pocket does not contribute strongly to the chiral discrimination. van der Waals interactions.

### TABLE 3

| Competitor ligand | Water | Protein | ATP:Ms²₅⁴⁺ | Glu²⁸₁ | Lys²⁸⁸ | Arg²¹⁷ | Asp²³³ | Glu²⁵⁵ | His⁴⁴⁸ | Arg⁴⁸⁹ | Asp⁴⁷⁵ | Glu⁴⁹⁴ |
|------------------|-------|---------|------------|--------|--------|-------|-------|-------|-------|-------|-------|-------|
| L-Asp inverted   | -6.2  | +13.1   | -0.2       | +16.0  | +10.6  | -6.9  | -10.6 | -3.4  | -1.9  | -0.4  | -0.4  | -0.9  |
| D-Asp            | -1.1  | +1.3    | +2.8       | -7.3   | -10.8  | +0.4  | +8.9  | +3.0  | +14.2 | +2.3  | -4.4  | -3.0  |
| D-Asp reversed    | -5.0  | +6.6    | +2.1       | +15.1  | +0.9   | +0.3  | -3.1  | +0.2  | -2.8  | -2.6  | -3.1  | -2.7  |
| Succinate (open loop) | -15.1 | -47.4  | -44.1      | +8.0   | -48.6  | -47.3 | +28.2 | +27.3 | -5.5* | -39.8 | +24.7 | +24.8 |
| Succinate (closed loop) | -25.3 | -41.1  | -37.8      | +66.7  | -62.8  | -46.7 | +40.4 | +26.6 | -104.8⁰ | -43.2 | +22.5 | +20.3 |

* For the L-Asp → succinate open loop system, His⁴⁴⁸ is closer than His⁴⁴⁸ to the ligand and so the His⁴⁴⁸ component is given.

** For the D-Asp → succinate closed loop system, both His⁴⁴⁸ and His⁴⁴⁸ are very close to the ligand and so the sum of both histidines is given; His⁴⁴⁸ and His⁴⁴⁸ contribute -46.8 and -58.0 kcal/mol, respectively.
**Chiral Specificity**

**Table 4**

Poisson-Boltzmann (PBFE) free energy component analysis for L-Asp and D-Asp binding to AspRS

Component analysis for the L-Asp and D-Asp ligand binding free energy estimates in AspRS (in kcal/mol), computed in each case from PBFE calculations on 250 MD structures from 1.0-ns native trajectories. The protein dielectric constant ε is set to 4, giving good agreement with MDFE ΔG L-Asp/D-Asp values (Table II). A negative ΔG component indicates ligand stabilization. The last two rows show L-Asp/D-Asp ΔG components (ΔG_{L-Asp} − ΔG_{D-Asp}); a positive value favors L-Asp over D-Asp.

| Ligand               | Total    | Protein | ATP-Mg<sup>3+</sup> | Glu<sup>171</sup> | Lys<sup>178</sup> | Arg<sup>417</sup> | Asp<sup>233</sup> | Glu<sup>235</sup> | His<sup>198</sup> | Arg<sup>499</sup> | Asp<sup>275</sup> | Glu<sup>492</sup> |
|----------------------|----------|---------|---------------------|------------------|------------------|-----------------|-----------------|------------------|-----------------|-----------------|-----------------|-----------------|
| L-Asp                | −46.8    | −32.8   | −9.2               | 11.7             | −10.1            | −6.2            | +16.0           | +19.4            | −10.4           | −17.4           | +18.3           | +17.8           |
| D-Asp                | −40.6    | −35.4   | −0.4               | 11.5             | −13.1            | −8.3            | +19.4           | +19.1            | −5.0            | −15.4           | +12.9           | +14.2           |
| D-Asp inverted       | −40.1    | −35.4   | +0.2               | +11.9            | −12.5            | −7.4            | +18.7           | +19.1            | −6.3            | −15.8           | +13.2           | +14.5           |
| ΔΔG components       |          |         |                    |                  |                  |                 |                 |                  |                 |                 |                 |                 |
| L-Asp/D-Asp          | +6.2     | −2.6    | +8.8               | −0.2             | −3.0             | −2.1            | +3.4            | −0.3             | +5.4            | +2.0            | −5.3            | −3.7            |
| L-Asp/D-Asp inverted | +6.7     | −2.6    | +9.4               | +0.2             | −2.4             | −1.2            | +2.7            | −0.3             | +4.1            | +1.6            | −5.1            | −3.3            |

**L-Asp Versus Succinate Binding to Open Loop and Closed Loop AspRS: MDFE Simulations**—Succinate binding to AspRS represents the most extreme case for ligand ammonium group specificity: in succinate, the ammonium has been removed altogether. As described above, the flexible, flipping loop in AspRS can close in response to amino acid binding, with the negative Glu<sup>171</sup> in the middle of the loop coordinating the NH<sub>3</sub> group of the ligand. Because succinate lacks the NH<sub>3</sub> group, we considered two models: binding to the open and the closed active site. In the present section, we describe MDFE simulations based on each model. In the next section, we describe experiments that yield the binding free energy difference, averaged over both loop states. Finally, in “L-Asp Versus Succinate Binding: Combining Simulations and Experiments,” we analyze the simulation and experimental data in combination, with the help of a thermodynamic cycle.

