Determinants of Cofactor Specificity in Isocitrate Dehydrogenase: Structure of an Engineered $NADP^+ \rightarrow NAD^+$ Specificity-Reversal Mutant[†,‡]

James H. Hurley,*[,§] Ridong Chen,[||] and Antony M. Dean[||]

Laboratory of Molecular Biology, National Institute of Diabetes, Digestive, and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0580, and Department of Biological Chemistry, The Chicago Medical School, 3333 Green Bay Road, North Chicago, Illinois 60064-3095

Received December 19, 1995; Revised Manuscript Received March 1, 1996[⊗]

ABSTRACT: The 7-fold mutation Cys201Met/Cys332Tyr/Lys344Asp/Tyr345Ile/Val351Ala/Tyr391Lys/Arg395Ser converts the cofactor specificity of *Escherichia coli* isocitrate dehydrogenase from a 7000-fold preference for $NADP^+$ to a 200-fold preference for NAD^+, with overall activity comparable to that of wild-type NAD^+-dependent isocitrate dehydrogenases. The structure of the NAD^+-dependent mutant has been determined and refined to a working *R*-factor of 0.186 at 1.9 Å resolution. The structure shows that $NADP^+$ affinity is destroyed by removing favorable interactions between the 2′-phosphate and Tyr345, Tyr391, and Arg395 and by adding a repulsive interaction with Asp344. NAD^+ affinity is enhanced by adding hydrogen bonds between Asp344 and the free 2′-hydroxyl. The favorable Asp344−2′-OH interaction requires a change in the pucker of the ribose to C2′ endo and a shift in the adenine ring. The ring shift is made possible by a series of changes in steric packing interactions. The linchpin for repacking in the adenosine binding site is residue 351. The side chain of this "second layer" residue dictates packing of the surrounding "first layer" residues which interact with the 2′ moiety and, in turn, directly determine specificity.

The *de novo* design of novel enzyme activities is the Holy Grail of protein engineering. While yet to be achieved, considerable progress has been made during the past 5 years in redesigning the activities of existing enzymes (Hedstrom, 1994). The dehydrogenases, by virtue of an accumulated wealth of structural and enzymological information together with a typically strong preference for either NAD^+ or $NADP^+$,[1] have become the targets of many redesign efforts. Most studies have dealt with enzymes that bind the dinucleotides in the classical Rossmann fold (Scrutton et al., 1990; Bocanegra et al., 1993; Nishiyama et al., 1993; Mittl et al., 1994). Recent studies have concentrated on engineering specificities in the structurally distinct dinucleotide binding domain of the decarboxylating dehydrogenases (Miyazaki & Oshima, 1994; Chen et al., 1995).

Decarboxylating dehydrogenases, which catalyze the Mg^{2+}- and $NAD(P)^+$-dependent oxidation of (2*R*,3*S*) 2-hydroxy acids followed by their Mg^{2+}-dependent decarboxylation, are ubiquitous in nature. The various isozymes of isocitrate dehydrogenase (IDH; EC 1.1.1.42) are involved with energy metabolism in the Krebs cycle, the production of glutamic acid, and provide NADPH for biosyntheses. Isopropylmalate dehydrogenases (IMDH; EC 1.1.1.85) are central to leucine biosynthesis. Despite sharing only 25% sequence identity, the crystallographic structures of *Escherichia coli* NADP-dependent IDH and *Thermus thermophilus* NAD^+-dependent IMDH reveal homodimers sharing a common, distinct protein fold (Hurley et al., 1989; Imada et al., 1991).

All residues involved in binding and catalysis at the 2(*R*)-malate core common to the substrates are rigidly conserved in the decarboxylating dehydrogenases, including the very divergent mammalian $NADP^+$-dependent isozyme (Haselbeck et al., 1992) which is only 16−24% (depending on gap penalties and alignment algorithms) identical to *E. coli* IDH. The catalytically inactive regulatory subunits of the octameric eukaryotic mitochondrial NAD^+-dependent IDH isozymes are the only exceptions. The nucleotide binding domains of the decarboxylating dehydrogenases, which are constructed from four loops, fall into three classes (Table 1; Chen et al., 1995): the NAD^+ binding domains of both eubacteria and eukaryotes form one class, while the $NADP^+$ binding domains of eubacteria and eukaryotes each form separate classes, the latter being the most divergent. Clearly, the decarboxylating dehydrogenases form a very ancient and very diverse class.

E. coli IDH has a 7000-fold preference for $NADP^+$, and *T. thermophilus* IMDH has a 140-fold preference for NAD^+. This strong discrimination between cofactors has made both enzymes attractive targets for protein engineering studies. Redesign of the *T. thermophilus* IMDH led to a triple mutant (Ser292Arg/Ser319Gly/Ile345Tyr, *E. coli* IDH numbering) with a 2-fold preference for $NADP^+$ but with overall activity reduced 1000-fold compared to wild type (Miyazaki & Oshima, 1994). That this attempt omitted engineering four

[†] Supported by USPHS Grant 1RO1-GM-48735 to A.M.D.

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (filename 1ISO) and will be made available prior to release at http://www-mslmb.niddk.nih.gov/.

* Corresponding author. Tel: (301) 402-4703. Fax: (301) 496-0201. E-mail: HURLEY@TOVE.NIDDK.NIH.GOV.

[§] National Institutes of Health.

[||] The Chicago Medical School.

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1996.

[1] Abbreviations: $NADP^+$, oxidized nicotinamide adenine dinucleotide phosphate; NAD^+, oxidized nicotinamide adenine dinucleotide; IDH, isocitrate dehydrogenase (EC 1.1.1.42); IMDH, isopropylmalate dehydrogenase (EC 1.1.1.85); MYDIAKS, *Escherichia coli* isocitrate dehydrogenase with the mutations Cys201Met/Cys332Tyr/Lys344Asp/Tyr345Ile/Val351Ala/Tyr391Lys/Arg395Ser; PCMBS, *p*-(chloromercuri)benzenesulfonic acid; AMP, adenosine 5′ monophosphate; NMN, nicotinamide adenine nucleotide; DDH, β-decarboxylating α-hydroxy acid dehydrogenase.

