BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

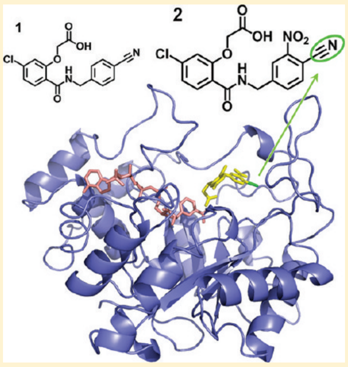
Electrostatic Fields near the Active Site of Human Aldose Reductase: 2. New Inhibitors and Complications Caused by Hydrogen Bonds

Lin Xu,[†] Aina E. Cohen,[‡] and Steven G. Boxer*,[†]

[†]Department of Chemistry, Stanford University, Stanford, California 94305-5080, United States
[‡]Stanford Synchrotron Radiation Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, United States

Supporting Information

ABSTRACT: Vibrational Stark effect spectroscopy was used to measure electrostatic fields in the hydrophobic region of the active site of human aldose reductase (*h*ALR2). A new nitrile-containing inhibitor was designed and synthesized, and the X-ray structure of its complex, along with cofactor NADP$^+$, with wild-type *h*ALR2 was determined at 1.3 Å resolution. The nitrile is found to be in the proximity of T113, consistent with a hydrogen bond interaction. Two vibrational absorption peaks were observed at room temperature in the nitrile region when the inhibitor binds to wild-type *h*ALR2, indicating that the nitrile probe experiences two different microenvironments, and these could be empirically separated into a hydrogen-bonded and non-hydrogen-bonded population by comparison with the T113A mutant, in which a hydrogen bond to the nitrile is not present. Classical molecular dynamics simulations based on the structure predict a double-peak distribution in protein electric fields projected along the nitrile probe. The interpretation of these two peaks as a hydrogen bond formation−dissociation process between the probe nitrile group and a nearby amino acid side chain is used to explain the observation of two IR bands, and the simulations were used to investigate the molecular details of this conformational change. Hydrogen bonding complicates the simplest analysis of vibrational frequency shifts as being due solely to electrostatic interactions through the vibrational Stark effect, and the consequences of this complication are discussed.



The organization of charged, polar, and polarizable residues in the tertiary structure of a protein creates large internal electrostatic fields, which are expected to affect many aspects of protein function.[1−3] Vibrational Stark effect (VSE) spectroscopy has been developed in recent years as a strategy to measure the magnitude and direction of protein electrostatic fields.[4−9] This method has been applied in several systems, including myoglobin,[4] human aldose reductase (*h*ALR2),[5,6] *Pseudomonas putida* ketosteroid isomerase (KSI),[7,8] and ribonuclease S (RNase S).[9] *h*ALR2 is a particularly attractive target for such studies as a large number of very high-resolution structures are available.[10−12] *h*ALR2 is a 36 kDa aldo-keto reductase that catalyzes the conversion of aldehydes derived from carbohydrate metabolism to corresponding alcohols using NADPH as the reducing reagent.[13,14] Because its involvement in the polyol pathway is believed to be responsible for many long-term diabetic complications, researchers have been working for decades on developing tight-binding inhibitors of *h*ALR2 for potential pharmacological treatment of diabetes.[10−20] Among a variety of *h*ALR2 inhibitors, there is a class of carboxylate-type inhibitors developed at the Institute for Diabetes Discovery (IDD), whose molecular scaffold allows for a variety of chemical modifications while still maintaining tight binding affinity.[19] This provides us the starting point for designing novel IDD-type inhibitors with Stark probes such as nitriles[5,6] to study the electrostatic environment in *h*ALR2.

The quantitative basis of this approach derives from the relationship between changes in protein electric fields, $\Delta\vec{F}_{\text{protein}}$ (MV/cm) and observed molecular vibrational frequency shifts, $\Delta\bar{\nu}^{\text{obs}}$ (cm^{-1}):

$$hc\Delta\bar{\nu}^{\text{obs}} = -\Delta\vec{\mu}\cdot\Delta\vec{F}_{\text{protein}} \tag{1}$$

where $|\Delta\vec{\mu}|$ (cm^{-1}/(MV/cm) is the Stark tuning rate, a measure of the sensitivity of the transition frequency to an electric field, h is Planck's constant, and c is the speed of light. We have shown that nitriles have a desirable combination of extinction coefficient in a relatively uncluttered region of the IR spectrum and a relatively large Stark tuning rate.[21] Nitriles can be incorporated on unnatural or modified amino acids or inhibitors.[5,6,22] The dot product in eq 1 can be reduced to a measure of the projection of the protein electric field along the −CN bond vector, $\Delta\vec{F}_{\parallel}^{\text{protein}}$, because $\Delta\vec{\mu}_{\text{CN}}$ has been shown to be parallel to the transition dipole moment, which in turn is parallel to the −CN bond.[23] The magnitude of the Stark tuning rate, $|\Delta\vec{\mu}_{\text{CN}}|$, is obtained by measuring the vibrational Stark spectrum in an external electric field, either for closely related compounds in organic glasses or in the protein,[9] and the

Received: June 17, 2011
Revised: August 5, 2011
Published: August 22, 2011

dx.doi.org/10.1021/bi200930f | *Biochemistry* 2011, 50, 8311−8322

direction of the −CN bond is obtained by X-ray crystallography. Thus, observed frequency shifts $(\Delta\bar{\nu}_{\mathrm{CN}}^{\mathrm{obs}})$ can be translated into values for the projection of changes in the protein electric field along the −CN bond axis, $\Delta\vec{F}_{\|}^{\mathrm{protein}}$, and this can be compared with simulations that are widely used to estimate such fields.

In part 1 of this series, a nitrile probe was delivered to the hydrophobic specificity pocket of *h*ALR2 through the binding of an IDD-type inhibitor, **1**, (5-chloro-2-{[(4-cyanobenzyl)-amino]carbonyl}phenoxy)acetic acid, that we prepared (Figure 1).[6] Changes in electrostatic fields due to mutations in the vicinity of the nitrile based on a modeled structure were quantified using nitrile stretching frequency shifts and compared with calculation results. According to the modeled structure of the *h*ALR2−**1**−NADP$^+$ complex, a hydrogen bond donor from an amino acid side chain in the vicinity of the nitrile probe might complicate the interpretation of observed frequency shifts as perturbations in electric fields in the sense described by eq 1, where $\Delta\vec{F}_{\mathrm{protein}}$ is meant to describe the electric field change due to a change in the organized environment, not specific, local chemical interactions such as hydrogen bonds, which may have an electrostatic component, but may also depend on overlap of the wave functions on the donor and acceptor, a contribution that is not included in eq 1.[8,9]

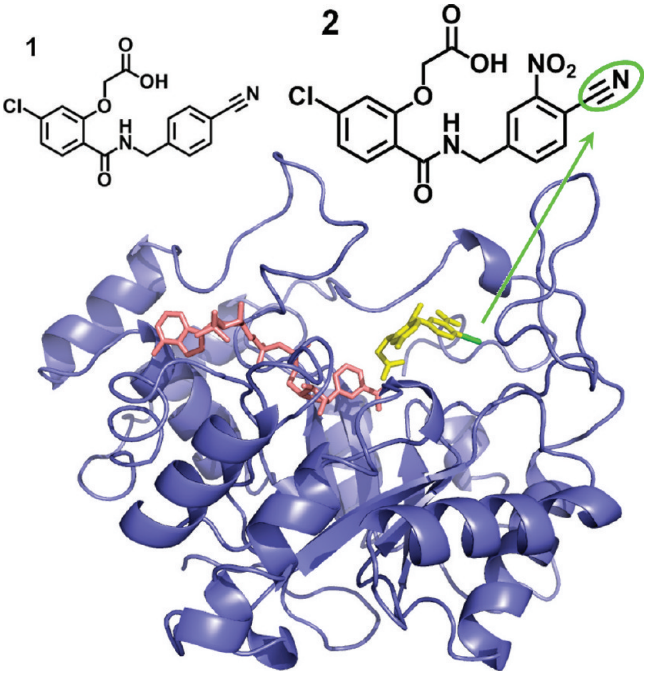


Figure 1. Structural model of X-ray data for the *h*ALR2−**2**−NADP$^+$ complex. The cofactor NADP$^+$ is colored red, inhibitor **2** yellow, and the nitrile probe green. The structure of inhibitor **1**[6] is shown at the top left.

In this work, we have modified the molecular structure of **1** by introducing the nitro functionality at the meta position of the amine moiety [inhibitor **2**, (5-chloro-2-{[(4-cyano-3-nitrobenzyl)amino]carbonyl}phenoxy)acetic acid (Figure 1)]. This molecule was initially designed for the purpose of increasing inhibitor binding affinity because it has been shown that adding a nitro group in the meta position increases the binding affinity of *h*ALR2 by 200-fold.[19,24] This modification also makes the orientation of the probe more unambiguous in the active site by preventing phenyl ring rotation. As described in the following, the structure of the tertiary complex of wild-type *h*ALR2-**2**-NADP$^+$ was determined by X-ray crystallography at 1.3 Å resolution. This structure shows that the nitrile probe is buried deep in the hydrophobic region of the binding pocket and well protected from water. It also shows that the probe forms a short contact with the side chain of residue T113. Interestingly, we found two IR peaks in the nitrile region when **2** bound to WT *h*ALR2, and IR spectra of two *h*ALR2 mutants helped in the assignment of their origin. Classical molecular dynamics (MD) simulations were performed to provide a molecular-level understanding of observed IR features. The complications due to hydrogen bonds for using VSE spectroscopy to study protein electrostatic fields are discussed.