We first test L-Asp/succinate binding specificity with the flipping loop open. As shown in Table 1, His<sup>449</sup>, next to the ligand (and adjacent to His<sup>448</sup>), remains uncharged with either L-Asp or succinate bound. The succinate binding pocket is quite flexible; succinate and nearby protein side chains have r.m.s. deviations (fluctuations) of 1.9 (0.6) and 2.3 (0.5) Å. The loss of the ligand ammonium causes Ser<sup>193</sup> and Glu<sup>195</sup> to swing away from the ligand. Also notable is the loss of an ordered water molecule on this side of the pocket. Lys<sup>198</sup> and Arg<sup>499</sup> maintain their strong stabilization of the ligand side chain, whereas neutral His<sup>449</sup> also coordinates the ligand. Overall, succinate is actually favored relative to L-Asp in the open loop pocket: ΔΔG = −2.4 kcal/mol (Table 2). Free energy components (Table 3) show that positive binding pocket residues and ATP:Mg<sup>2+</sup> are responsible.

With bound L-Asp, the flipping loop is known to be predominately closed, although the exact populations of the open and closed forms are not known (see below). We therefore performed simulations of succinate binding to closed loop AspRS. We first computed histidine charge states (Table 1) and found that both His<sup>448</sup> and His<sup>449</sup> are charged with bound succinate (but not L-Asp). The charging of His<sup>448</sup> and His<sup>449</sup> allows the loop to close despite the negative succinate charge. Indeed, the positive His<sup>449</sup> sits between Glu<sup>171</sup> of the flipping loop and succinate, counteracting their negative charges and replacing the L-Asp ammonium (Fig. 2D). Enhanced mobility of the histidine residues, however, relative to the L-Asp situation, means the histidines share stabilization of the ligand carboxylates, with His<sup>448</sup> and His<sup>449</sup> interchanging roles.
charged histidines, the L-Asp/succinate MDFE free energy difference constant, we can then deduce that the binding free energy is equal to 21 kcal/mol (Table 1), giving a total dissociation constant of 5.2 kcal/mol. For the ATP-PP\textsubscript{i} exchange reaction, approximation of the Asp dissociation constant by \( K_m \) should be reasonable, because the reaction is known to be limited by the catalytic step, not the substrate binding step (65). The same approximation is quite accurate for the homologous LysRS (64).

To compute the L-Asp/succinate binding free energy difference in the closed loop active site, we first compare L-Asp to succinate with both His\textsuperscript{448} and His\textsuperscript{449} charged. In a second step, we add the His\textsuperscript{449} uncharging free energy with bound L-Asp, to get the net, closed-loop, L-Asp/succinate \( \Delta G \). With the charged histidines, the L-Asp/succinate MDFE free energy difference is just +0.4 kcal/mol (Table 2). The free energy components (Table 4) show a large, pro-L-Asp contribution of Glu\textsuperscript{171} and a compensating, pro-succinate contribution of the charged histidines. To obtain the net \( \Delta G \), however, we must add the +12.3 kcal/mol of His\textsuperscript{449} uncharging energy with bound L-Asp (Table 1), giving a total \( \Delta G \) of +12.7 kcal/mol. Thus, the closed AspRS state strongly prefers Asp over succinate, whereas the open state (above) preferred succinate over Asp. To resolve the competition between them, additional information is needed, which is provided by the kinetic measurements below.