S0006-2960(95)03001-7 CCC: $12.00

Table 1: Sequence Alignments in Loops Comprising the Adenosine Binding Pocket[m]

$NADP^+$-specific	283[a]		294	316		325	334			344		351	357	388		395
E. coli IDH[b]	DAFLQQILLRPAEY			Q-VGGIGIAPGAN			LFE-ATHGTAPKY------AGQDKVNPGS-IILS							TVTYDFERLM		
Vibrio sp. IDH[c]	DAMLQQVLLRPAEY			Q-VGGIGIAPGAN			VFE-ATHGTAPKY------AGKNKVNPGS-VILS							TVTYDFERLM		
B. subtilis IDH[d]	DIFLQQILTRPNEF			Q-VGGIGIAPGAN			IFE-ATHGTAPKY------AGLDKVNPSS-VILS							VVTYDFARLM		
S. cerevisiae mitoIDH[e]	DDMVAQMLKSKGGY			G-FGSLGLMTSVL			ESD-RAHGTVTRHLTDYDKGRETSTNSIASIFAW							IMTKDLALIL		
S. cerevisiae cytoIDH[f]	DDMVAQMLKSKGGY			G-FGSLGLMTSVL			ESE-AAHGTVTRHFRQHQQGKETSTNSIASIFAW							IMTKDLALIL		
NAD^+-specific		†		† † †	†				**		*				*	*
Engineered *E. coli* IDH	DAFLQQILLRPAEY			Q-VGGIGIAPGAN			LFE-ATHGTAPDI------AGQDKANPGS-IILS							TVTKDFESLM		
T. thermophilus IMDH[g]	DAMAMHLVRSPARF			L-PGSLGLLPSAS			VFE-PVHGSAPDI------AGKGIANPTA-AILS							TPPPDLGGSA		
S. cerevisiae IMDH[h]	DSAAMILVKNPTHL			I-PGSLGLLPSAS			LYE-PCHGSAPDL-------PKNKVNPIA-TILS							IRTGDLGGSN		
N. crassa IMDH[i]	DSAAMLLVKNPRAL			I-PGSIGLLPSAS			IYE-PIHGSAPDI------SGKGIVNPVG-TILS							TKTKDLGGNA		
P. putida TDH[j]	DILCARFVLQPERF			C-AGTIGIAPSAN			LFE-PVHGSAPDI------FGKNIANPIA-MIWS							SVTPDMGGTL		
S. cerevisiae cataIDH[k]	DNSVLKVVTNPSAY			LSAGSLGLTPSAN			IFE-AVHGSAPDI------AGQDKANPTA-LLLS							NRTGDLAGTA		
S. cerevisiae reguIDH[l]	DNASMQAVAKPHQF			L-IGGPGLVAGAN			VFEPGSRHVGLDI------KGQNVANPTA-MILS							HTTRDIGGSS		

[a] *E. coli* IDH numbering. [b] Thorsness & Koshland, 1987. [c] Ishii et al., 1993. [d] Imai et al., 1987. [e] Haselbeck & McAlister-Henn, 1991; mitochondrial isozyme. [f] Loftus et al., 1994; cytosolic isozyme. [g] Kagawa et al., 1987. [h] Andreadis et al., 1984. [i] Li et al., 1993. [j] Tipton & Beecher, 1994. [k] Cupp & McAlister-Henn, 1991; catalytic subunit. [l] Cupp & McAlister-Henn, 1992; regulatory subunit. [m] Sites where mutagenesis was unsuccessful are indicated with a † and sites where mutagenesis was successful are indicated with an *(Chen et al., 1995).

Table 2: Kinetic Parameters of Wild-Type and Mutant Enzymes

	NADP			NAD			
enzyme	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (*A*) ($\mu M^{-1}\ s^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (*B*) ($\mu M^{-1}\ s^{-1}$)	specificity (*A*/*B*)
E. coli IDH at 21 °C[a]							
$K^{344}Y^{345}V^{351}Y^{391}R^{395}C^{332}C^{201}$ (wild type)	17	80.5	4.7	4700.0	3.22	0.00069	6900
$D^{344}I^{345}$—^{351}—^{391}—^{395}—^{332}—^{201}	7300	6.3	0.00086	3300.0	2.59	0.00078	1.1
$D^{344}I^{345}A^{351}$—^{391}—^{395}—^{332}—^{201}	6400	18.0	0.0028	850.0	9.39	0.011	0.25
$D^{344}I^{345}A^{351}K^{391}S^{395}Y^{332}M^{201}$	5800	4.70	0.00081	99.0	16.2	0.164	0.005
T. thermophilus IMDH at 65 °C[b]							
wild type	12300	29.9	0.00243	40	13.6	0.34	0.007
S. cerevisiae IDH at 24 °C[c]							
wild type				210	40	0.190	

The singe-letter amino acid code is used to denote amino acid residues. All apparent standard errors are less than 15% of the estimates. [a](Chen et al., 1995); [b](Miyazaki and Oshima, 1994); [c](Cupp and McAlister Henn, 1993).

key determinants of specificity underscores the difficulty in identifying those few critical residues in the molecular noise of the 65% divergence around the nucleotide binding domains.

We previously identified five determinants of cofactor specificity by comparing the structures of the binary complexes of *E. coli* IDH with $NADP^+$ and *T. thermophilus* IMDH with NAD^+ (Hurley et al., 1991; Hurley & Dean, 1994). Using this as a guide, the cofactor specificity of *E. coli* IDH has been completely inverted from a 7000-fold preference for $NADP^+$ to a 200-fold preference for NAD^+ (Table 2; Chen et al., 1995). The kinetic parameters and cofactor preference of the redesigned IDH are similar to those of naturally occurring IMDHs and NAD^+-specific IDHs. The high activity and overall 1.4×10^6-fold change in cofactor preference, achieved with just a handful of structurally targeted replacements, dramatize the usefulness of having in hand a three-dimensional structural comparison at the outset of redesign efforts. To further explore what lessons this exceptional reversal in specificity might have for enzyme design, we have determined the structure of the NAD^+-dependent mutant Cys201Met/Cys332Tyr/Lys344Asp/Tyr345Ile/Val351Ala/Tyr391Lys/Arg395Ser/(MYDIAKS) IDH bound to NAD^+.

To understand precisely the structural determinants of cofactor preference, one must be able to relate specific structural differences to changes in binding constants. This is hampered in structural comparisons of naturally occurring enzymes, where there are often extensive structural differences due to overall sequence differences. IDH and IMDH, whose structures differ by about 2.5 Å rms for C_αs in alignable regions, clearly fall into this class. The largest structural differences observed do not necessarily have any role in specificity differences, but it is virtually impossible to deconvolute gratuitous changes from changes critical for specificity. In a second goal of this study, we hoped to be able to dissect these two categories of structural variation by focusing on the structure of a designed enzyme having the minimum number of amino acid replacements needed to invert specificity.