■ MATERIALS AND METHODS

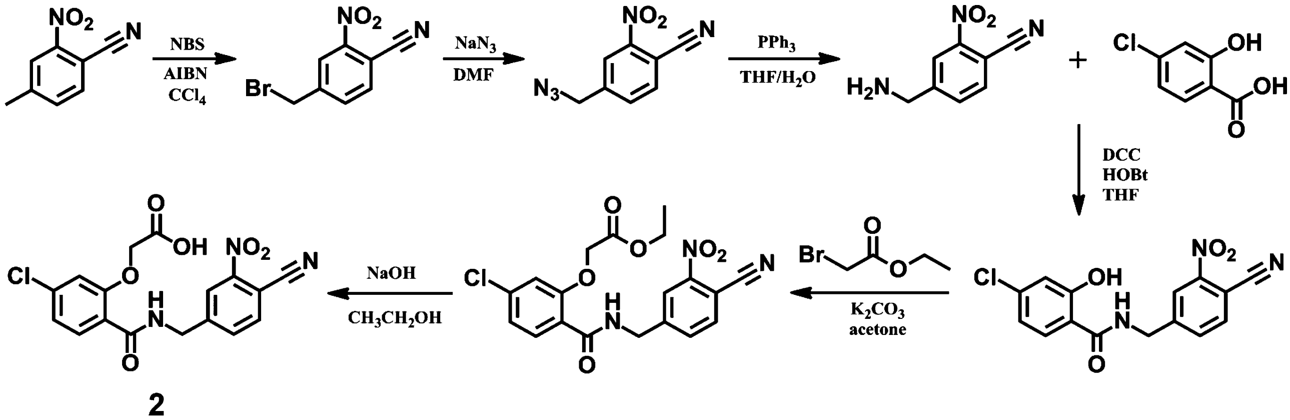
Inhibitor Synthesis. The nitrile-containing inhibitor **2**, (5-chloro-2-{[(4-cyano-3-nitrobenzyl)amino]carbonyl}phenoxy)-acetic acid, was synthesized according to the procedure described in Scheme 1. All starting materials were purchased from Aldrich and used without further purification. All reactions were conducted under nitrogen protection. Products from each step were purified by either flash column chromatography (FCC) using silica gel (Aldrich, 230−400 mesh, 60 Å) or high-performance liquid chromatography (HPLC) (Shimadzu). Proton nuclear magnetic resonance spectra (^{1}H NMR) were recorded on a 300 MHz Varian Gemini instrument and mass spectra on a ZQ single-quadrupole electron spray mass spectrometry (MS) instrument.

4-Cyano-3-nitrobenzyl Bromide. *N*-Bromosuccinimide (NBS, 2.4 g) and azobisisobutyronitrile (AIBN, 180 mg) were added to a 16 mL solution of 2 g of 4-methyl-2-nitrobenzonitrile in anhydrous carbon tetrachloride (CCl_4). The reaction mixture was heated at 80 °C overnight. The solution was filtered to remove solids and put under vacuum to remove excess CCl_4. The residue was dissolved in ethyl acetate (EtOAc) and applied to FCC equilibrated in hexane. The product was eluted with 60% (v/v) EtOAc in hexane. The product was identified by the appearance of new peaks on NMR spectra: 4-methyl-2-nitrobenzonitrile (8.3 ppm, s, 1 H; 8.1 ppm, d, 1 H; 7.8 ppm, d, 1 H) versus 4-cyano-3-nitrobenzyl bromide (8.5 ppm, s, 1 H; 8.2 ppm, d, 1 H; 8.1 ppm, d, 1 H; 4.9 ppm, s, 2 H).

4-Cyano-3-nitrobenzyl Azide. The bromide product was dissolved in 60 mL of anhydrous dimethylformamide (DMF); 500 mg of sodium azide was added in small portions over 10 min. The reaction mixture was stirred at room temperature for 24 h. Sodium bromide was removed by filtration through Celite; the filtrate was extracted with EtOAc and water, and the organic layer was dried over $MgSO_4$. The solids of $MgSO_4$ were removed by filtration, and excess solvent was removed under vacuum. The product was identified by the appearance of new peaks on the NMR spectrum of 4-cyano-3-nitrobenzyl azide (8.4 ppm, s, 1 H; 8.2 ppm, d, 1 H; 8.0 ppm, d, 1 H; 4.8 ppm, s, 2 H).

4-Cyano-3-nitrobenzylamine. The azide product was dissolved in a 12 mL mixture of tetrahydrofuran (THF) and water (10:1, v/v). Triphenylphosphine (1.8 g) was added in small portions over 10 min, and the reaction mixture was stirred at room temperature overnight. Excess solvent was removed

Scheme 1. Synthetic Pathway of 2, (5-Chloro-2-{[(4-cyano-3-nitrobenzyl)amino]carbonyl}phenoxy)acetic Acid



under vacuum, and the residue was dissolved in dichloromethane and extracted with a 0.2 M HCl aqueous solution. The aqueous phase was combined and cooled to 0 °C; 1.0 M NaOH was added until the pH of the solution reached 10, and the product was extracted with dichloromethane, washed with saturated NaCl (aqueous), and dried over $MgSO_4$. The solids were removed by filtration. Excess solvent was removed under vacuum. The product was identified by NMR (8.4 ppm, s, 1 H; 8.1 ppm, d, 1 H; 7.9 ppm, d, 1 H; 3.9 ppm, s, 2 H; 2.1 ppm, s, 2 H) and MS [m/z 177.9 (ES^+), 178.2 g/mol (calculated)].

4-Chloro-N-(4-cyano-3-nitrobenzyl)-2-hydroxylbenzamide. The amine product was dissolved in a 6 mL solution of 1.0 g of 4-chloro-2-hydroxybenzoic acid and 100 mg of hydroxybenzotriazole (HOBt) in anhydrous THF. A 2 mL solution of 1.1 g of *N,N'*-dicyclohexylcarbodiimide (DCC) in THF was added dropwise over several minutes, and the reaction mixture was stirred at room temperature overnight. Solids were removed by filtration. The residue was subjected to reverse phase HPLC (30 to 60% acetonitrile in water, 5 mL/min, 30 min), and excess solvent was removed by lyophilization.

Ethyl (5-Chloro-2-{[(4-cyano-3-nitrobenzyl)amino]carbonyl}phenoxy)acetate. The amide product was dissolved in an 8 mL solution of 300 mg of K_2CO_3 in anhydrous acetone, and 200 μL of ethyl bromoacetate was added. The solution was heated at 56 °C overnight; 1.0 M HCl was added until the pH of the solution reached 1, and the mixture was extracted with EtOAc. The organic phase was combined, washed with saturated NaCl, and dried over $MgSO_4$. The solids of $MgSO_4$ were removed by filtration. Excess solvent was removed under vacuum. The residue was subjected to reverse phase HPLC (30 to 60% acetonitrile in water, 5 mL/min, 30 min). Excess solvent was removed by lyophilization.

(5-Chloro-2-{[(4-cyano-3-nitrobenzyl)amino]carbonyl}phenoxy)acetic Acid. The product was dissolved in 7.5 mL of ethanol and treated with 1.2 mL of 2.0 M NaOH. The solution was stirred at room temperature for 4 h; 1.0 M HCl was added until the pH of the solution reached 7. Most solvent was removed under vacuum. The concentrated solution of product is adjusted to pH 1, and a yellow precipitate appeared, which was dissolved and extracted with EtOAc. The organic phase was combined and washed with saturated NaCl and dried over $MgSO_4$. The solids of $MgSO_4$ were removed by filtration. Excess solvent was removed under vacuum. The residue was further subjected to reverse phase HPLC (40 to 50% acetonitrile in water, 10 mL/min, 60 min). Excess solvent was removed by lyophilization. The final product was a white powder and was characterized by NMR (9.3 ppm, t, 1 H; 8.4 ppm, s, 1 H; 8.1 ppm, d, 1 H; 7.9 ppm, d, 1 H; 7.8 ppm, d, 1 H; 7.3 ppm, s, 1 H; 7.2 ppm, d, 1 H; 4.9 ppm, s, 2 H; 4.7 ppm, d, 2 H) and MS [m/z 390.0 (ES^+), 389.7 g/mol (calculated)].

Protein Expression and Purification. Expression and purification were conducted as described previously.[6] The gene for WT *h*ALR2 was obtained from A. Podjarny in the pET-15b expression vector (Novagen). Mutations were performed using the Quikchange mutagenesis kits (Stratagene) with polymerase chain reaction primers purchased from Elim Biopharmaceuticals, Inc. All plasmids were transformed into *Escherichia coli* strain BL21(DE3) (Novagen). Cells were incubated at 37 °C for 3 h. Expression was then induced by addition of IPTG (1 mM), followed by growth for an additional 4 h. The cells were then pelleted, flash-frozen, and stored at −20 °C. To purify the protein, the cell lysate was resuspended and sonicated. After the cell debris had been removed, the lysate was applied to a Ni-NTA agarose column (Qiagen), and His-tagged protein was eluted using an imidazole gradient. The His tag was cut off using thrombin from bovine plasma (Aldrich). The cleaved product was loaded onto a 5 mL Hitrap Q HP anion exchange column (GE Healthcare) and purified by fast protein liquid chromatography (FPLC) with gradient elution (0 to 120 mM NaCl, 5 mL/min, 12 min). The purity of the final cleaved product was confirmed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis and mass spectrometry.

Kinetic Measurements of *h*ALR2. Kinetic experiments were performed using DL-glyceraldehyde as a substrate and NADPH as a cofactor in 20 mM Hepes buffer (pH 7) at room temperature as previously described.[6] Reaction rates were measured by following the decrease in NADPH adsorption at 340 nm over 5 min. The data were fit to the Michaelis−Menten equation to determine k_{cat}. Inhibition curves were determined at a saturating substrate concentration. The slope of the decrease in NADPH absorption was divided by blank and fit to eq 2 to yield the apparent inhibition constant

$$\frac{v_i}{v_0} = \frac{E - I - K_I + \sqrt{(E - I - K_I)^2 + 4EK_I}}{2E} \quad (2)$$

where E is the enzyme concentration, I is the inhibitor concentration, and K_I is the apparent inhibition constant.[25,26] All data were averaged from two parallel experiments.