**L-Asp Versus Succinate Binding**

**Kinetic Measurements**—To study succinate binding, initial rates of the L-Asp-dependent ATP-PP\textsubscript{i} exchange reaction were measured in the presence of various succinate and L-Asp concentrations. In the conditions employed, the \( K_m \) value for L-Asp was 320 mM. Addition of succinate to the reaction mixture led to a decrease in the reaction rate. At an L-Asp concentration of 100 mM, 50% inhibition was obtained at a succinate concentration of 210 \( \pm \) 10 mM. At higher concentrations of L-Asp (up to 2 mM), the succinate concentration resulting in a 50% inhibition of the enzyme remained equal to 210 \( \pm \) 10 mM. Therefore, inhibition of AspRS by succinate is weak and mainly non-competitive. Indeed, Fig. 4 shows the ratio \( (V_2/V_1)_{\text{exp}} \) of the rates at two Asp concentrations, [Asp\textsubscript{1}]=100 mM and [Asp\textsubscript{2}]=2 mM, as a function of succinate concentration, [succ]. In the case of competitive inhibition, we expect that the ratio would have the form \( (V_2/V_1)_{\text{comp}} = f([\text{succ}]) = ([\text{Asp}]_2/[\text{Asp}]_1) ([\text{Asp}]_1 + K_m + [\text{succ}] K_m/K_0) ([\text{Asp}]_2 + K_m + [\text{succ}] K_m/K_0) \), where \( K_m \) is the inhibition constant. Because [Asp\textsubscript{1}] is much greater than \( K_m \), \( f([\text{succ}]) \approx (1/[\text{Asp}]_1) ([\text{Asp}]_1 + K_m + [\text{succ}] K_m/K_0) \), a linear function of [succ]. The function \( f \) is plotted in Fig. 5, using a range of \( K_m \) values. The experimental ratio \( (V_2/V_1)_{\text{exp}} \) is weakly dependent on [succ]. The smallest value of \( K_m \) that is consistent with the experimental behavior is 2000 mM. Therefore, the lack of dependence of succinate inhibition on the L-Asp concentration indicates that the succinate dissociation constant \( K_d \) is at least 2 M, and succinate inhibition is mainly non-competitive. By assuming the \( K_m \) value for L-Asp approximates the dissociation constant, we can then deduce that the binding free energy difference, \( \Delta \Delta G_{\text{exp}} \), between L-Asp and succinate is at least \( RT \log K_d/K_m \approx 5.2 \) kcal/mol. For the ATP-PP\textsubscript{i} exchange reaction, approximation of the Asp dissociation constant by \( K_m \) should
Combining Simulations and Experiments—In this section, we combine the experimental and simulation data to understand the role of the two enzyme conformations, open and closed, more clearly. Our analysis is based on the thermodynamic cycle in Fig. 5A. Closed, left, and open, right, represent the enzyme restricted to its closed or open states, respectively. “AspRS,” in the center, represents the unrestricted enzyme, which occupies a mixture of the two states. The vertical arrows represent binding of succinate and unbinding of aspartate, assumed to be in equal, low concentrations. The free energy for the central arrow is the experimental \( \Delta G_{\text{exp}} \) for which only a lower bound could be measured: \( \Delta G_{\text{exp}} > 5.2 \text{ kcal/mol} \). The free energies for the left- and right-hand vertical arrows were estimated by MDFE, above: \( \Delta G_c = 12.7 \text{ kcal/mol} \) and \( \Delta G_o = -2.4 \text{ kcal/mol} \). Horizontal arrows establish a conformational restriction, eliminating either the open (left) or closed (right) states. Such a restriction necessarily leads to an increase in free energy: \( \Delta G_o, \Delta G_c, \Delta G_{\text{exp}} \), and \( \Delta G'_{\text{exp}} \) are all positive. Indeed, the free energy associated with the lower left arrow, for example, obeys the simple relation (66),

\[
\exp(-\Delta G_o/k_B T) = f_{\text{closed}} \exp(-U/kT)/[f_{\text{closed}} \exp(-U/kT) + f_{\text{open}} \exp(-U/kT)] \quad \text{(Eq. 1)}
\]

where \( k_B \) is Boltzmann’s constant, \( T \) is the temperature, and the subscripts indicate integrals over the regions of conformational space corresponding to the open or closed loop state. From Equation 1, we can obtain Equation 2, below.

\[
\exp(-\Delta G_o/k_B T) + \exp(-\Delta G_o/k_B T) = \exp(-\Delta G'_{\text{exp}}/k_B T)
\]

Thus, each pair of free energies is linked. Each exponential in this relationship represents the population of the corresponding conformation in the AspRS-Asp or AspRS-succinate complex; the populations sum to one as expected.