MATERIALS AND METHODS

MYDIAKS IDH crystals (Hurley et al., 1989) were equilibrated in artificial mother liquor supplemented with 100

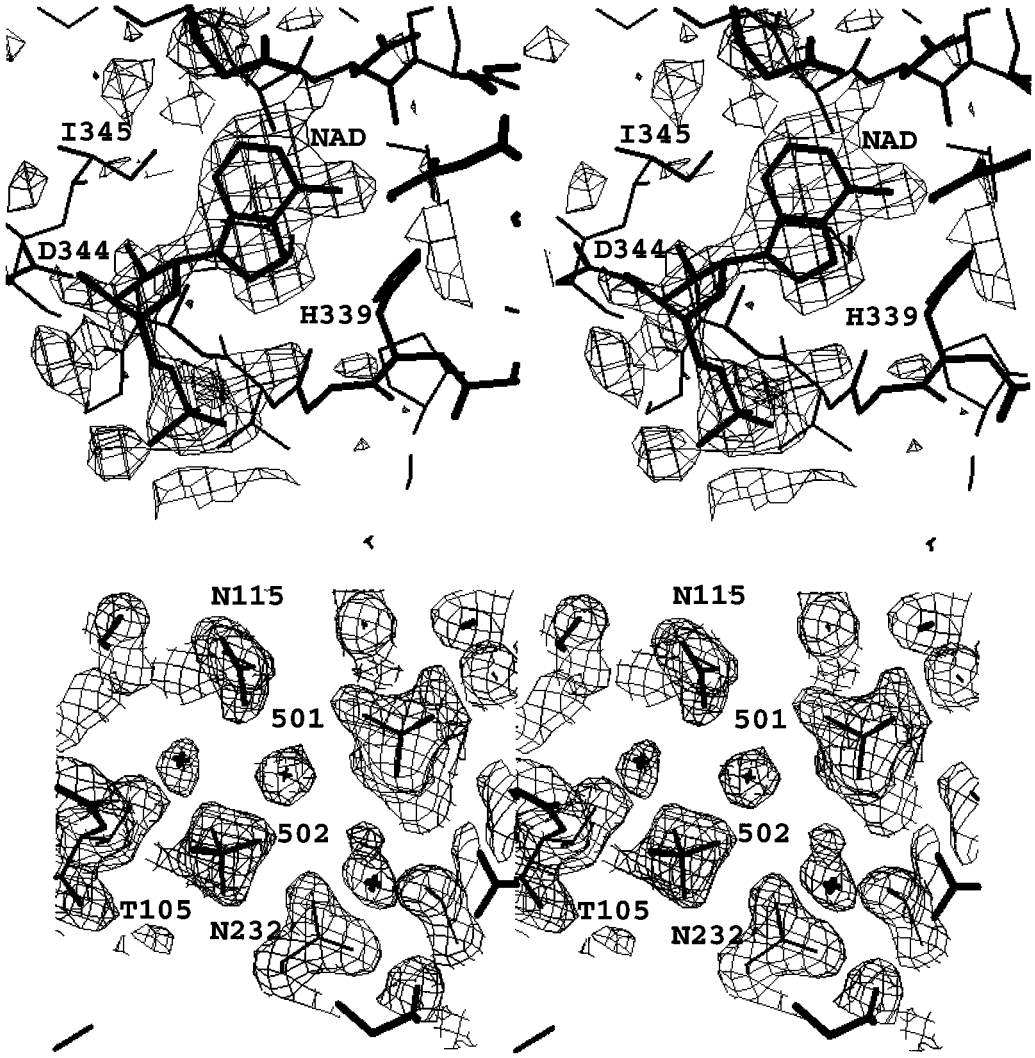

FIGURE 1: (A, top) Positive $(F_o - F_c)\alpha_{calc}$ difference density contoured at 2.0 standard deviations in the NAD^+ binding site. Phases and F_c coefficients are calculated from a model refined at 2.2 Å after including all mutated side chains and sulfate ions but prior to including water molecules and NAD^+. All data in the resolution range from 100 to 2.2 Å were used, with four-parameter scaling (Tronrud et al., 1987) of amplitudes. The final refined model, excluding solvent molecules, is superimposed. (B, bottom) Final $(2F_o - F_c)\alpha_{calc}$ electron density contoured at 1.0 standard deviation showing sulfates 502 and 501 superimposed on the final refined model.

mM NAD^+ (Sigma) for at least 1 h. A crystal 400 μm along the largest dimension was transferred to mother liquors with NAD^+ and 10%, 20%, and 30% glycerol for 5, 5, and 45 min, respectively. The crystal was then mounted in a rayon loop and flash-frozen in a dry nitrogen stream at −185 °C. Data were collected in 1° oscillations on an R-AXIS IIC at 140 mm using mirror-focused Cu Kα radiation. Essentially no radiation damage was observed; however, two crystals were used to obtain a complete data set due to the constraints of one-circle data collection geometry. Data were processed and scaled with Denzo and scalepack (Otwinowski, 1988). Based on autoindexing, cell parameters in $P4_32_12$ are $a = b = 102.4$ Å and $c = 150.6$ Å.

The *a* and *b* axes shrink by nearly 3% relative to room temperature, and a molecular replacement search was required to position the model. A rotation and translation search was carried out with Amore (Navaza, 1994) using data from 10 to 4 Å. The solution was rotated by 2° and translated by 4 Å in the *ab* plane relative to the search model and had an *R*-factor of 37% against 10−4 Å data. The model was refined with X-PLOR (Brunger, 1992b) by rigid body minimization of an eight-fragment model using 6−4 Å data, which reduced the working *R*-factor to 0.31 for this resolution range. This was followed by slow-cooling simulated annealing from 2000 to 300 K against 95% of the measured 6−2.2 Å data (no weak data cutoff). Side chains replaced by mutagenesis were modeled as Ala at this stage of refinement. Following simulated annealing, electron density for three sulfate ions from the crystallization medium was identified, and the ions were included in the model. Altered amino acid side chains were also added in the model at this stage using O (Jones et al., 1991) based on $F_o - F_c$ and $2F_o - F_c$ electron density maps. After one additional round of refinement including data to 1.9 Å, the adenosine monophosphate (AMP) moiety of NAD^+ was located in unbiased $F_o - F_c$ difference density (Figure 1A) and included in the model. Eleven side chains were placed which had not been located in previous 2.5 Å room-temperature structures. Two side chains, Asn18 and Gln20, remain impossible to place in density.

After refinement of the protein, sulfate ions, and AMP moiety converged, 345 water molecules were added by an automated procedure (Mason and Hurley, unpublished program), and 40 more waters were added manually. Waters were added to electron density peaks of 3.5σ or greater where hydrogen bonds could be formed. Roughly 20 waters in the

Table 3: Crystallographic Data as a Function of Resolution

d_{min}[a] (Å)	R_{merge}[b]	completeness[c] reflections	completeness[c] (%)	$\langle I\rangle/\sigma\langle I\rangle$[d]	R_{cryst}[e]
4.09	0.031	6790	99.8	60.1	0.222
3.25	0.043	6490	99.9	49.6	0.159
2.84	0.060	6389	99.9	27.7	0.173
2.58	0.082	6359	99.6	19.0	0.182
2.39	0.102	6320	99.6	13.3	0.191
2.25	0.118	6277	99.3	11.6	0.194
2.14	0.152	6222	98.9	9.3	0.205
2.05	0.181	6226	98.7	7.3	0.214
1.97	0.223	6208	98.6	5.9	0.230
1.90	0.273	6069	97.0	4.5	0.259
1.90	0.055	63350	99.1	22.6	0.196