X-ray Crystallography. The hanging drop method was used to obtain crystals (EasyXtal Tool 24 culture plates,

Table 1. Measured Kinetic Constants, Nitrile Vibrational Frequencies, and Full Widths at Half-Maximum (fwhm) of Inhibitor **2** in the Specificity Pocket of WT and Mutant *h*ALR2 at Room Temperature

sample	k_{cat} (s^{-1})	K_I (nM)	frequency (cm^{-1})		fwhm (cm^{-1})	
buffer	–	–	2242.2 ± 0.1		10.6 ± 0.1	
T113A	0.20 ± 0.01	30 ± 1	2237.8 ± 0.1		5.8 ± 0.1	
WT	0.33 ± 0.01	107 ± 7	2238.2 ± 0.1	2252.3 ± 0.1	11.4 ± 0.4	9.5 ± 0.2
T113S	0.19 ± 0.01	32 ± 1	2238.8 ± 0.1	2247.4 ± 0.3	6.4 ± 0.1	8.2 ± 1.1

Qiagen). WT *h*ALR2 was cocrystallized with the oxidized form of coenzyme NADP$^+$ (Sigma) and inhibitor **2** at room temperature (1:2:2 protein:inhibitor:coenzyme ratio). Hanging drops were made by the mixing mother liquor [0.06 M citric acid, 0.04 M Bis-tris propane, and 12% (w/v) PEG 3350 (pH 4.2)] with a holoenzyme solution (10 mg/mL, 20 mM Hepes, pH 7.0). Crystals were observed after equilibration for 3 days. Cryofreezing was conducted via quick transfer into a stabilization solution (25% PEG 3350), then into a cryoprotecting solution (55% PEG 3350), and finally into liquid nitrogen. Experiments were conducted at X-ray beamline 9-2 at the Stanford Synchrotron Radiation Light source (SSRL). Data sets were processed with HKL2000 and PHENIX. Data collection and refinement statistics are provided as Supporting Information (Table S1). The Protein Data Bank entry is 3T42.

IR Measurements of *h*ALR2. IR spectra were recorded with a Bruker Vertex 70 FT-IR spectrometer at both room temperature and low temperatures. Samples were illuminated through a notch filter that transmitted 2000−2500 cm^{-1} light, and spectra were recorded with a liquid nitrogen-cooled indium antimonide detector at 1 cm^{-1} resolution. For room-temperature measurements, the sample cell was made of sapphire windows separated by 75 and 100 μm Teflon spacers. Solutions of holoenzyme (1:1:2 protein:**2**:NADP$^+$ ratio) were equilibrated for 1 h at 4 °C, flushed with an excess of 20 mM Hepes buffer (pH 7), and concentrated to a final concentration of 2 mM. Variable-temperature IR spectra were recorded using a temperature-controlled water-jacketed cell from Lake Shore Cryotronics, Inc. Ethylene glycol was added to the water bath to push temperatures below 0 °C. The samples were prepared by dissolving lyophilized powder of a 4 mM holoenzyme solution in an equal volume of a 40% (w/w) trehalose/water mixture. The use of trehalose ensures a fluid state below 0 °C, and this process also forms an optical-quality glass at cryogenic temperatures. Equilibration for 5 min at each temperature was allowed before data were collected. To obtain a spectrum at 77 K, the sample cell was made of sapphire windows separated by 30 and 60 μm Teflon spacers. The cell was filled with a solution of 4 mM WT *h*ALR2 holoenzyme or a mixture of 4 mM WT *h*ALR2 and 4 mM T113A holoenzyme in 40% trehalose buffer and immediately immersed in liquid nitrogen in a custom cryostat.[6] All spectra were averaged on three parallel measurements and fit using one or two Gaussian functions with Matlab.

Stark Spectrum of Inhibitor 2. The sample cell was made of sapphire windows coated with 4.5 nm of Ni separated by a pair of 30 μm spacers and attached by copper wires to a high-voltage DC power supply (Trek Instruments Inc.).[6] The cell was filled with a solution of 50 mM inhibitor **2** in 2-methyltetrahydrofuran (MeTHF) and immediately immersed in liquid nitrogen in a custom liquid nitrogen cryostat. Applied fields of 0.7−1.6 MV/cm were synchronized with a home-built control unit; 128 scans with the field applied were alternated with 128 scans with no applied field, and the field-on-minus-field-off difference spectrum was obtained. The spectra were fit as described previously.[5,6,21] The Stark tuning rate of the nitrile probe on **2** reported is the average of three separate experiments.

Molecular Dynamics Simulations. The crystallographic structure of the *h*ALR2−**2**−NADP$^+$ complex was used in MD simulations. Amino acid mutations were created using Pymol.[27] The calculation strategy used here follows closely the approach developed in our lab for the ribonuclease S system containing nitrile probes.[9] Molecular dynamics simulations were conducted in Gromacs 3.3.1 using the AMBER-'99 force field.[28] Inhibitor **2** and the cofactor were parametrized using the Antechamber and Leap package with GAFF atom type and AM1BCC charges.[29] More details about the parametrization of **2** can be found in the Supporting Information (Figure S1 and Table S3). The protein was solvated with explicit SPC water molecules.[30] Nonbonded cutoffs of 10 Å were used. Simulations were conducted with periodic boundary conditions, and the calculation of long-range electrostatics was treated with the particle mesh Ewald model. Simulations were equilibrated (20 ps energy minimization, followed by 20 ps of heavy-atom position restrained refinement) and run using the Nose-Hoover thermostat and the Parinello-Raman barostat. Replica exchange molecular dynamics (REMD, 4 ns) were performed with 24 replicas from 298 to 339.45 K, ordered in temperature with a spacing of 1.75 K per replica. Electric fields at the midpoint of the nitrile bond were calculated at every 2 fs step for all MD trajectories. Atomic positions were saved at every 2 ps step only for the 298 K trajectories. Therefore, in the 4 ns 298 K trajectories, 2 × 10^6 electric field data were recorded and 2 × 10^3 structures were saved. Structures were taken every 60 ps from the 298 K trajectories of the REMD simulation over the last 3 ns, generating 50 unique starting structures to run 50 independent parallel 1 ns simulations (50 × 1 ns). In the 50 × 1 ns simulation, 25 × 10^6 electric field data were recorded; structures were not saved because of limited hard drive space. Electric field histograms were plotted using the 50 × 1 ns simulation data set (the vertical axis represents the percentage of electric field counts within a bin of 1 MV/cm, based on all 25 × 10^6 electric field data) and fit using one or two Gaussian functions. Correlation plots between electric field data and geometrical parameters were based on 4 ns 298 K trajectories from the REMD simulations, with 2 × 10^3 data points, each representing one electric field value and the corresponding structural data at that time point.

■ RESULTS AND DISCUSSION

Catalysis and Structure. All proteins studied in this work displayed normal catalytic activity and strong binding affinity toward inhibitor **2** (Table 1, columns 2 and 3). For example, WT *h*ALR2 showed a k_{cat} of 0.33 s^{-1}, close to the result reported previously.[6,14] **2** bound to WT *h*ALR2 with an apparent inhibition constant (K_I) of 107 nM, compared to a value of 2 μM for binding of inhibitor **1** to WT *h*ALR2;[6] the binding affinity has been increased by approximately 20-fold.

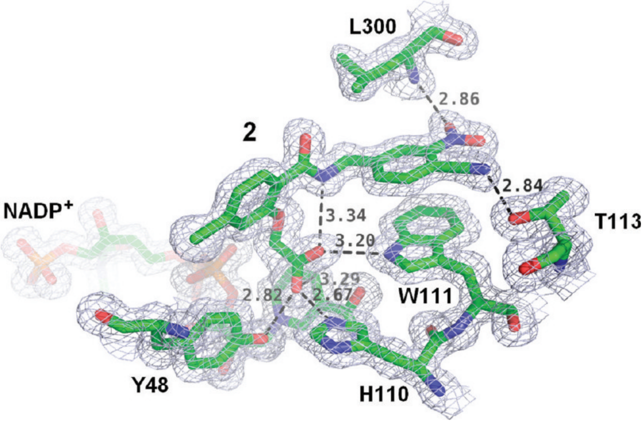


Figure 2. Electron density map for the inhibitor binding pocket ($2F_{obs} - F_{calc}$ contoured at 1σ in blue mesh), identifying relevant residues involved in inhibitor binding.

The fact that **2** binds strongly to WT and mutant *h*ALR2 ensures the feasibility of using this inhibitor to carry the probe into the protein active site for IR experiments and for X-ray crystallography. [We attempted to cocrystallize WT *h*ALR2 with inhibitor **1** (Figure 1). Only crystals of the *h*ALR2−NADP$^+$ complex were obtained, likely reflecting the much poorer binding of **1** than **2** at the active site. Likewise, soaking native crystals of the *h*ALR2−NADP$^+$ complex in the presence of **1** failed.]

The *h*ALR2 holoenzyme folds as an eight-stranded α/β-barrel structure (Figure 1). The active site is a well-resolved region located at the C-terminal loop (Figure 2), and **2** binds similarly to previously studied *h*ALR2−IDD inhibitor−NADP$^+$ complexes.[11,12] The carboxylate moiety of **2** is tightly anchored via hydrogen bond interactions with Y48, H110, and W111, as well as electrostatic interaction with cofactor NADP$^+$. There is also hydrophobic π-stacking interaction between the amine moiety of **2** and the W111 side chain. The nitro functionality forms a hydrogen bond with the L300 backbone amide, which accounts for the enhanced binding affinity of **2**, consistent with earlier work on nitro-substituted IDD inhibitors.[24] The nitrile probe is buried in the hydrophobic interior and well protected from close-range interaction with solvent molecules. A short contact exists between the nitrile group and the T113 side chain (N···O distance of 2.84 Å), which is a common binding mode when an inhibitor opens the specificity pocket of *h*ALR2 (e.g., similar effects are seen for IDD594, zopolrestat and IDD388[11,18,31]). For residue T113, no sign of multiple-side chain occupancy was observed in the electron density map.

IR Spectroscopy. The vibrational Stark spectrum of inhibitor **2** is shown in Figure 3A. The magnitude of the Stark tuning rate of the nitrile probe on **2** is found to be 0.53 $\mathrm{cm^{-1}/(MV/cm)}$, which is somewhat smaller than what has been observed previously for nitrile groups on phenyl rings (e.g., inhibitor **1** has a Stark tuning rate of 0.77 $\mathrm{cm^{-1}/(MV/cm)}$).[5,6] This is likely due to the electron withdrawing property of the adjacent nitro group. Detailed analysis of the Stark spectrum can be found in the Supporting Information.