By considering the right-hand part of the cycle, we can obtain an upper bound for \( \Delta G_o \) and establish that the closed loop is strongly preferred with bound \( \ell \)-Asp. Indeed, we have \( \Delta G_o = \Delta G_{\text{exp}} + \Delta G'_{\text{exp}} - \Delta G_{\text{circ}} \), with \( \Delta G_{\text{circ}} = -2.4 \text{ kcal/mol} \), \( \Delta G_{\text{exp}} > 5.2 \text{ kcal/mol} \), and \( \Delta G'_{\text{exp}} > 0 \). It follows that \( \Delta G_o > 7.6 \text{ kcal/mol} \). If we allow for the MDFE uncertainty on \( \Delta \Delta G_o \) (2 kcal/mol), we obtain \( \Delta G_o > 5.6 \text{ kcal/mol} \). An open loop population of 50% in the Asp-AspRS complex would give \( \Delta G_o = kT \log[2] \), about 0.4 kcal/mol at room temperature. The larger \( \Delta G_o \) value indicates that the open loop population is very much smaller, essentially zero when \( \ell \)-Asp is bound to AspRS. This result agrees with the observation of a closed loop in the Pyrococcus kodakaraensis Asp-AspRS crystal structure (12). Note that the crystal structure showed only that the closed loop population was over 50%; in fact it is essentially 100%.

By considering the left-hand part of the cycle, we can now obtain an upper bound for \( \Delta \Delta G_{\text{exp}} \). Because the closed state is predominant in the AspRS-Asp complex, \( \Delta G_o \) cannot be larger than \( k_B T \log[2] = 0.4 \text{ kcal/mol} \). In fact, using \( \Delta G_o > 5.6 \text{ kcal/mol} \) and Equation 2, we find \( \Delta G_o < 0.01 \text{ kcal/mol} \). Therefore, \( \Delta \Delta G_{\text{exp}} < 12.7 \text{ kcal/mol} \). Thus, simple thermodynamic relations combine with the MDFE data to narrow the possible range for \( \Delta \Delta G_{\text{exp}} \) considerably.

If we approximate \( \Delta G_o \) by 0 kcal/mol and we vary \( \Delta \Delta G_{\text{exp}} \) in the allowed range, \( 5.2 < \Delta \Delta G_{\text{exp}} < 12.7 \text{ kcal/mol} \), all the other free energies are completely determined: \( \Delta G'_{\text{exp}} = \Delta G_{\text{exp}} - \Delta \Delta G_{\text{exp}} \), \( \Delta G'_{\text{circ}} = -k_B T \log[1 - \exp(-\Delta G'_{\text{exp}}/kT)] \), \( \Delta G_o = \Delta \Delta G_{\text{exp}} + \Delta G'_{\text{circ}} - \Delta G_{\text{circ}} \). These three free energies are plotted in Fig. 5B. We see that if \( \Delta \Delta G_{\text{exp}} = 5.2 \text{ kcal/mol} \), the succinate complex with the closed loop has a very high free energy (6 kcal/mol). If \( \Delta \Delta G_{\text{exp}} = 12.7 \text{ kcal/mol} \), the succinate complex with the open loop has a very high free energy (6 kcal/mol). Thus, the data do not allow us to infer which state binds succinate most strongly. Nevertheless, we saw that Asp strongly prefers the closed state. Because succinate binds much less strongly than Asp to this state, its overall binding free energy to AspRS is much weaker than that of Asp.

Control Simulations

\( \ell \)-Tyr and \( d \)-Tyr Binding to TyrRS—As a control, we also simulated another synthetase that has a measurable activity for a \( d \)-amino acid: tyrosyl-tRNA synthetase. The binding free energy difference between \( \ell \)-Tyr and \( d \)-Tyr was computed by MDFE, as above. Twenty runs were performed. The computed \( \Delta G \) was \( +3.1 \text{ kcal/mol} \), favoring \( \ell \)-Tyr (supplementary materials Table SM2). This low specificity is consistent with the known \( d \)-Tyr tRNA\( \text{Tyr} \) acylation rates (26). Indeed, converting the ratio of the experimental \( \ell \)-Tyr and \( d \)-Tyr acylation rates into a binding free energy difference gives a \( \Delta G \) value of \( +2.2 \text{ kcal/mol} \), close to the MDFE estimate. What is more, the specificity computed for TyrRS is lower than that computed for AspRS, in agreement with the experiment. Indeed, the experimental \( \ell \)-Asp and \( d \)-Asp acylation rates correspond (see above) to a \( \Delta G \) of 4.9 kcal/mol, 1.8 kcal/mol larger than for TyrRS. This is very close to the difference obtained by MDFE (\( \Delta G = 3.1 \text{ kcal/mol} \) for TyrRS; \( \Delta G = 4.7 \) or 5.9 kcal/mol for AspRS). Thus, the correspondence between binding free energy differences and ratios of acylation rates appears to hold qualitatively, and the TyrRS simulations provide further support for the validity and accuracy of the MDFE methodology.