[a] The indicated resolution is the upper limit for each bin. All low-resolution data are included in the first bin. The bottom row includes all data from the beam stop (100 Å) to 1.9 Å. [b] $R_{merge} = \Sigma|I_i - \langle I\rangle|/\Sigma\langle I\rangle$, where the summation is over all observations. Each data set was obtained from a single crystal. [c] Completeness is the percentage of all reflections measured with $\langle I\rangle/\sigma\langle I\rangle > -1.0$ within a resolution bin. [d] $\langle I\rangle/\sigma\langle I\rangle$ is the ratio, within a resolution bin, of the mean intensity of reflections to the mean standard deviation. [e] All measured reflections within a resolution bin are used to calculate $R_{cryst} = \Sigma|F_c - F_o|/\Sigma|F_o|$, where F_c is scaled to F_o using four parameters.

second hydration shell were included where justified by multiple hydrogen bonds to other waters, strong spherical electron density, and low B-factors following refinement. Water B-factors do not exceed 55 Å^2. A test set consisting of 5% of the measured reflections was reserved for free *R*-factor calculation (Brunger, 1992a) and used to monitor the addition of solvent molecules and individual B-factors to the model. Final working and free *R*-factors from 6 to 1.9 Å are 0.186 and 0.218. The crystallographic *R*-factor against all data from 100 to 1.9 Å is 0.196 (Table 3) using four-parameter scaling (Tronrud et al., 1987). Bond lengths and angles deviate from standard values by 0.013 Å and 2.53° rms, respectively.

RESULTS

Overall Structure of MYDIAKS IDH Refined at 1.9 Å. The overall structure of MYDIAKS refined against low-temperature data at 1.9 Å closely resembles that of the 2.5 Å wild-type IDH structure determined at room temperature. In spite of significant non-isomorphism, the coordinates of 414 C_αs differ by only 0.4 Å rms. The tentative identification of Pro262 as *cis* is confirmed by the higher resolution structure. Eleven polar amino acid side chains on the protein surface for which electron density was not seen in earlier maps have now been located. Only one residue, Arg96, is in a disallowed conformation, close to the left-handed helical region of the Ramachandran plot. This residue is well ordered ($\langle B\rangle = 18$ Å^2), and the conformation is similar in the wild-type structure. No large difference electron density features are observed around Arg96, suggesting that the conformation in the refined structure is correct. Asp259 is in poorly defined density and is one of three residues in generously allowed regions of the Ramachandran plot. The conformation of Asp259 is therefore less reliable than the rest of the structure. The average protein B-factor for the structure is 16.5 Å^2, compared to 28.0 Å^2 for the room-temperature structure of the IDH−NADP$^+$ complex. The average B-factor for all atoms, including the cofactor and solvent molecules, is 18.2 Å^2.

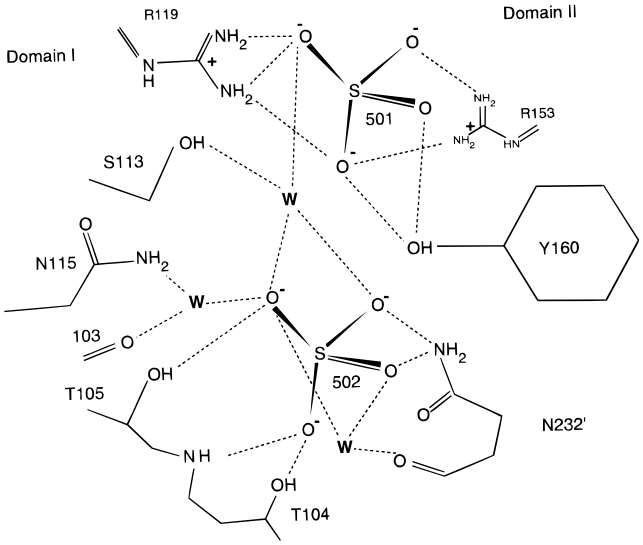

FIGURE 2: Schematic of sulfate binding sites in the active site cleft, illustrating bridging interactions between domain I (left) and domain II (right). Residues of subunit II are indicated with a prime.

Three sulfate ions bind in or near the active site cleft (Figure 1B). These ions bridge the two domains of IDH (Figure 2) and participate in lattice contacts. These three strong electron density features were present but not interpretable in previous room-temperature structures but were clearly interpretable as sulfate or phosphate ions in the high-resolution low-temperature structure. On the basis of their protein ligands and the presence of 2 M ammonium sulfate in the equilibration medium, these ions have been treated as sulfates in the refined structure. These density features are also present in the 2.2 Å structure of wild-type IDH refined against low-temperature data (J. H. Hurley, unpublished). Average B-factors for the three sulfate ions are 52.3, 68.9, and 27.5 Å^2. The high values for the first two suggest they may be bound at partial occupancy, but the shapes of the electron density leave little doubt regarding their identity. Sulfate 501 is bound directly by the side chains of Arg119, Arg153, and Tyr160 and by a water-mediated hydrogen bond to Ser113. Sulfate 502 makes direct contacts to the side chains of Thr104, Thr105, and Asn232′ (where the prime indicates the second IDH monomer) and to the main chain nitrogen of 104. Sulfate 502 also makes water-mediated hydrogen bonds to the main chain at 103 and 232′ and to the side chains of Ser113 and Asn115. Sulfates 501 and 502 interact with each other and Ser113 through the same water molecule. Sulfate 503 is bound by main chain peptide nitrogens of Gly108 and Gly110 and makes water-mediated interactions with the main chain at 106, 109, and 111. Sulfate 503 also participates in lattice contacts *via* an interaction with the side chain of Lys235 of a symmetry-related molecule.

Mutant Side Chain Conformations. The side chains of all seven mutated amino acids are well localized in the structure. Cys201, which is buried and inaccessible to reaction with PCMBS (Hurley et al., 1989), was replaced by Met. The introduced Met ($\langle B\rangle = 10.0$ Å^2) is accommodated by a preexisting cavity in the structure and is essentially completely buried. There are no long-range structural changes around this site, but the adjacent g helix moves 0.4 Å toward the site of the mutation. Cys332 is partially exposed and was one of the two sites of reaction

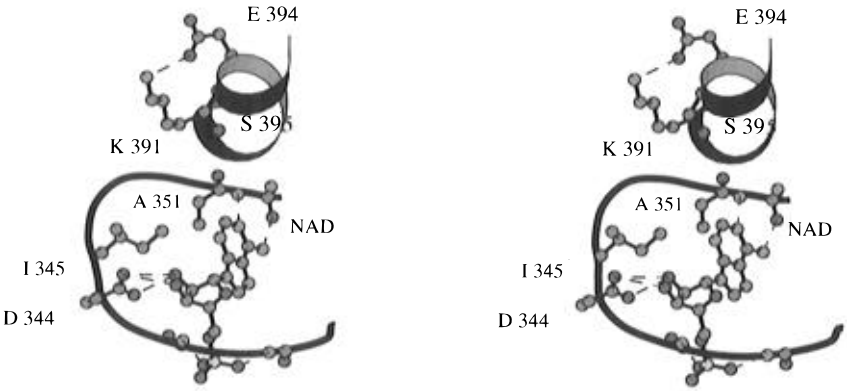

FIGURE 3: Stereoview (Kraulis, 1991) of the adenosine binding portion of the NAD^+ binding site of MYDIAKS-IDH. Oxygen atoms are colored red, nitrogens are in blue, carbon is in green, and phosphorus is in pink. Hydrogen bonds are indicated by dashed lines.