FTIR spectra of **2** bound to WT and mutant *h*ALR2 were recorded to obtain information about electrostatic fields and/or hydrogen bond interactions in the vicinity of the nitrile probe. As seen in Figure 3B and Table 1 (columns 4 and 5), when **2** is bound to WT *h*ALR2, two peaks separated by ~14 $\mathrm{cm^{-1}}$ were observed in the nitrile region at room temperature. On the basis of the short contact observed in the crystallographic structure with the hydroxyl group of T113, it seems likely that the higher-energy, more intense band is due to the hydrogen-bonded form. To test this idea, two single-point mutations were examined, also shown in Figure 3B. T113A, in which the side chain hydroxyl group is removed, displays a single symmetric peak whose frequency is close to the lower-frequency peak of WT. T113S, which also has a side chain hydroxyl group, displays a major peak close to mutant T113A, with a small but reproducible shoulder on the high-energy side. These results are consistent with a well-documented result for simple nitriles in hydrogen-bonding and non-hydrogen-bonding solvents: the nitrile frequency typically shifts to higher wavenumbers upon accepting hydrogen bonds from hydroxylic solvent molecules.[32−45] Qualitatively, Purcell and Drago[32] have argued that when nitriles form a complex with Lewis acids (including a hydrogen bond), the electron lone pair is alleviated from the nitrogen nonbonding orbital that overlaps with nitrile bonding orbitals, resulting in an increase in the nitrile force constant and a blue shift in the nitrile stretch frequency.

To improve our understanding of the origin of the two peaks of WT *h*ALR2, we obtained IR spectra as a function of temperature (Figure 3C and Table S2 of the Supporting Information). The high-frequency (hydrogen-bonded) peak shifts to higher frequencies and narrows with a decrease in temperature. The blue shift of the hydrogen-bonded peak could be explained by the strengthening of hydrogen bonds at low temperatures,[9] with an accompanying decrease in bandwidth due to an increasingly more restrained system. The low-frequency (non-hydrogen-bonded) peak becomes broader and more asymmetric as the temperature is lowered, possibly suggesting that there may be more than one non-hydrogen-bonded configuration (the center frequency of the non-hydrogen-bonded peak first shifts slightly to the blue and then to the red). The area ratio of the high-frequency peak to the low-frequency peak increases as the temperature is lowered (the total integrated area increases by ~1.3-fold), and at 77 K, essentially only a single high-frequency peak is observed. This suggests that the hydrogen-bonded form is more favorable at low temperatures. We also measured the IR spectrum for mutant T113A at the same sample concentration at 77 K: a single peak, which is 17 $\mathrm{cm^{-1}}$ lower in frequency than the WT peak, is observed. This suggests that in the WT spectrum, the high-frequency peak at 77 K should be attributed to the hydrogen-bonded form. The ratio of the area of the nitrile peak of WT to that of T113A at 77 K is 1.7 ± 0.1. It is reasonable to assume that most of the population is in the hydrogen-bonded form at 77 K for **2** bound to WT *h*ALR2, so we can obtain an approximate value of the ratio of the integrated peak area for hydrogen-bonded and non-hydrogen-bonded nitrile species at 77 K. The result is consistent with previous reports of simple model compounds: the integrated intensities for the hydrogen-bonded forms are higher than those of the non-hydrogen-bonded forms. For example, the −CN stretch for CH_3CN has twice the integrated intensity in H_2O that it does in the pure solvent,[33] and for HCN, the CN stretching intensity increases by a factor of 1.4 upon formation of a complex with HF.[34] Assuming that the ratio of intensities does not change drastically with temperature, the relative populations of the hydrogen-bonded and non-hydrogen-bonded forms of **2** in WT *h*ALR2 at room temperature can be estimated from IR absorption. The ratio of the area of the two peaks at room temperature is 1.9 ± 0.3 [the value used here was obtained in

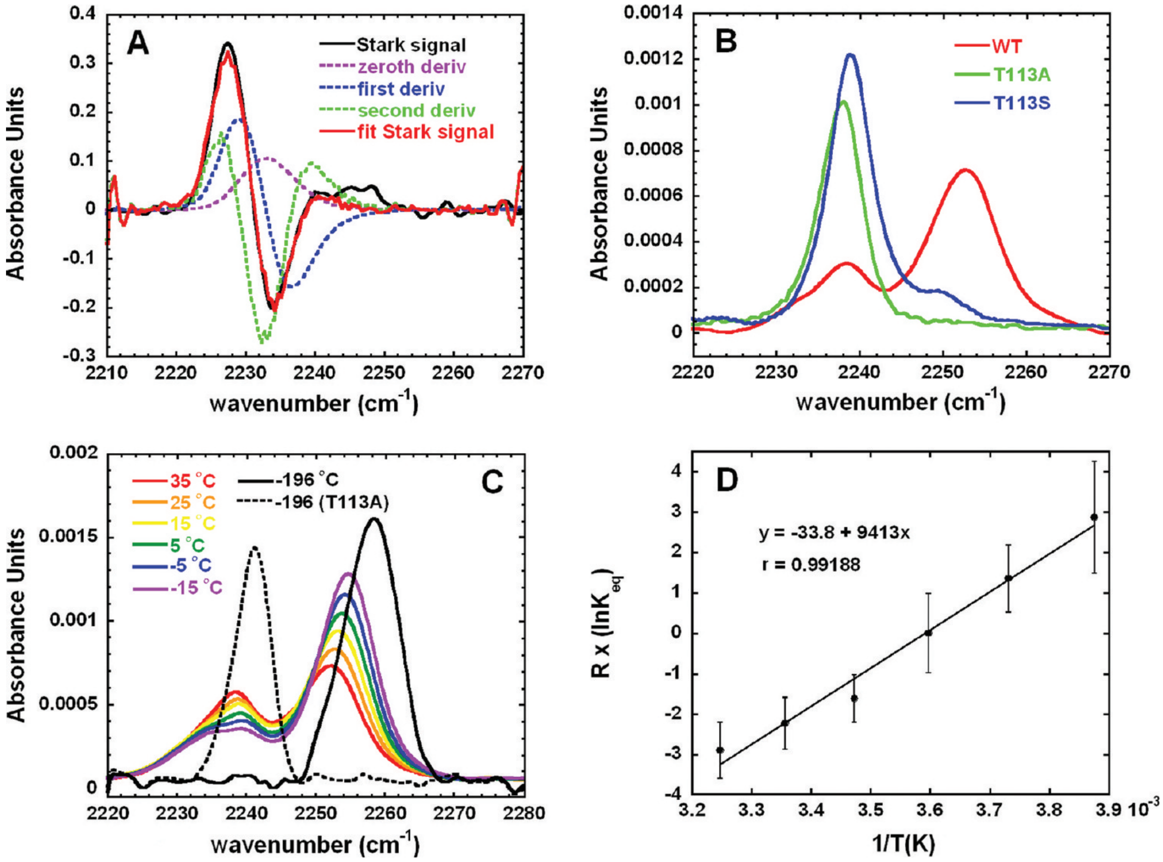


Figure 3. (A) Vibrational Stark spectrum of inhibitor **2** dissolved in MeTHF at 77 K. The black curve is the difference spectrum (field on minus field off) scaled to sample an absorbance of 1 and an applied field strength of 1.0 MV/cm. The purple, blue, and green curves represent contributions from the zeroth, first, and second derivatives, respectively, of the vibrational absorption spectrum of **2**. The red curve is the overall fit, and the second-derivative contribution gives a $|\Delta\vec{\mu}_{CN}|$ of 0.53 cm^{-1}/(MV/cm). More detailed analysis of the Stark spectrum can be found in the Supporting Information. (B) Vibrational absorption spectra in the nitrile region of **2** in the active sites of WT and two mutants of *h*ALR2 at room temperature. (C) Temperature dependence of IR absorption spectra of **2** bound to WT *h*ALR2 in 40% trehalose buffer and compared with the low-temperature absorbance of **2** bound to the T113A mutant. (D) van't Hoff plot derived from variable-temperature IR spectra.

20 mM Hepes buffer; the use of trehalose changed the area ratio at 25 °C slightly (see Table S2 of the Supporting Information)], from which we estimate a ratio of the two species of 1.1, which can be used later for comparisons with room-temperature MD simulations. We can calculate the equilibrium constant, K_{eq}, of the equilibrium

$$\text{non-hydrogen-bonded CN species} \leftrightharpoons \text{hydrogen-bonded CN species}$$

at various temperatures. Fitting the temperature dependence of K_{eq} to the van't Hoff equation

$$R \times \ln K_{eq} = -\frac{\Delta H^{\ominus}}{T} + \Delta S^{\ominus} \tag{3}$$

where R is the molar gas constant, $\Delta H^{\ominus}$ is the standard enthalpy change, and $\Delta S^{\ominus}$ is the standard entropy change, we can extract thermodynamic parameters for the reaction shown above. A van't Hoff plot is shown in Figure 3D, giving a $\Delta H^{\ominus}$ of −9.4 kJ mol^{-1} and a $\Delta S^{\ominus}$ of −34 J mol^{-1} K^{-1}. These results will be discussed later. In addition, the equilibrium constant at 77 K can be calculated using $\Delta H^{\ominus}$ and $\Delta S^{\ominus}$ to be ~4 × 104, consistent with the fact that at cryogenic temperature only a single high-frequency peak is observed.

MD Simulations and Electrostatics of the Wild Type.

To further investigate the molecular origins of these spectra, we have performed classical MD simulations. Classical MD applies Newton's laws of motion to a group of interacting atoms using a variety of force fields. In hydrogen bond interactions, both electrostatic interactions (dominant at long range) and exchange repulsion (short range) are important components. At the outset, we attempt to employ a very simple model to calculate IR absorption spectral shifts using the Stark tuning rate as the bridge between calculated changes in electric fields and IR shifts: we extract information about electric fields from the MD output and from this obtain information about the structures, dynamics, and populations that relate to room-temperature IR results. In the particular case of hydrogen bonds to nitriles, quantum effects are generally considered to be more significant than electrostatics.[32−45] As mentioned earlier, Purcell and Drago explained the observed blue shifts of nitrile frequency upon formation of hydrogen bonds by loss of an electron lone pair from the nonbonding orbital of the nitrile triple bond on the basis of molecular orbital calculations.[32] Reimers and Hall noted that when there is strong and directional solvent−solute interaction, such as a hydrogen bond or proton transfer, this interaction will be of primary importance, and the nitrile frequency becomes less sensitive to its electrostatic environments.[35] Therefore, we note at the

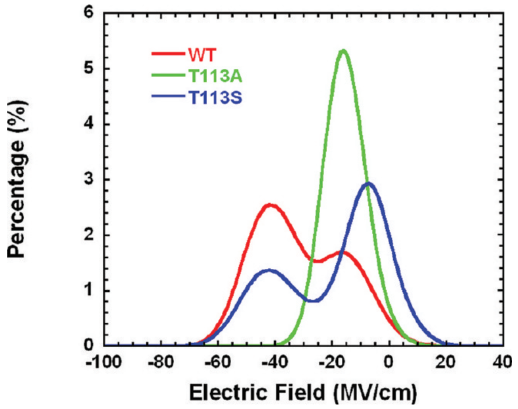


Figure 4. Calculated electric field histogram projected on the nitrile probe in WT and mutant *h*ALR2 bound to **2** extracted from MD simulations and using eq 1 (cf. Figure 3B).