\( \ell \)-Asp Binding to Q199E-AspRS—These simulations provide an additional test for MDFE. The single mutant Q199E-AspRS is known to be active for aspartyl-adenylate formation and the Michaelis constants measured for the native and mutant proteins are 2.5 and 9.1 mM, respectively (10). As discussed, their ratio can be used to estimate (at least roughly) the change in Asp binding free energy, 0.8 kcal/mol. Before computing the change in the \( \ell \)-Asp binding free energy due to the mutation, we first considered the protonation states of all nearby histidines. As shown in supplementary materials Table SM2, introducing the negative Glu199 induces the uptake of a proton by His528, located just 3.1 Å away. His528 is completely conserved in AspRSs. The positive charge of the proton effectively neutralizes the negative Glu199 charge. The binding free energy change was therefore computed by alchemically transforming Glu199 into Glu and simultaneously changing His528 to its doubly protonated state, using the thermodynamic cycle shown in supplementary materials Fig. SM3 and charge sets in Fig. SM2. The
transformation was performed both in the presence and absence of the ligand Asp. The resulting free energy changes were subtracted to obtain the Asp binding free energy change (44). Four free energy simulations were done, of 5 ns each, performing the transformation in either direction (Gln to Glu or the reverse). The change in the Asp binding free energy due to the mutation was computed to be +1.0 kcal/mol (supplemental materials Table SM1), in very good agreement with the experimental estimate of +0.8 kcal/mol (where the positive sign indicates reduced Asp binding affinity). Poisson-Boltzmann estimates give +0.1 kcal/mol (supplemental materials Table SM1), in good agreement. Artificially constraining His528 to keep its cationic group destabilizes the ligand side chain coordination to these positive residues. In species (primarily Archaea) that lack this E. coli AspRS histidine loop, we may anticipate a slightly lower chiral specificity. Arg489 is the class II aaRS invariant arginine, conserved in all AspRS enzymes and functionally irreplaceable.

The TyrRS control simulations are also consistent with the TyrRS structure (12, 15); others are less obvious. For example, we saw that His448 becomes doubly protonated when D-Asp is present, yet D-Asp could acylate tRNA with its side chain. (i) D-Asp has a computed binding free energy 5.9 ± 2 kcal/mol higher than L-Asp. (ii) Inverted D-Asp binding is slightly better than regular binding (ΔΔG = 4.7 ± 2 kcal/mol), raising the possibility that D-Asp could acylate tRNA with its side chain. (i) and (ii) show that the reduced AspRS activity for D-Asp (26) can be explained by unfavorable binding. (iii) Succinate binding is weak and probably very weak: ΔΔG = 9 ± 3.8 kcal/mol; this estimate combined simulations, experiments, and thermodynamic arguments. (iv) Inverted L-Asp binding is extremely weak. The simulations have provided a detailed model to explain these preferences.

The moderate L-Asp/D-Asp binding free energy difference makes it difficult to pinpoint binding pocket mutations that would strongly favor either stereoisomer, as shown by the low individual free energy components in Tables 3 and 4. Only His448 and, to a lesser extent, Arg489 show a significant preference for L-Asp. Displacement of the backbone ammonium group destabilizes the ligand side chain coordination to these positive residues. In species (primarily Archaea) that lack this E. coli AspRS histidine loop, we may anticipate a slightly lower chiral specificity. Arg489 is the class II aaRS invariant arginine, conserved in all AspRS enzymes and functionally irreplaceable.

The extent of discrimination against inverted L-Asp is somewhat surprising: rotating L-Asp into the inverted orientation changes the binding free energy by 10 kcal/mol. Yet this rotation can be viewed as a simple shift of the ammonium group from the α to the β carbon. One might have anticipated that steric overlap between the displaced ammonium and the protein would be the main source of the binding free energy penalty. In fact, both the ligand and protein shift during the simulations to alleviate steric overlap (Fig. 3), so that van der Waals interactions make a minor contribution (31%) to the overall binding free energy difference. Instead, the large discrimination is accomplished by Coulomb interactions with Glu171 of the flipping loop, which enforces the correct L-Asp orientation despite being itself a flexible and mobile group. The succinate analysis showed, indirectly, that with bound L-Asp, the flipping loop is closed essentially 100% of the time.

Fig. 6 illustrates how Glu171, His448, and Arg489 mirror the pseudosymmetry of L-Asp. Overall, the experimental L-Asp binding geometry appears to be unique in that (a) the flipping loop residue Glu171 greatly stabilizes the backbone ammonium group, unlike inverted L-Asp and succinate; and (b) His448 greatly stabilizes the side chain carboxylate (with some help from Arg489), unlike regular and inverted D-Asp.