(out of five cysteines) with PCMBS (Hurley et al., 1989). The bulkier Tyr replacement ($\langle B \rangle = 15.4$ Å^2) at 332 contacts the C-terminal end of the c helix and pushes it away by 0.7 Å. These structural changes do not extend beyond the c helix and adjacent residues.

The introduced side chains Asp344 ($\langle B \rangle = 30.3$ Å^2) and Ile345 ($\langle B \rangle = 16.4$ Å^2) point into the adenosine binding site. The C_δ of Ile345 partially fills the void left by the Val351Ala replacement. Relative to the wild-type Tyr, the side chain of Lys391 ($\langle B \rangle = 17.6$ Å^2) flips entirely away from the adenosine binding site and forms a 3.0 Å salt bridge with Asp394. The short side chain of Ser395, replacing Arg, is well removed from the adenosine binding site. One $NADP^+$ binding residue which was not replaced in MYDIAKS, Arg292′, is in essentially the same conformation as in wild type but has no direct interactions with NAD^+. Arg292′ forms a bidentate salt bridge with Glu295′ that appears to lock the Arg side chain in place regardless of the presence or absence of the $NADP^+$ 2′-phosphate. The average B-factor for Arg292′ is 35.4 Å^2 in MYDIAKS as compared to 47.8 Å^2 in the wild-type $NADP^+$ complex, but this difference is essentially identical to the difference in average protein B-factors and does not suggest any real difference in flexibility.

Conformation and Binding of NAD$^+$. The adenosine 5′-monophosphate (AMP) moiety of NAD^+ is well ordered and clearly localized in electron density (Figure 1A). The cofactor binds in the same pocket in which $NADP^+$ binds to wild-type IDH (Figure 3). The NMN (nicotinamide mononucleotide) moiety could not be located and is presumed to be disordered. Indeed, several "tendrils" of uninterpretable electron density emanate from the AMP 5′-phosphate, suggestive of partially occupied NMN groups superimposed on partially occupied bound water molecules. Likewise, the NMN moiety of $NADP^+$ could not be located in the previous structure of its binary complex with wild-type IDH. The NMN group is better localized in ternary complexes with isocitrate, where turnover was inhibited by Ca^{2+} (Stoddard et al., 1993) or mutagenesis (Bolduc et al., 1995). In these structures there is an interaction between the isocitrate γ-carboxylate and the positively charged nicotinamide ring. The disordered NMN moiety observed here is expected on the basis of the absence of substrate in the complex. The last well-ordered group common to both binary complexes, the 5′-phosphate, only shifts 0.8 Å between the superimposed structures, based on the phosphorus position. This is comparable to the estimated coordinate error given the elevated B-factors for the adenosine 5′-phosphates, 54.6 Å^2 (phosphorus) for MYDIAKS and 73.9 Å^2 (phosphorus) for wild type. This does not suggest that any large structural changes are propagated beyond the adenosine moiety to the catalytic site.

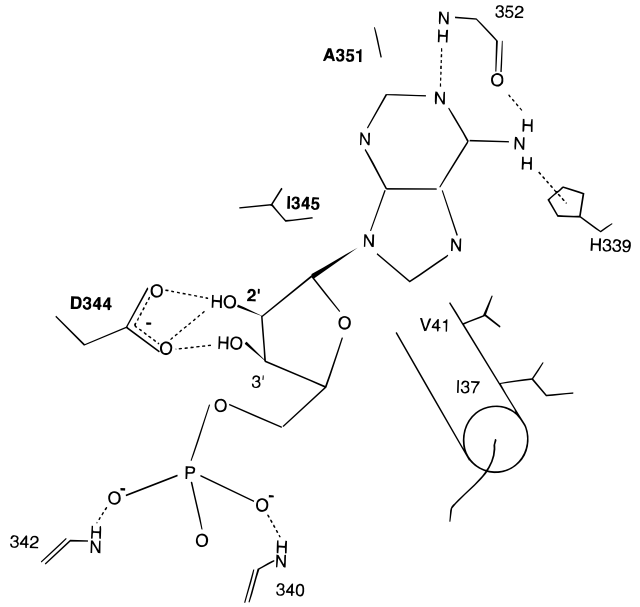

FIGURE 4: Schematic of the MYDIAKS IDH NAD^+ binding site showing interactions with the well-ordered AMP moiety. The unique 2′-hydroxyl of NAD^+ and mutated side chains are labeled in boldface.

The adenine ring is in the *anti* conformation (Figure 4), as found previously in both the IMDH and wild-type IDH cofactor complexes. The adenine ring is shifted substantially "inward" toward its principal protein ligands (Figure 5), with N3 shifted by 1.6 Å. The adenosine ribose follows the ring, with C1′ moving 1.5 Å in the same direction. Most dramatically, the adenosine ribose O2′ moves 3.7 Å due to the combined effects of the overall movement and of the change in sugar pucker from the C3′ endo to the C2′ endo conformation. The changes are not due to a gross movement of the entire cofactor binding site, since the main chain at the key binding residue Asp344 moves 0.7 Å in the opposite direction, toward the cofactor.

DISCUSSION

Implications of the Structure Refined at 1.9 Å Resolution. The first sulfate binds at a subsite of the isocitrate binding site corresponding roughly to the β- and γ-carboxylates of

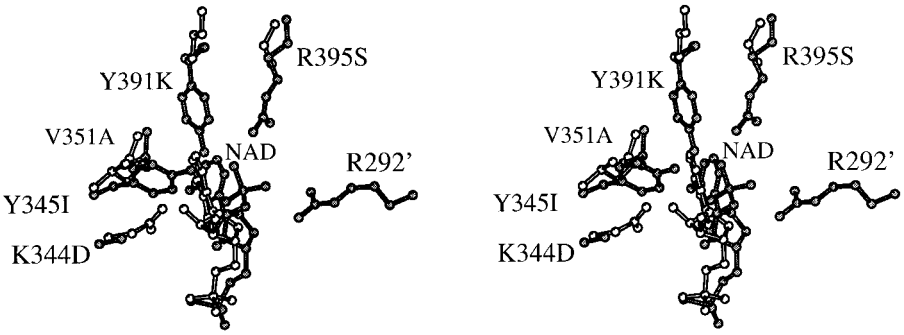

FIGURE 5: Stereoview of superimposed structures of wild-type (shaded bonds) and MYDIAKS (open bonds) IDH bound to $NADP^+$ and NAD^+, respectively, illustrating direct determinants of specificity. Mutations are labeled. Arg292′ is shown only for wild type, since the conformation is essentially identical in MYDIAKS. The poorly ordered Lys344 side chain of wild type is omitted.

the substrate. This suggests that sulfate could be a competitive inhibitor with respect to isocitrate at sufficiently high concentrations. More intriguingly, by bridging the two domains (Figure 2), sulfates 501 and 502 might provoke the same type of conformational changes as the substrate. An apparent paradox in the structural biology of the decarboxylating dehydrogenases has been that IMDH structures determined in the absence of isopropylmalate (Imada et al., 1991; Hurley & Dean, 1994) have displayed open or partially open conformations, while IDH structures display a closed conformation whether or not substrate is bound. One possibility is that there is a fundamental difference between the conformational flexibility of IDH and IMDH.