Table 2. Calculated Peak Positions and fwhm Values for the Electric Field Distribution along the Nitrile Probe in WT and Mutant *h*ALR2. Measured and Predicted Changes in Electric Fields (Non-hydrogen-bonded Peaks vs Mutant T113A) and fwhm Values (MV/cm)[a]

	MD			experiment[a]	
	position	$\Delta\vec{F}_{\|\|}^{\text{protein}}$	fwhm	$\Delta\vec{F}_{\|\|}^{\text{protein}}$	fwhm
WT	−15.6	0.5	23.6	0.8	21.5
	−41.9	–	22.1	–	–
T113A	−16.1	–	15.2	–	10.9
T113S	−7.4	8.7	20.8	1.9	12.1
	−41.5	–	24.5	–	–

[a] The measured $\Delta\vec{F}_{||}^{\text{protein}}$ and fwhm are calculated using the observed spectra and the Stark tuning rate of **2** in MeTHF ($|\Delta\vec{\mu}_{\text{CN}}|$ = 0.53 cm^{-1}/(MV/cm).

outset that modeling the results as entirely electrostatic in origin will incorrectly assign the spectral shifts of the hydrogen-bonded and non-hydrogen-bonded populations. However, this anticipated discrepancy does not imply that the force field used in the simulations is wrong, because the energies of a nitrile in MD are not determined solely by electrostatics. Especially, the Lennard-Jones parameters in our simulations were calibrated by Corcelli et al. against a DFT calculation of the interaction between acetonitrile and water in a linear hydrogen bonding geometry for a more accurate description of the hydrogen bond interaction between the lone pair of the nitrogen atom and water.[44] The validity of the structural and population information then largely depends on the transferability of these parameters into a chemically different system (different donor, acceptor, and hydrogen bond configurations), which can be tested by comparing simulated spectral features with experimental results. In our case, the most direct linkage between them is the population information. That is, if the populations that emerge from our analysis are similar to those observed experimentally, this would suggest that the Lennard-Jones parameters are valid so that the energetics of a nitrile falls into a reasonable regime.

Classical MD simulations on the nanosecond time scale were performed, and electric fields at every MD step were recorded giving a histogram of protein electric fields projected along the nitrile probe, as shown in Figure 4 and Table 2 (columns 2 and 4). MD predicts a double-peak field distribution for WT *h*ALR2, consistent with the two peaks observed in the room-temperature IR spectrum (Figure 3B). For T113A, MD predicts a single peak, again consistent with the single peak observed for T113A. Although MD will not identify the two peaks in the WT spectrum, we can deduce their identity through the comparison with T113A: the single peak of T113A is close to the more positive field peak of WT *h*ALR2; therefore, this peak is the non-hydrogen-bonded peak, while the more negative field peak is the hydrogen-bonded peak. More quantitatively, the MD simulations predict that in WT *h*ALR2, the population ratio of hydrogen-bonded to non-hydrogen-bonded forms is 1.5, which is quite close to the experimental value estimated above to be 1.1. As for T113S, MD predicts a double-peak field distribution, and the area ratio of the more positive peak is larger than that of WT *h*ALR2.

With the basic confirmation that the simulations provide reasonable agreement for the populations and fields, we can exploit the MD trajectories to further investigate the microscopic interactions between individual atoms that lead to the field distributions in a given protein, something that is not possible in the experiments. Specifically, we extracted structural information from the MD output and used this to establish correlations between computed electric fields and geometric parameters. Figure 5A shows the dependence of the electric field projected on the inhibitor −CN bond on the dihedral angle representing T113 side chain rotation for WT *h*ALR2. The relatively constant angle values indicate that the T113 side chain conformation does not change much during course of the simulation. As seen in Figure 5B, however, the two-distance correlation (the distance between the T113 hydroxyl hydrogen atom and the nitrogen atom of the probe and the distance between the hydrogen atom and the oxygen atom of the T113 carbonyl group) displays two-state behavior. To explain this, we propose a model of a hydrogen bond motif involving one donor and two acceptors. This is based on the observation of a short contact between the nitrile probe and T113 hydroxyl group (as seen in the low-temperature X-ray structure), and also on the fact that threonine [serine also (see below)] is frequently found to form an intraresidue hydrogen bond.[11,46] Dynamically, the hydroxyl group switches between hydrogen bonds to the nitrile probe and the amide carbonyl group of T113. Panels C and D of Figure 5 show the correlation between the computed electric fields and the angle and distance parameters of the hydrogen bond between the T113 hydroxyl group and the nitrile probe. In the field−angle correlation, the two-state behavior is seen again. All of the electric field data are sorted in Figure 5E by using geometrical criteria for hydrogen bond formation [angle of ≤60° and distance of ≤3.5 Å (see the legend and panel F of Figure 5, which shows snapshots)]. This results in two distinguished peaks. The more negative field peak corresponds to hydrogen-bonded nitrile species and the other to non-hydrogen-bonded species. This is consistent with the observation that the single peak of T113A on the electric field histogram is close to the more positive field peak of WT *h*ALR2. As expected, the electrostatic model we employed predicts a red shift in the frequency of the nitrile probe upon hydrogen bond formation, failing to capture the blue shift observed in IR experiments due to the dominating quantum effects in nitrile hydrogen bond formation. Finally, the calculated lifetimes for the two species are both on the picosecond time scale (9.7 ps for the hydrogen-bonded form

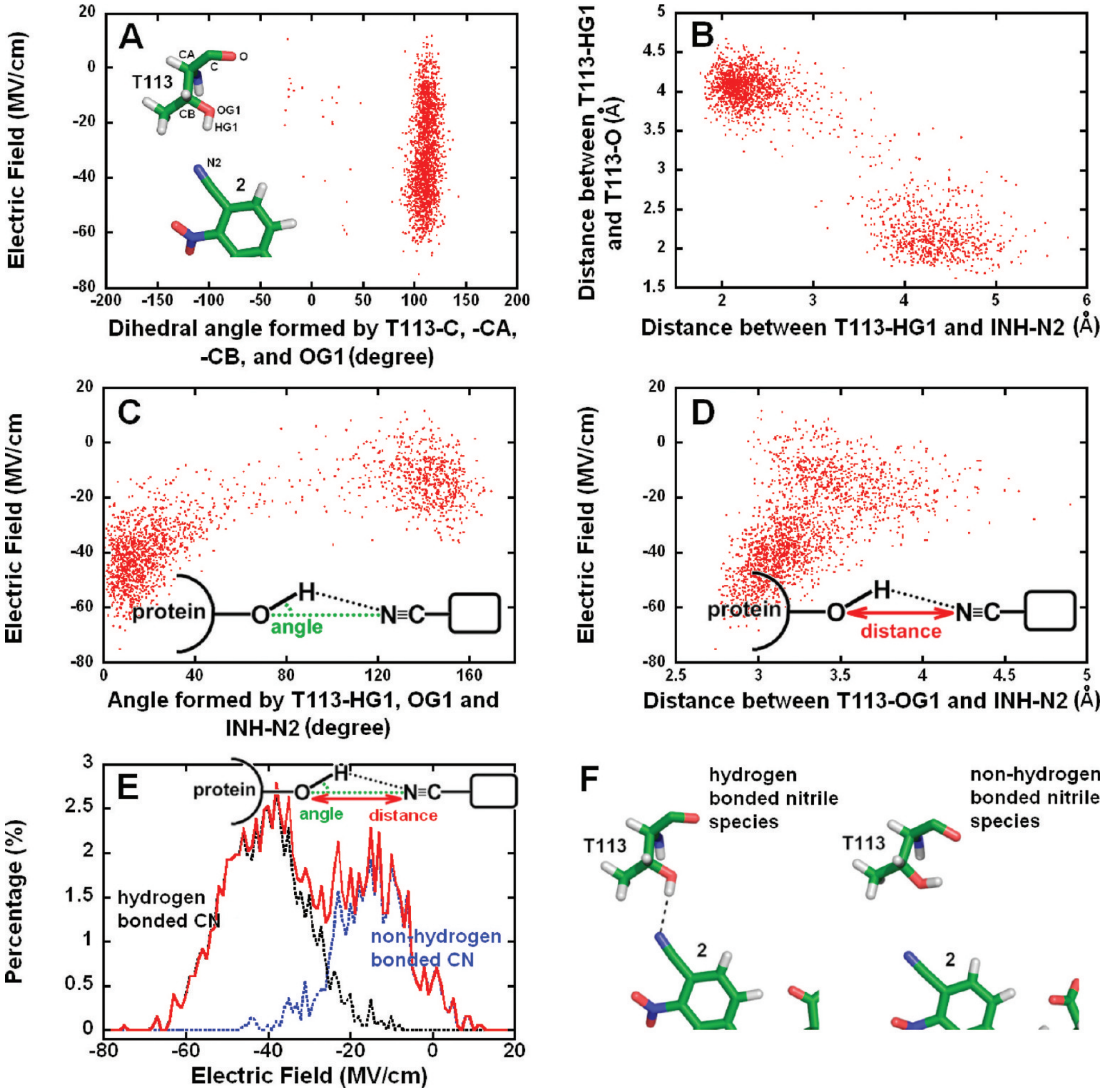


Figure 5. MD trajectory analysis for WT *h*ALR2. (A) Correlation of the calculated electric field projected on the inhibitor −CN bond with the dihedral angle formed by the C, CA, CB, and OG1 atoms of T113. Each data point represents one electric field value and the corresponding structural data at that time point. (B) Correlation of two distances in the MD simulation: the distance between T113 HG1 and INH N2 and the distance between T113 HG1 and T113 O (see the inset of panel A for labeling). (C) Correlation of the calculated electric field projected on the inhibitor −CN bond and the angle formed by the HG1 and OG1 atoms of T113 and the N2 atom of INH. (D) Correlation of calculated electric field projected on the inhibitor −CN bond and distance between T113 OG1 and INH N2. (E) Calculated electric fields projected on the inhibitor −CN bond sorted according to hydrogen bond geometrical criteria. The red solid curve represents the whole data set (2001 data points from 4 ns REMD); the black dashed curve represents configurations in which the inhibitor nitrile is hydrogen-bonded while the carbonyl group of T113 is not hydrogen-bonded (61%, angle of ≤60°, distance of ≤3.5 Å), and the blue dashed curve represents configurations in which the nitrile probe is not hydrogen-bonded while the T113 carbonyl group is hydrogen-bonded (39%). (F) Detailed representation of the T113−OH···**2**−CN species (left) and T113−OH···T113−C=O species (right).

and 6.2 ps for the non-hydrogen-bonded form).[47] In summary, by analyzing the MD trajectories for WT *h*ALR2, we propose that the source of the two peaks in the protein electric field distribution projected on the probe −CN bond is the dynamical making and breaking of a hydrogen bond between the nitrile probe and T113 side chain hydroxyl group, which viewed from the nitrile IR band gives rise to the two peaks seen in the IR spectrum, and further that the T113 hydroxyl group is always hydrogen-bonded, either to the inhibitor nitrile group or to its amide carbonyl group.