Several of the above interactions are visible in the L-Asp: AspRS crystal structure (12, 15); others are less obvious. For example, we saw that His448 becomes doubly protonated when the flipping loop closes, bringing Glu171 into the binding pocket. Note that the experimental evidence for the doubly protonated His448 is indirect, and relies partly on MD simulations. Specifically, Asn inhibition measurements are well reproduced

DISCUSSION

Free Energy Simulations as a Complement to Experiments—MD free energy simulations have been extensively validated for problems such as redox properties, acid-base equilibria, enzyme mechanism, and protein-ligand binding (21–25, 32–39, 44–47, 67). We recently studied the preference of AspRS for Asp over the neutral analogue Asn (24), and found good agreement between MDFE and measurements of AspRS inhibition by Asn. Predictions of the magnesium dependence of AspRS kinetics (25) were also confirmed experimentally. Here, we reported control simulations of both AspRS and TyrRS. For Q199E-AspRS, we compared the computed reduction in L-Asp affinity to an experimental estimate based on the ratio of Asp Michaelis constants for pyrophosphate exchange (10). The good agreement supports both the simulation model and the close relation between the K_m and K_d ratios for this enzyme. For the homologous class IIb LysRS, the Lys dissociation constant is known to be well approximated by the Michaelis constant (64). The TyrRS control simulations are also consistent with the experiment.

Thus validated, the MDFE simulations provide structural and thermodynamic information that would be difficult to obtain from experiments alone. One example is their ability to characterize weakly populated states, such as the minor flipping loop conformation: simulating both the major and minor conformations with bound succinate was necessary to fully characterize the major (closed) state with bound L-Asp. Another example is their ability to identify couplings between amino acids (including long-range couplings), which can require elaborate, multiple mutant cycles for experimental characterization. This is illustrated by His528, whose protonation state changes when Gln199 is mutated. More generally, the free energy component analysis (Tables 3 and 4) provides a qualitative measure of couplings throughout the binding pocket.

The Basis for AspRS Specificity—The main question posed here is the mechanism by which AspRS achieves its preference for L-Asp in the correct orientation, with respect not only to D-Asp, but to inverted L-Asp and to succinate (Fig. 1). This question is slightly broader than that of chiral L-Asp/D-Asp discrimination. Asp has a pseudosymmetry with respect to backbone/side chain exchange, which enhances the chemical similarity between the ligands and binding modes (Fig. 1A). We recall the binding preferences: (i) D-Asp has a computed binding free energy 5.9 ± 2 kcal/mol higher than L-Asp. (ii) Inverted D-Asp binding is slightly better than regular binding (ΔΔG = 4.7 ± 2 kcal/mol), raising the possibility that D-Asp could acylate tRNA with its side chain. (i) and (ii) show that the reduced AspRS activity for D-Asp (26) can be explained by unfavorable binding. (iii) Succinate binding is weak and probably very weak: ΔΔG = 9 ± 3.8 kcal/mol; this estimate combined simulations, experiments, and thermodynamic arguments. (iv) Inverted L-Asp binding is extremely weak. The simulations have provided a detailed model to explain these preferences.

The moderate L-Asp/D-Asp binding free energy difference makes it difficult to pinpoint binding pocket mutations that would strongly favor either stereoisomer, as shown by the low individual free energy components in Tables 3 and 4. Only His448 and, to a lesser extent, Arg489 show a significant preference for L-Asp. Displacement of the backbone ammonium group destabilizes the ligand side chain coordination to these positive residues. In species (primarily Archaea) that lack this E. coli AspRS histidine loop, we may anticipate a slightly lower chiral specificity. Arg489 is the class II aaRS invariant arginine, conserved in all AspRS enzymes and functionally irreplaceable.

The extent of discrimination against inverted L-Asp is somewhat surprising: rotating L-Asp into the inverted orientation changes the binding free energy by 10 kcal/mol. Yet this rotation can be viewed as a simple shift of the ammonium group from the α to the β carbon. One might have anticipated that steric overlap between the displaced ammonium and the protein would be the main source of the binding free energy penalty. In fact, both the ligand and protein shift during the simulations to alleviate steric overlap (Fig. 3), so that van der Waals interactions make a minor contribution (31%) to the overall binding free energy difference. Instead, the large discrimination is accomplished by Coulomb interactions with Glu171 of the flipping loop, which enforces the correct L-Asp orientation despite being itself a flexible and mobile group. The succinate analysis showed, indirectly, that with bound L-Asp, the flipping loop is closed essentially 100% of the time.

Fig. 6 illustrates how Glu171, His448, and Arg489 mirror the pseudosymmetry of L-Asp. Overall, the experimental L-Asp binding geometry appears to be unique in that (a) the flipping loop residue Glu171 greatly stabilizes the backbone ammonium group, unlike inverted L-Asp and succinate; and (b) His448 greatly stabilizes the side chain carboxylate (with some help from Arg489), unlike regular and inverted D-Asp.