The present structure suggests an alternative explanation for the occurrence of closed conformations of IDH in the absence of substrate, namely, that the high concentrations of sulfate in all available IDH structures have stabilized the closed conformation by binding in the active site and bridging the two domains. Single and double mutations at several positions interacting with the sulfates (Ser113, Asn115, Tyr160) do not affect the ability of IDH to crystallize in the closed form in the absence of substrate (Hurley et al., 1990; Lee et al., 1995); hence not all of the sulfate−protein interactions are required for the putative stabilization of the closed form. IMDH was first crystallized in an open conformation from ammonium sulfate (Imada et al., 1991). Only one (corresponding to Arg119) out of the five sulfate binding residues on domain I of IDH is conserved in IMDH, however. Two bound sulfates are reported in the 2.2 Å structure of unliganded IMDH, but neither one is in the active site cleft. Further investigation will clearly be required to resolve the significance of the domain bridging sulfate ions.

Specificity Determinants: "First Layer" Interactions. Specificity is governed first by residues that interact directly with the unique 2′-hydroxyl and phosphate groups of NAD^+ and $NADP^+$, second by more distant residues which modulate the effects of the first group, and third by remote residues. For convenience, we will refer to the first class as "first layer" residues. Residues which are in direct contact with the first layer but not in contact with the unique cofactor moieties we refer to as "second layer". Residues beyond the second layer will be referred to as "long range".

The introduction of the Asp344 side chain appears to have all of the intended direct effects on specificity, both positive and negative. It is poised to electrostatically repel a 2′-phosphate group if it were present on the cofactor. The carboxylate also forms a triple hydrogen bond with the adenosine ribose. A single hydrogen bond is formed between one carboxylate oxygen and the 3′-hydroxyl. Both carboxylate oxygens are within hydrogen-bonding distance of the 2′-hydroxyl, forming hydrogen bonds in the strongly favored *syn* conformation (Gandour, 1981). The free 2′-hydroxyl is the only moiety unique to NAD^+, and there is little doubt that the highly favorable interaction geometry between Asp344 and the hydroxyl contributes decisively to NAD^+ specificity. This feature of NAD^+ binding sites has been noted previously (Wierenga et al., 1985; Mittl et al., 1994; Hurley & Dean, 1994), and its inclusion was almost certainly critical to the redesign of IDH.

Discrimination against the presence of a 2′-phosphate is also achieved by removing hydrogen bonds with Tyr345 and Tyr391 and a salt bridge with Arg395. The replacement of Tyr345 with a hydrophobic Ile side chain and Arg395 with the much shorter Ser destroys any potential for electrostatic or hydrogen-bonding interactions. The effect of the Tyr391Lys mutation is more subtle. The introduced Lys side chain forms a salt bridge with Asp394 in the mutant structure. Although modeling suggests that an interaction could still be formed with the 2′-phosphate by a large movement of the Lys, this could only occur at the expense of disrupting a preexisting salt bridge, therefore contributing less to affinity. The creation of the Lys−Asp salt bridge was fortunate but not foreseen. It was found not only unnecessary but also slightly deleterious to replace Arg292′ (Chen et al., 1995). It is easy to imagine that the changes at residues 344, 345, 391, and 395 so thoroughly destroy the 2′-phosphate binding site that any further disruptions become superfluous.

Specificity Determinants: "Second Layer" and Long-Range Interactions. The structure reveals several crucial indirect effects of Val351Ala in changing specificity. The adenine ring of NAD^+ approaches the C_α of residue 351 more closely than does $NADP^+$. This movement is coupled to the change in pucker of the adenosine ribose of NAD^+ and its near approach to Asp344, suggesting an important role for the ring shift in specificity. The Val351Ala mutant was designed to avoid obstructing this ring shift. The NAD^+ adenine in the mutant structure approaches within 2.9 Å of the Val side chain in the superimposed wild-type structure, an unfavorably short distance. The closest point of contact is to the C_β of Val351. Although the C_β is common to the Ala replacement, the absence of a side chain allows the Ala to relax by moving 1.3 Å away from the adenosine. The position of the MYDIAKS Ala351 C_β corresponds within 0.5 Å to that of the wild-type Val351 $C_{\gamma 2}$. The Val is incapable of such an adjustment because its $C_{\gamma 2}$ is tightly

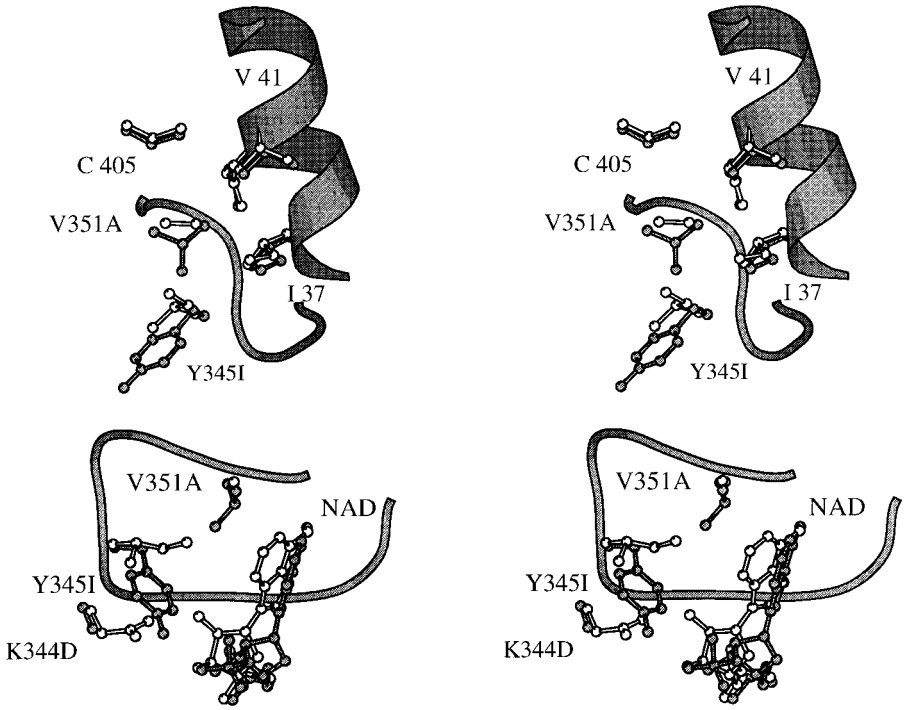

FIGURE 6: Stereoviews of superimposed structures of wild-type (shaded bonds) and MYDIAKS (open bonds) IDH illustrating (A, top) the hydrophobic wall against which the side chain of residue 351 packs that is formed by residues 37, 40, 41, 345, and 405. (B, bottom) Potential severe steric conflicts between wild-type Val351 and both mutant Ile345 and the adenine ring of mutant-bound NAD^+ and between wild-type Tyr345 and both mutant Asp344 and the ribose of mutant-bound NAD^+. The Ala351 side chain is obscured in (B) because its C_β coincides with the $C_{\gamma 2}$ of the wild-type Val351, as seen in (A). Coil and helix secondary structures are very similar in both structures and for clarity are shown for wild-type only.

packed against Ile37 (3.6 Å to C_β), Asp40 (4.4 Å to C_β), Val41 (3.8 Å to $C_{\gamma 1}$), and Cys405 (4.8 Å to S_γ) (Figure 6A).