Following this analysis, the thermodynamic parameters obtained earlier actually correspond, in the context of the entire protein−inhibitor complex, to the equilibrium between two types of hydrogen bonds, a hydrogen bond between T113 and its amide carbonyl group and one between T113 and the inhibitor nitrile:

$$\text{T113−OH}\cdots\text{T113−C=O species}$$
$$\leftrightharpoons \text{T113−OH}\cdots\mathbf{2}\text{−CN species}$$

From other studies, hydrogen bond energies are typically

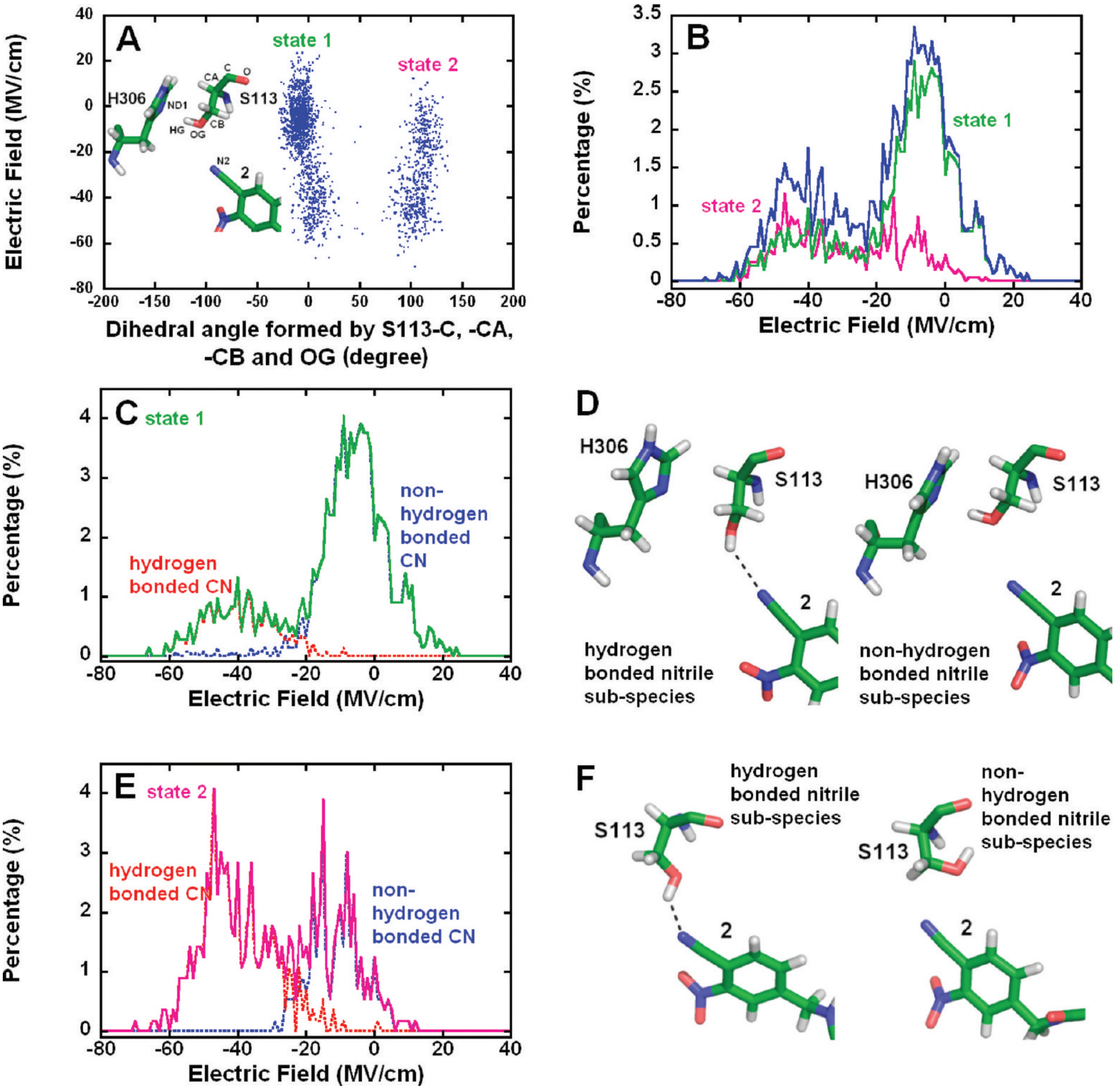


Figure 6. MD trajectory analysis of *h*ALR2 T113S. (A) Correlation of electric fields with the dihedral angle formed by the C, CA, CB, and OG atoms of S113. (B) Electric fields sorted according to serine side chain conformation. The blue curve represents the full data set (2001 data points). The green curve (72%) represents the data set for state 1. The pink curve (28%) represents the data set for state 2. (C) Electric fields in state 1 sorted according to the hydrogen bond formation criteria. The green solid curve represents the entire data set (1436 data points); the red dashed curve represents subconfigurations in which the inhibitor nitrile is hydrogen-bonded while the imidazole nitrogen atom of H306 is not hydrogen-bonded (23%, angle of ≤60°, distance of ≤3.5 Å), and the blue dashed curve represents subconfigurations in which the nitrile probe is not hydrogen-bonded while the imidazole nitrogen atom of H306 is hydrogen-bonded (77%). (D) S113−OH···**2**−CN subspecies in state 1 (left) and S113−OH···H306−imidazole subspecies (right). (E) Electric fields in state 2 sorted according to the hydrogen bond formation criteria. The peak solid curve represents the full data set (565 data points). The red dashed curve represents subconfigurations in which the inhibitor nitrile is hydrogen-bonded while the carbonyl group of S113 is not hydrogen-bonded (63%, angle of ≤60°, distance of ≤3.5 Å). The blue dashed curve represents subconfigurations in which the nitrile probe is not hydrogen-bonded while the S113 carbonyl group is hydrogen-bonded (37%). (F) S113−OH···**2**−CN subspecies in state 2 (left) and S113−OH···S113−C═O subspecies (right).

approximately 10−30 kJ mol^{-1}.[48−53] The enthalpy change obtained from the van't Hoff plot in Figure 3D for the difference between a hydrogen bond with the nitrile as the acceptor and one with the amide carbonyl as the acceptor ($\Delta H^{\Theta} = -9.4$ kJ mol^{-1}) is reasonable in terms of the order of magnitude. The negative sign indicates that the nitrile nitrogen atom is a better hydrogen bond acceptor than the amide oxygen atom in this case, although previous studies, e.g., ab initio calculations[50] and Abraham's empirical hydrogen bond basicity scale,[54] suggested the contrary. This discrepancy is likely due to an idiosyncratic behavior of this restrained system. The negative entropy change ($\Delta S^{\Theta} = -34$ J mol^{-1} K^{-1}) implies that the configuration on the left in Figure 5F is less disordered. It can be interpreted as forming a stronger hydrogen bond limits the movement of molecules, leading to a decrease in the system entropy.

MD Simulations and Electrostatics of T113S. For T113S, a similar approach was used, and the results suggest both similarities and differences between this mutant and WT *h*ALR2. We first examine the dependence of electric fields on the S113 side chain rotation. As seen in Figure 6A, the data set is composed of two populations, denoted state 1 and state 2. Compared with threonine in WT *h*ALR2, in which the methyl group largely blocks side chain rotation (Figure 5A), the serine side chain is more flexible in T113S, occupying two possible conformations, as illustrated in panels D (state 1) and F (state 2) of Figure 6. By sorting electric field data according to the dihedral angle parameter (Figure 6B), we find that the two populations contribute differently to the electric field distribution. The prediction of multiple side chain conformations for S113 is consistent with a recent paper on crystallographic studies of *h*ALR2 mutants bound with IDD inhibitors with halo substituents on phenyl para sites.[31] The two side chain conformations of S113 in our system predicted by MD simulations resemble the two side chain conformations of S113 detected in structures of T113S−IDD388 or −IDD594 complexes, and the predicted percentages of two configurations as depicted in panels D and F of Figure 6 are close to the measured fractional occupancies of the S113 side chain in crystal structures of T113S−IDD388 or −IDD594 complexes. (In our system, the predicted percentages of state 1 and state 2 are 72 and 28%, respectively. In the crystal structure of the T113S−IDD388 complex, the percentages of the two corresponding configurations are 72 and 28%, respectively, and in the T113S−IDD594 complex, they are 68% and 32%, respectively.[31]) In state 1, there is a hydrogen bond motif similar to that in WT *h*ALR2. In this case, the imidazole nitrogen atom in H306 acts as one hydrogen bond acceptor while the other acceptor is the nitrogen atom on the probe nitrile, and the hydrogen bond donated by the S113 side chain hydroxyl group switches between them dynamically. If we use the same hydrogen bond criteria (angle of ≤60° and distance of ≤3.5 Å) to sort electric field data for state 1 (Figure 6C), the peak at a more negative electric field corresponds to the S113−OH···**2**−CN subspecies; the other corresponds to the S113−OH···H306−imidazole subspecies. Snapshots for the two subspecies in state 1 are given in Figure 6D. In the case of state 2, there is also a similar hydrogen bond motif, in which the donor is the S113 side chain hydroxyl group, one acceptor is the nitrogen atom on the probe nitrile, and the other acceptor is the amide carbonyl oxygen atom of S113. Electric field data are sorted using the same criteria (angle of ≤60° and distance of ≤3.5 Å), and two separated peaks are seen again (Figure 6E). Snapshots for the S113−OH···**2**−CN species and S113−OH···S113−C═O species in state 2 are given in Figure 6F, which has a configuration similar to that of WT *h*ALR2 (cf. Figure 5F). Overall, this could explain the IR spectrum of **2** bound to T113S, a major peak with a high-frequency tail, attributed to hydrogen bonding, but more complicated than that of WT: unlike WT *h*ALR2, in T113S, the serine residue displays two possible side chain conformations, each interpreted as a hydrogen bond making and breaking process, and state 1, the dominant configuration, supports a higher ratio of non-hydrogen-bonded nitrile species.