Several of the above interactions are visible in the L-Asp: AspRS crystal structure (12, 15); others are less obvious. For example, we saw that His448 becomes doubly protonated when the flipping loop closes, bringing Glu171 into the binding pocket. Note that the experimental evidence for the doubly protonated His448 is indirect, and relies partly on MD simulations. Specifically, Asn inhibition measurements are well reproduced.
by MDFE simulations where His448 is protonated, but not by simulations where it is neutral; see Ref. 24 for details. His449 becomes doubly protonated when the ligand ammonium is removed (succinate complex). Strong, but compensating interactions were seen, involving groups of residues or cofactors: e.g. Lys198 and Asp233, ATP, Asp475 and Glu482. Thus, “hidden” interactions are revealed, the most crucial being the histidine switches. Such hidden effects and compensating interactions make it difficult to guess the precise effect of a point mutation ahead of time; for example, we had not anticipated that the Q199E mutation would lead to proton uptake by nearby His528.

Similarly, the role of steric overlap in inverted L-Asp binding is smaller than might have been anticipated from a simple structural model.

The amino acids discussed here are all partly or completely conserved in AspRSs, suggesting that the observed effects are rather general. The main exceptions are the absence of His448 and His449 in Archaea and the absence of Glu171 in some eukaryotes. It remains to be seen whether these organisms have reduced specificity or whether alternate interactions can provide the same level of specificity.

**The Strength of Electrostatic Interactions**—By considering all four possible positions on the amino acid ligand, as well as ammonium deletion, we could investigate a secondary question: the strength of electrostatic interactions in an enzyme active site. Although the essential role of electrostatics in enzyme activity is well established (37–43), a great deal can still be learned from specific systems. This and earlier studies (21–25) show that AspRS amino acid specificity arises from a complex network of electrostatic interactions, including long-range interactions (24, 25). One interesting feature is the magnitude of solvent shielding, which effectively rescues the electrostatic interactions and is impossible to guess from a crystal structure. It is reflected by the smaller PBFE components (Table 4), compared with the MDFE components (Table 3). The PBFE components incorporate solvent shielding, and are reduced by a factor of four. This reduction factor is consistent with earlier studies of dielectric shielding in AspRS (24). Another interesting feature is the tight coupling within the interaction network, illustrated by the ability of all four active site histidines to buffer charge modifications. This charge buffering can be viewed as an electrostatic induced fit. Despite the buffering, inserting a positive or a negative charge on the cognate L-Asp ligand has a drastic effect on the binding affinity. We showed earlier that the binding free energy difference between L-Asp and L-Asn is about 21 kcal/mol (26); the difference between L-Asp and succinate is about 9 kcal/mol (±3.8 kcal/mol). Thus, the AspRS active site discriminates very effectively against both positive and negative charge insertion.

**CONCLUSIONS**

We have continued here to probe the interactions within the AspRS active site and their relation to enzyme specificity. AspRS has the special problem of discriminating against inverted ligand orientations, in addition to D-amino acids. Despite weak ligand binding and sizeable error bars for the computed free energies, the main qualitative effects could be identified. The AspRS active site is well protected against both positive and negative charge insertion. Discrimination against a displacement of the substrate ammonium group, a charge-conserving process, is more variable, with the flipping loop, the active site histidines, and nearby charged side chains all playing a role. There is thus a strong protection against inverted L-Asp, but only moderate protection against D-Asp, in agreement with the experiments. Our analysis underlines the considerable complexity that protein electrostatics and molecular recognition can exhibit. With the increased availability of reliable methods for free energy simulations (32–39, 68, 69) and continuing increases in computer power, integrated simulation and experimental studies will become more common. A combined approach could
lead to a successful, rational engineering of aminoacyl-tRNA synthetase specificity in the near future.

Acknowledgments—We thank Martin Karplus for the CHARMM program. Simulations were performed at the CINES Supercomputer Center of the French Ministry of Research, at Tyndall National Institute using in-house clusters provided by Science Foundation Ireland (SFI) and also at the SFI/HEA Irish Centre for High-End Computing (ICHEC).