A second key role for Val351Ala is to accommodate the correct packing of the introduced Ile345. The wild-type Val351 and the mutant Ile345, in their respective conformations, approach within 1.8 Å of each other, a severe violation of packing constraints (Figure 6B). The introduced Ile might avoid this collision by displacing Asp344. This would destroy the key Asp344−2′-hydroxyl interaction, however. By contrast, the Cβ of Ala351 approaches no closer than 4.1 Å to Ile345.

The importance of Val351Ala explains why the double mutation Lys344Asp/Tyr345Ile destroys $NADP^+$ binding without significantly promoting NAD^+ binding. The corresponding residue is Val or Ile in all known $NADP^+$-dependent IDHs and Ala in most, but not all, NAD^+-dependent DDHs. For instance, this residue is Val in the IMDHs of *Saccharomyces cerevisiae* and *Saccharomyces pombe* (Table 1). In view of the critical role of the Val351Ala mutation in the specificity of MYDIAKS, it is intriguing that this residue is not absolutely conserved among the NAD^+-dependent DDHs. This accords with the observation that there are, in general, many ways to solve the packing problem (Ponder & Richards, 1987). The Val351Ala mutation is apparently the simplest way to sterically accommodate the rest of the NAD^+ binding site, but other possibilities might be mutations which alter the length or conformation of the loop 337−352 or mutations in the residues (37, 40−41, 345, 405) against which 351 packs. The yeast IMDH sequences manifest additional changes in both these regions.

The structure shows that the Tyr345Ile mutation eliminates unfavorable contacts with NAD^+ as well as favorable contacts with $NADP^+$. The wild-type Tyr345 ring approaches within 1.9 Å of the mutant conformation of the NAD^+ adenosine 2′-hydroxyl and within 2.3 Å of the Asp344 carboxylate. The Tyr391 hydroxyl approaches within 2.3 Å of the adenosine ring. All of these contacts are close enough to be highly unfavorable, and the rigid and bulky structures of the Tyr side chains make them difficult to relieve. In combination, these steric interactions could account for much of the discrimination against NAD^+ by wild-type IDH.

The two cysteine replacements are far from the cofactor binding site. Cys201 and Cys201′ are 38 and 33 Å from the adenosine binding site, and Cys 332 and Cys332′ are 33 and 48 Å from the adenosine binding site, respectively. The mutations induce structural shifts of less than 1 Å which do not extend farther than 15 Å from the site of mutation. The effects on kinetics are modest and not highly cofactor-specific. Kinetics were previously reported for a 7-fold mutant including Cys201Ile rather than Cys201Met, but the results are essentially identical. Cys201Met/Cys332Tyr increases performance by factors of 4.5 and 7.8 for $NADP^+$ and NAD^+, respectively, compared to the 5-fold mutants (Chen et al., 1995). The tactic of mutating arbitrarily chosen cysteines appears to have been helpful, but the mutant structure provides no obvious explanation for the subtle activating effects seen. Similarly, substantial long-range effects on substrate specificity have been observed in α-amylase (Holm et al., 1990) and in serine proteases (Altschuh et al., 1994; Mace et al., 1995). Mace et al. (1995) report marked activation of α-lytic protease toward abnormal substrates by mutations at several sites as far as 21 Å from the active site. The structural mechanism of long-range specificity changes is not well understood in any system but might promote conformations suitable for low-energy binding through effects on flexibility or dynamics.

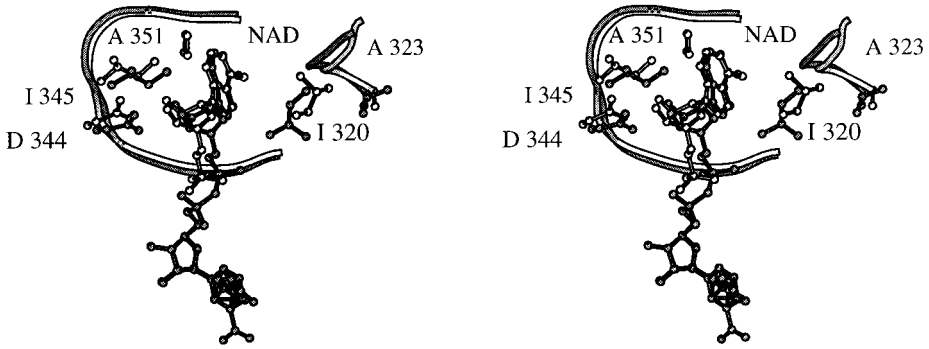

FIGURE 7: Stereoview of structures of the IMDH−NAD^+ complex (shaded bonds) and MYDIAKS (open bonds) IDH illustrating additional van der Waals contacts formed between NAD^+ and the loop of *T. thermophilus* IMDH corresponding to IDH 318−323. The superposition was based on 43 C_αs in the cofactor binding sites of IDH (322−329 and 332−366) and IMDH (256−263 and 266−300) whose coordinates deviate by 0.8 Å rms. Both observed conformers are shown for the nicotinamide ring of NAD^+ bound to IMDH (Hurley & Dean, 1994).

Implications for Naturally Occurring NAD^+-Dependent Isocitrate Dehydrogenases. The most prominent difference between the NAD^+ binding sites of IMDH and the mutant IDH is found in the loop 317−324 (Figure 7). In IMDH, the residues equivalent to 317−320 are in a different conformation from IDH, which positions Leu254 (320 in IDH numbering) of IMDH to form a hydrophobic contact with the NAD^+ adenine ring. In IDH, by contrast, the loop is removed from the adenosine binding site, leaving one face of the ring exposed to solvent. The conformation of the loop appears to be determined by Leu257 (323 in IDH) of IMDH, whose side chain is sterically incompatible with the loop conformation found in IDH. Leu257 is conserved throughout the IMDH sequences, but not in NAD^+-dependent IDHs. Extensive efforts to reengineer this loop or replace it entirely resulted in enzymes with markedly reduced activity (Chen et al., 1995), however. The reasons for this are unclear, but one possibility is that residues 317−323, which form the second connection between the two domains of IDH, have a role in interdomain alignment. The structure makes it clear that contacts between the adenine and the 317−323 loop are not essential for NAD^+ specificity. No structure is currently available for a naturally occurring NAD^+-dependent IDH. The MYDIAKS structure, as an NAD^+-dependent enzyme with sequence features intermediate between $NADP^+$-dependent IDHs and NAD^+-dependent IMDHs, probably resembles the three-dimensional structure of natural NAD^+-dependent IDHs more closely than any available structure of a natural DDH.