Concluding Comments about Complications Caused by Hydrogen Bonds. As discussed earlier, many experiments reveal that the nitrile stretch frequency shifts to higher frequencies upon accepting a hydrogen bond. Using a purely electrostatic model, when the −CN bond interacts with a hydroxyl group in a configuration as in Figure 5F, the difference dipole moment of the nitrile bond $\Delta\vec{\mu}_{CN}$ (pointing from the nitrogen atom to carbon atom) is aligned along the electric field exerted by the hydroxyl dipole, and a red shift in energy is expected according to eq 1. As a result, this purely electrostatic model predicts a frequency shift in the opposite direction to what is observed when a hydrogen bond is formed. For the same reason, when there are hydrogen bond interactions between the nitrile probe and the surrounding environment, the idea of using the nitrile as a Stark probe to measure electrostatic fields in proteins solely on the basis of eq 1 is questionable: environment-induced frequency shifts may no longer respond linearly to electric fields but are more dominated by the quantum mechanical nature of such interactions, and these effects may depend upon details of the geometry of a particular hydrogen bond.[35,45,55] We recently presented a method to separate effects of hydrogen bonds on IR frequency shifts from the environmental electrostatic contribution in the case of −SCN probes by correlating the ^{13}C NMR chemical shift and the IR frequency.[8,56] This approach is relatively simple to apply to −SCN probes as ^{13}C can be directly introduced during labeling, but it would be more difficult to apply to a complex inhibitor such as **2** because of the synthetic effort, though we are currently developing a synthetic strategy that may allow facile preparation of ^{13}C-labeled versions of aromatic nitrile-containing inhibitors. One interesting application would be using ^{13}C-labeled inhibitor **1** to acquire tandem IR and NMR measurements in the systems studied in part 1 of this series, where an X-ray structure could not be obtained, so that we can detect whether the nitrile probe accepts hydrogen bonds from nearby amino acid residues in the absence of crystallographic structures. If a comparison is made within a series of mutants, pH changes or inhibitor changes, and their added electrostatic contribution is the goal, then a demonstration that a hydrogen bond is retained or lost by the NMR−IR correlation is very useful for vetting the data. It would also be interesting to measure $|\Delta\vec{\mu}_{CN}|$ for both hydrogen-bonded and non-hydrogen-bonded populations to determine whether the Stark tuning rate is affected by a hydrogen bond. It has been shown in one case that hydrogen-bonded nitrile probes have a Stark tuning rate similar to those of non-hydrogen-bonded ones.[56] We attempted to measure the vibrational Stark effect of inhibitor **2** in situ in WT *h*ALR2, but the relatively low extinction coefficient and limited solubility of *h*ALR2 in buffered solution precluded this measurement.

While the interpretation of frequency shifts of electric fields with calculated shifts may be compromised when the nitrile is hydrogen-bonded, in cases where the nitrile is not hydrogen-bonded, we can compare changes in electric fields $(\Delta\vec{F}_{\parallel}^{\text{protein}})$ measured by IR given the Stark tuning rate with predictions made by MD (Table 2, columns 3 and 5). It can be seen that the calculation strategy performs well for WT *h*ALR2 and T113S in terms of the direction, but not magnitude. Treating the line width as being due to inhomogeneous broadening from the distribution of local electric fields interacting with $\Delta\vec{\mu}_{CN}$, we found the simulation reproduces the observed trend for the three non-hydrogen-bonded peaks (Table 2, columns 4 and 6); i.e., T113A shows the narrowest fwhm, while WT *h*ALR2 shows the broadest fwhm. Thus, the basic computational strategy is qualitatively successful, but not quantitatively, likely because of errors in the force field and/or insufficient sampling in the MD simulations. There are limited examples in the literature demonstrating the improvement of force fields in the pursuit of precise assignment of structural features based on electric field calculations, but several important examples have been presented using the system of carbon monoxide photodissociation in myoglobin.[57−60] A genuine QM/MM simulation might be able to better capture the properties of the hydrogen-bonded populations. This system, which contains a hydrogen-bonded and non-hydrogen-bonded nitrile in the same macromolecular system, can serve as a test system for establishing benchmarks in advanced simulation methodology, and the independent predictions of the time scale for making and breaking hydrogen bonds can, in principle, be tested using two-dimensional IR techniques.[61] Such experiments are currently in progress.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

(1) Analysis of Stark spectroscopy data, (2) crystallographic data, (3) temperature-dependent IR data, and (4) paramerization of inhibitor **2**. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sboxer@stanford.edu. Phone: (650) 723-4482. Fax: (650) 723-4817.

Funding

This work was supported in part by a grant from the National Institutes of Health (GM27738) to S.G.B.

■ ACKNOWLEDGMENTS

We thank Drs. Aaron Fafarman and Lauren Webb for useful discussions about all aspects of this work. We thank Prof. Michael Van Zandt for useful discussions about inhibitor design. We also thank Dr. Sushant Malhotra for helpful discussions about the inhibitor synthesis pathway and Dr. Jorge Zuniga from the Brunger lab at Stanford for assisting us in setting up robot screening trays for obtaining the initial crystallization conditions. Portions of this research were conducted at the Stanford Synchrotron Radiation Light source, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

■ ABBREVIATIONS

*h*ALR2, human aldose reductase; VSE, vibrational Stark effect; IR, infrared; IDD, Institute for Diabetes Discovery; FCC, flash column chromatography; DMF, dimethylformamide; MeTHF, 2-methyltetrahydrofuran; MD, molecular dynamics.

■ REFERENCES

(1) Honig, B., and Nicholls, A. (1995) Classical Electrostatics in Biology and Chemistry. *Science 268*, 1144−1149.

(2) Warshel, A., and Papazyan, A. (1998) Electrostatic effects in macromolecules: fundamental concepts and practical modeling. *Curr. Opin. Struct. Biol. 8*, 211−217.

(3) Simonson, T. (2001) Macromolecular electrostatics: Continuum models and their growing pains. *Curr. Opin. Struct. Biol. 11*, 243−252.

(4) Park, E. S., Andrews, S. S., Hu, R. B., and Boxer, S. G. (1999) Vibrational Stark Spectroscopy in Proteins: A Probe and Calibration for Electrostatic Fields. *J. Phys. Chem. B 103*, 9813−9817.

(5) Suydam, I. T., Snow, C. D., Pande, V. S., and Boxer, S. G. (2006) Electric Fields at the Active Site of an Enzyme: Direct Comparison of Experiment with Theory. *Science 313*, 200−204.

(6) Webb, L. J., and Boxer, S. G. (2008) Electrostatic Fields Near the Active Site of Human Aldose Reductase: 1. New Inhibitors and Vibrational Stark Effect Measurements. *Biochemistry 47*, 1588−1598.

(7) Sigala, P. A., Fafarman, A. T., Bogard, P. E., Boxer, S. G., and Herschlag, D. (2007) Do Ligand Binding and Solvent Exclusion Alter the Electrostatic Character within the Oxyanion Hole of an Enzymatic Active Site. *J. Am. Chem. Soc. 129*, 12104−12105.

(8) Fafarman, A. T., Sigala, P. A., Herschlag, D., and Boxer, S. G. (2010) Decomposition of Vibrational Shifts of Nitriles into Electrostatic and Hydrogen-Bonding Effects. *J. Am. Chem. Soc. 132*, 12811−12813.

(9) Fafarman, A. T., and Boxer, S. G. (2010) Nitrile Bonds as Infrared Probes of Electrostatics in Ribonuclease S. *J. Phys. Chem. B 114*, 13536−13544.

(10) Podjarny, A., Cachau, R. E., Schneider, T., Van Zandt, M. C., and Joachimiak, A. (2004) Subatomic and Atomic Crystallographic Studies of Aldose Reductase: Implications for Inhibitor Binding. *Cell. Mol. Life Sci. 61*, 763−773.

(11) Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M. C., Sibley, E. O., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A., and Podjarny, A. (2004) Ultrahigh Resolution Drug Design. I: Details of Interactions in Human Aldose Reductase-Inhibitor Complex at 0.66 Å. *Proteins 55*, 792−804.

(12) Ruiz, F., Hazemann, I., Darmanin, C., Mitschler, A., Van Zandt, M., Joachimiak, A., El-Kabbani, O., and Podjarny, A. (2011) The Crystallographic Structure of Aldose Reductase IDD393 Complex Confirms Leu300 as a Specificity Determinant, manuscript to be submitted.

(13) Bohren, K. M., Page, J. L., Shankar, R., Henry, S. P., and Gabbay, K. H. (1991) Expression of Human Aldose and Aldehyde Reductases. *J. Biol. Chem. 266*, 24031−24037.

(14) Bohren, K.M., Grimshaw, C. E., Lai, C. J., Harrison, D. H., Ringe, D., Petsko, G. A., and Gabbay, K. H. (1994) Tyrosine-48 Is the Proton Donor and Histidine-110 Directs Substrate Stereochemical Selectivity in the Reduction Reaction of Human Aldose Reductase: Enzyme Kinetics and Crystal Structure of the Y48H Mutant Enzyme. *Biochemistry 33*, 2021−2032.

(15) Kador, P. F. (1988) The Role of Aldose Reductase in the Development of Diabetic Complications. *Med. Res. Rev. 8*, 325−352.

(16) Srivastava, S. K., Ramana, K. V., and Bhatnagar, A. (2005) Role of Aldose Reductase and Oxidative Damage in Diabetes and the Consequent Potential for Therapeutic Options. *Endocr. Rev. 26*, 380−392.

(17) Kaul, C. L., and Ramarao, P. (2001) The Role of Aldose Reductase Inhibitors in Diabetic Complications: Recent Trends. *Methods Find. Exp. Clin. Pharmacol. 23*, 465−475.