REFERENCES

1. Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002) Biochemistry, 5th Ed., W. H. Freeman, New York
2. Ibbá, M., Franchklyn, C., and Cusack, S. (eds) (2005) Aminoacyl-tRNA Synthetases, Landes Bioscience, Georgetown, TX
3. Arnez, J. G., and Moras, D. (1994) in RNA–Protein Interactions (Nagai, K., Matsu, I., eds) pp. 52–81, Oxford University Press, Oxford, UK
4. Meinnel, T., Mchulam, Y., and Blanquet, S. (1995) in tRNA: Structure, Biosynthesis, and Function (Soll, D., and Raj Bhandary, T. L., eds) pp. 251–291, ASM Press, Washington, D. C.
5. Schimmel, P., and Ribas de Pouplana, L. (2000) Trends Biochem. Sci. 25, 207–209
6. Pallanck, L., Pak, M., and Schulman, L. H. (1995) in RNA: Structure, Biosynthesis and Function (Soll, D., and Raj Bhandary, T. L., eds) pp. 371–393, ASM Press, Washington, D. C.
7. Arnez, J. G., and Moras, D. (1997) Trends Biochem. Sci. 22, 211–216
8. Franchklyn, C. S. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 9560–9562
9. Ruff, M., Krishnaaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991) Science 252, 1682–1689
10. Cavarelli, J., Eriani, B., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J. C., and Moras, D. (1994) EMBO J. 13, 327–337
11. Peterman, A., Delarue, M., Thierry, J. C., and Moras, D. (1994) J. Mol. Biol. 244, 158–167
12. Schmitt, E., Moulinier, L., Fujiwara, S., Imanaka, T., Thierry, J. C., and Moras, D. (1998) EMBO J. 17, 5227–5237
13. Eiler, S., Docks-Breggeon, A. C., Moulinier, L., Thierry, J. C., and Moras, D. (1999) EMBO J. 18, 6532–6542
14. Briand, C., Poterszman, A., Eiler, S., Webber, G., Thierry, J. C., and Moras, D. (2000) J. Mol. Biol. 299, 1051–1060
15. Moulinier, L., Eiler, S., Eriani, G., Gangloff, J., Thierry, J. C., Gabriel, K., McClain, W. H., and Moras, D. (2001) EMBO J. 20, 5290–5301
16. Ador, L., Jaeger, S., Geslain, R., Martin, F., Cavarelli, J., and Eriani, G. (2004) Biochemistry 43, 7028–7037
17. Berthet-Colominas, C., Seignovert, L., Hartlein, M., Grotli, M., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Naka, K., Kline, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. I., Cammi, R., Polcini, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Darenberg, J. K., Zakrzewski, V. G., Dauprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Mallick, D. K., Rabuck, A. D., Rachvachari, K., Foresman, J., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacomb, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2003) Gaussian 03, revision B.04, Gaussian, Inc., Pittsburgh, PA
18. Becke, A. D. (1993) J. Chem. Phys. 98, 5648–5652
19. Reed, A. E., Weinhold, F., J., and Curtiss, L. A. (1988) Chem. Rev. 88, 899–926
20. MacKerrell, A. D., Bashford, D., Bellot, M., Dunbrack, D. L., Evanseck, J. D., and Field, M. J. (1998) J. Phys. Chem. B 102, 3586–3616
21. Archontis, G., and Simonson, T. (2005) Biochem. J. 388, 3884–3904
22. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) J. Comp. Chem. 4, 187–217
23. Kobayashi, T., Takimura, T., Sekine, R., Vinvent, K., Kamata, K., Sakamoto, K., Nishimura, S., and Yokoyama, S. (2005) J. Mol. Biol. 346, 105–117
24. Simonson, T. (2000) J. Phys. Chem. B 104, 6509–6513
25. Beglov, D., and Roux, B. (1994) J. Phys. Chem. 98, 9050–9063
26. Stote, R. H., States, D. J., and Karplus, M. (1991) J. Chem. Phys. 88, 2419–2433
27. Kralakis, P. J. (1991) Appl. Crystlallogr. 24, 946–950
28. Boresch, S., Archontis, G., and Karplus, M. (1994) Proteins Struct. Funct. Genet. 20, 25–33
Chiral Specificity

62. Yang, A.-S., and Honig, B. (1993) J. Mol. Biol. 231, 459–474
63. Alexov, E. G., and Gunner, M. R. (1997) Biophys. J. 72, 2075–2093
64. Takita, T., Nakagoshi, M., Inouye, K., and Tonomura, B. (2003) J. Mol. Biol. 325, 677–695
65. Giege, R., and Rees, B. (2005) in Aminoacyl-tRNA Synthetases (Ibba, M., Francklyn, C., Cusack, S., eds) pp. 210–226, Landes Bioscience, Georgetown, TX
66. McQuarrie, D. (1975) Statistical Mechanics, Harper and Row, New York
67. Tajkhorshid, E., Baudry, J., Schulten, K., and Suhai, S. (2000) Biophys. J. 78, 683–693
68. Strajbl, M., Shurki, A., and Warshel, A. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 14834–14839
69. Warshel, A. (2003) Annu. Rev. Biophys. Biomol. Struct. 32, 425–443