Lessons for Enzyme Design. The structure of MYDIAKS IDH dramatizes the critical role of residue 351 in controlling specificity indirectly by steric packing with first-layer residues. Our observation of key second-layer specificity determinants adjacent to the most prominent direct interactions reinforces similar observations on trypsin (Perona et al., 1995). In spite of a wealth of structural information at the outset of the design effort, these packing interactions were only partly anticipated. This illustrates the extraordinary difficulty in anticipating these subtle second-layer effects. In the more usual situation, there is no recourse to a crystal structure of an enzyme cousin which already has the desired specificity. Future enzyme engineering efforts will be aided by carefully considering not only those residues in direct contact with substrates but also the surrounding steric packing interactions as well. Systematic searches for acceptable packing combinations (Ponder & Richards, 1987) and iterative cycles of design and structure determination may prove helpful.

A second major finding is that mutations distant from the binding site can affect specificity without the propagation of any large structural changes between the two sites. The structure therefore rules out any mechanism involving large-scale structural changes in the bound conformation. The most probable mechanisms involve instead destabilization of other conformations which bind more poorly to the cofactor, an increase in the rate of transition to the preferred conformation, or both. Agard and co-workers (Mace et al., 1995) have suggested this type of mechanism in α-lytic protease based on exhaustive mutagenesis of a loop distant from the active site. Our finding of marked effects on specificity in the absence of discernible structural changes provides direct evidence in support of this type of communication in a very different enzyme system. Whatever the underlying structural mechanism, random mutagenesis of distant residues is clearly an important complement to conventional structure and homology-based design. The finding that different types of mutations at distant sites can affect specificity in favorable ways is cause for optimism, suggesting that exhaustive mutagenesis and screening may not always be necessary.

ACKNOWLEDGMENT

We thank A. Greer for excellent technical assistance.

REFERENCES

Altschuh, D., Tessier, D. C., & Vernet T. (1994) *Protein Eng. 7*, 769−775.
Andreadis, A., Hsu, Y. P., Hermodson, M., Kohlhaw, G., & Schimmel, P. (1984) *J. Biol. Chem. 259*, 8059−8062.
Bocanegra, J. A., Scrutton, N. S., & Perham, R. N. (1993) *Biochemistry 32*, 2727−2740.
Bolduc, J. M., Dyer, D. H., Scott, W. G., Singer, P., Sweet, R. M., Koshland, D. E., Jr., & Stoddard, B. L. (1995) *Science 268*, 1312−1318.
Brunger, A. T. (1992a) *Nature 355*, 472−474.
Brunger, A. T. (1992b) *X-PLOR manual version 3.1: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
Chen, R., Greer, A., & Dean, A. M. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 11666−11670.
Cupp, J. R., & McAlister-Henn, L. (1991) *J. Biol. Chem. 266*, 22199−22205.
Cupp, J. R., & McAlister-Henn, L. (1992) *J. Biol. Chem. 267*, 16417−16423.
Gandour, R. D. (1981) *Bioorg. Chem. 10*, 169−176.
Haselbeck, R. J., & McAlister-Henn, L. (1991) *J. Biol. Chem. 266*, 2339−2345.
Haselbeck, R. J., Colman, R. F., & McAlister-Henn, L. (1992) *Biochemistry 31*, 6219−6223.

Hedstrom, L. (1994) *Curr. Opin. Struct. Biol. 4*, 608−611.
Holm, L., Koivula, A. K., Lehtovaara, P. M., Hemminki, A., & Knowles, J. K. C. (1990) *Protein Eng. 3*, 181−191.
Hurley, J. H., & Dean, A. M. (1994) *Structure 2*, 1007−1016.
Hurley, J. H., Thorsness, P., Ramalingam, V., Helmers, N., Koshland, D. E., Jr., & Stroud, R. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8635−8639.
Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E., Jr., & Stroud, R. M. (1990) *Science 249*, 1012−1016.
Hurley, J. H., Dean, A. M., Koshland, D. E., Jr., & Stroud, R. M. (1991) *Biochemistry 30*, 8671−8678.
Imada, K., Sato, M., Tanaka, N., Katsube, Y., Matsuura, Y., & Oshima, T. (1991) *J. Mol. Biol. 222*, 725−738.
Imai, R., Sekiguchi, T., Hosoh, Y., & Tsuda, K. (1987) *Nucleic Acids Res. 15*, 4988.
Ishii, A., Suzuki, M., Sahara, T., Takada, Y., Sasaki, S., & Fukunaga, N. (1993) *J. Biol. Chem. 175*, 6873−6880.
Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.
Kagawa, Y., Nojima, H., Nukiwa, N., Ishizuka, M., Nakajima, T., Yasuhara, T., Tanaka, T., & Oshima, T. (1984) *J. Biol. Chem. 259*, 2956−2960.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−960.
Lee, M. E., Dyer, D. H., Klein, O. D., Bolduc, J. M., Stoddard, B. L., & Koshland, D. E., Jr. (1995) *Biochemistry 34*, 378−384.
Li, Q., Jarai, G., Yaghmai, B., & Marzluf, G. A. (1993) *Gene 136*, 301−305.
Loftus, T. M., Hall, L. V., Anderson, S. L., & McAlister-Henn, L. (1994) *Biochemistry 33*, 9661−9667.
Mace, J. E., Wilk, B. J., & Agard, D. A. (1995) *J. Mol. Biol. 251*, 116−134.
Mittl, P. R. E., Berry, A., Scrutton, N. S., Perham, R. N., & Schulz, G. E. (1994) *Protein Sci. 3*, 1504−1514.
Miyazaki, K., & Oshima, T. (1994) *Protein Eng. 7*, 401−403.
Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
Nishiyama, M., Birktoft, J. J., & Beppu, T. (1993) *J. Biol. Chem. 268*, 4656−4660.
Otwinowski, Z. (1988) *DENZO. A program for automatic evaluation of film densities*, Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT.
Perona, J. J., Hedstrom, L., Rutter, W. J., & Fletterick, R. J. (1995) *Biochemistry 34*, 1489−1499.
Ponder, J., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
Scrutton, N. S., Berry, A., & Perham, R. N. (1990) *Nature 343*, 38−43.
Stoddard, B. L., Dean, A. M., & Koshland, D. E., Jr. (1993) *Biochemistry 32*, 9310−9316.
Thorsness, P. E., & Koshland, D. E., Jr. (1987) *J. Biol. Chem. 262*, 10422−10425.
Tipton, P. A., & Beecher, B. S. (1994) *Arch. Biochem. Biophys. 313*, 15−21.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr., Sect. A 43*, 489−501.
Wierenga, R. K., De Maeyer, M. C. H., & Hol, W. G. J. (1985) *Biochemistry 24*, 1346−1357.

BI953001Q