(18) Urzhumtsev, A., Tete-Favier, F., Mitschler, A., Barbanton, J., Barth, P., Urzhumtseva, L., Biellmann, J.-F., Podjarny, A., and Moras, D. (1997) A 'specificity' pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil. *Structure 5*, 601−612.

(19) Van Zandt, M. C., Sibley, E. O., McCann, E. E., Combs, K. J., Flam, B., Sawicki, D. R., Sabetta, A., Carrington, A., Sredy, J., Howard, E., Mitschler, A., and Podjarny, A. (2004) Design and synthesis of highly potent and selective (2-arylcarbamoyl-phenoxy)-acetic acid inhibitors of aldose reductase for treatment of chronic diabetic complications. *Bioorg. Med. Chem. 12*, 5661−5675.

(20) Steuber, H., Zentgraf, M., Podjarny, A., Heine, A., and Klebe, G. (2006) High-resolution Crystal Structure of Aldose Reductase Complexed with the Novel Sulfonyl-pyridazinone Inhibitor Exhibiting an Alternative Active Site Anchoring Group. *J. Mol. Biol. 356*, 45−56.

(21) Andrews, S. S., and Boxer, S. G. (2000) Vibrational Stark Effects of Nitriles I. Methods and Experimental Results. *J. Phys. Chem. A 104*, 11853−11863.

(22) Fafarman, A. T., Webb, L. J., Chuang, J. I., and Boxer, S. G. (2006) Site-Specific Conversion of Cysteine Thiols into Thiocyanate Creates an IR Probe for Electric Fields in Proteins. *J. Am. Chem. Soc. 128*, 13356−13357.

(23) Andrews, S. S., and Boxer, S. G. (2002) Vibrational Stark Effects of Nitriles II. Physical Origins of Stark Effects from Experiment and Perturbation Models. *J. Phys. Chem. A 106*, 469−477.

(24) Steuber, H., Heine, A., and Klebe, G. (2007) Structural and Thermodynamic Study on Aldose Reductase: Nitro-substituted Inhibitors with Strong Enthalpic Binding Contribution. *J. Mol. Biol. 368*, 618−638.

(25) Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., Van Dorsselaer, A., Podjarny, A., and Motas, D. (1999) Production of crystals of human aldose reductase with very high resolution diffraction. *Acta Cystallogr. 55*, 721−723.

(26) Kurono, M., Fujiwara, I., and Yoshida, K. (2001) Stereospecific interaction of a novel spirosuccinimide type aldose reductase inhibitor, AS-3201, with aldose reductase. *Biochemistry 40*, 8216−8226.

(27) DeLano, W. L. http://www.pymol.org.

(28) Sorin, E. J., and Pande, V. S. (2005) Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. *Biophys. J. 88*, 2472−2493.

(29) http://ambermd.org/antechamber/antechamber.html.

(30) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) Interaction models for water in relation to protein hydration. In *Intermolecular Forces* (Pullman, B., Ed.) p 331, Kluwer Academic Publishers, Dordrecht, The Netherlands.

(31) Koch, C., Heine, A., and Klebe, G. (2011) Tracing the Detail: How Mutations Affect Binding Modes and Thermodynamics Signatures of Closely Related Aldose Reductase Inhibitors. *J. Mol. Biol. 406*, 700−712.

(32) Purcell, K. F., and Drago, R. S. (1966) Studies of the Bonding in Acetonitrile Adducts. *J. Am. Chem. Soc. 88*, 919−924.

(33) Bertie, J. E., and Lan, Z. (1997) Liquid Water-Acetonitrile Mixtures at 25 °C: The Hydrogen-Bonded Structure Studied through Infrared Absolute Integrated Absorption Intensities. *J. Phys. Chem. B 101*, 4111−4119.

(34) Curtiss, L. A., and Pople, J. A. (1973) Molecular Orbital Calculation of Some Vibrational Properties of Ilse Complex between HCN and HF. *J. Mol. Spectrosc. 48*, 413−426.

(35) Reimers, J.R., and Hall, L. E. (1999) The Solvation of Acetonitrile. *J. Am. Chem. Soc. 121*, 3730−3744.

(36) Eaton, G., Pena-Nunez, A. S., Symons, M. C. R., Ferrario, M., and McDonald, I. R. (1988) Spectroscopic and Molecular Dynamics Studies of Solvation of Cyanomethane and Cyanide Ions. *Faraday Discuss. Chem. Soc. 85*, 237−253.

(37) Nyquist, R. A. (1990) Solvent-Induced Nitrile Frequency Shifts: Acetonitrile and Benzonitrile. *Appl. Spectrosc. 44*, 1405−1407.

(38) Fawcett, W. R., Liu, G., and Kessler, T. E. (1993) Solvent-Induced Frequency Shifts in the Infrared Spectrum of Acetonitrile in Organic Solvents. *J. Phys. Chem. 97*, 9293−9298.

(39) Jamroz, D., Stangret, J., and Lindgren, J. (1993) An Infrared Spectroscopic Study of the Preferential Solvation in Water-Acetonitrile Mixtures. *J. Am. Chem. Soc. 115*, 6165−6168.

(40) Getahun, Z., Huang, C., Wang, T., De Leon, B., DeGrado, W. F., and Gai, F. (2003) Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment. *J. Am. Chem. Soc. 125*, 405−411.

(41) Huang, C., Wang, T., and Gai, F. (2003) Temperature dependence of the CN stretching vibration of a nitrile-derivatized phenylalanine in water. *Chem. Phys. Lett. 371*, 731−738.

(42) Maienschein-Cline, M. G., and Londergan, C. H. (2007) The CN Stretching Band of Aliphatic Thiocyanate is Sensitive to Solvent Dynamics and Specific Solvation. *J. Phys. Chem. A 111*, 10020−10025.

(43) Aschaffenburg, D. J., and Moog, R. S. (2009) Probing Hydrogen Bonding Environments: Solvatochromic Effects on the CN Vibration of Benzonitrile. *J. Phys. Chem. B 113*, 12736−12743.

(44) Lindquist, B. A., Haws, R. T., and Corcelli, S. A. (2008) Optimized Quantum Mechanics/Molecular Mechanics Strategies for Nitrile Vibrational Probes: Acetonitrile and para-Tolunitrile in Water and Tetrahydrofuran. *J. Phys. Chem. B 112*, 13991−14001.

(45) Chaban, G. M. (2004) Anharmonic Vibrational Spectrospcopy of Nitriles and Their Complexes with Water. *J. Phys. Chem. A 108*, 4551−4556.

(46) Eswar, N., and Ramakrishnan, C. (2000) Deterministic features of side-chain main-chain hydrogen bonds in globular protein. *Protein Eng. 13*, 227−238.

(47) Kim, Y. S., and Hochstrasser, R. M. (2005) Chemical exchange 2D IR of hydrogen-bond making and breaking. *Proc. Natl. Acad. Sci. U.S.A. 102*, 11185−11190.

(48) Wofford, B. A., Eliades, M. E., Lieb, S. G., and Bevan, J. W. (1987) Determination of dissociation energies and thermal functions of hydrogen-bond formation using high resolution FTIR spectroscopy. *J. Chem. Phys. 87*, 5674−5680.

(49) Feyereisen, M. W., Feller, D., and Dixon, D. A. (1996) Hydrogen Bond Energy of the Water Dimmer. *J. Phys. Chem. 100*, 2993−2997.

(50) Rablen, R. P., Lockman, W. J., and Jorgensen, L. W. (1998) Ab Initio Study of Hydrogen-Bonded Complexes of Small Organic Molecules with Water. *J. Phys. Chem. A 102*, 3782−3797.

(51) Lantz, M. A., Jarvis, S. P., Tokumoto, H., Martynski, T., Kusumi, T., Nakamura, C., and Miyake, J. (1999) Streching the α-helix: A direct measure of the hydrogen-bond energy of a single-peptide molecule. *Chem. Phys. Lett. 315*, 61−68.

(52) Wang, Z., Wu, C., Lei, H., and Duan, Y. (2007) Accurate an Initio Study on the Hydrogen-Bond Pairs in Protein Secondary Structures. *J. Chem. Theory Comput. 3*, 1527−1537.

(53) Zhang, Y., and Wang, C. (2008) Estimation on the Intramolecular 10-Membered Ring N−H···O=C Hydrogen-Bonding Energies in Glycine and Alanine Peptides. *J. Comput. Chem. 30*, 1251−1260.

(54) Abraham, M. H., Grellier, P. L., Prior, D. V., Morris, J. J., and Taylor, P. J. (1990) Hydrogen Bonding. Part 10. A Scale of Solute Hydrogen-Bond Basicity using log *K* Values for Complexation in Tetrachloromethane. *J. Chem. Soc., Perkin Trans. 2 4*, 521−529.

(55) Choi, J., Oh, K., Lee, H., Lee, C., and Cho, M. (2008) Nitrile and thiocyanate IR probes: Quantum chemistry calculation studies and multivariate least-square fitting analysis. *J. Chem. Phys. 128*, 134506.

(56) Fafarman, A. T., Sigala, P. A., Schwans, J. P., Fenn, T. D., Herschlag, D., and Boxer, S. G. Quantitative, Directional Measurement of Electric Field Heterogeneity Within an Enzyme Active Site, manuscript submitted.

(57) Meller, J., and Elber, R. (1998) Computer Simulations of Carbon Monoxide Photodissociation in Myoglobin: Structural Interpretation of the B States. *Biophys. J. 74*, 789−802.

(58) Nutt, D. R., and Meuwly, M. (2003) Theoretical Investigation of Infrared Spectra and Pocket Dynamics of Photodissociated Carbonmonoxy Myoglobin. *Biophys. J. 85*, 3612−3623.

(59) Anselmi, M., Aschi, M., Nola, A. D., and Amadei, A. (2007) Theoretical Characterization of Carbon Monoxide Vibrational Spectrum in Sperm Whale Myoglobin Distal Pocket. *Biophys. J. 92*, 3442−3447.

(60) Plattner, N., and Meuwly, M. (2008) The Role of Higher CO-Multipole Moments in Understanding the Dynamics of Photodissociated Carbonmonoxide in Myoglobin. *Biophys. J. 94*, 2505−2515.

(61) Zheng, J., and Fayer, M. D. (2007) Hydrogen Bond Lifetimes and Energetics for Solute-Solvent Complexes Studied with 2D-IR Vibrational Echo Spectroscopy. *J. Am. Chem. Soc. 129*, 4328−4335.