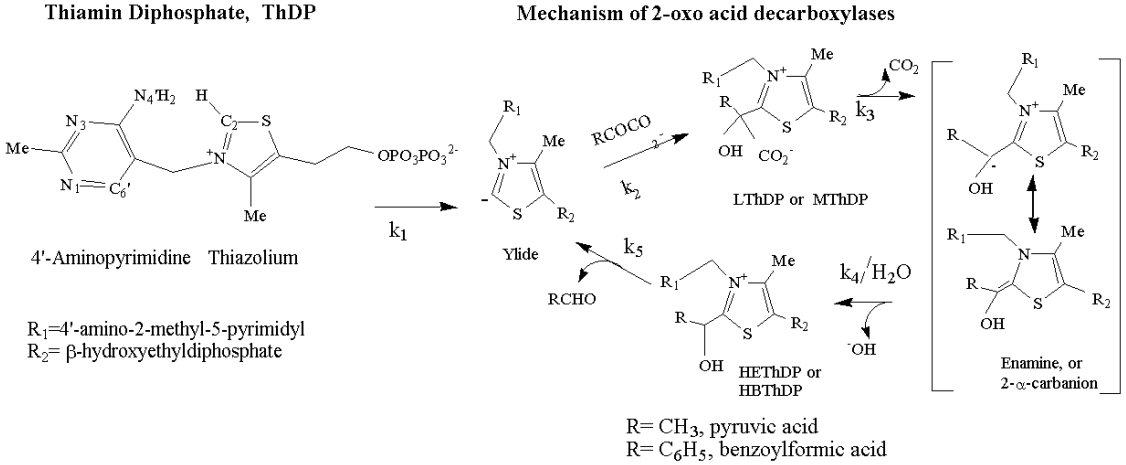
STRUCTURE-FUNCTION STUDIES IN THIAMIN DIPHOSPHATE-DEPENDENT 2-OXO ACID DECARBOXYLATING ENZYMES

F. Jordan, E. Sergienko, N. Nemeria, Min Liu, Jue Wang, Fusheng Guo, and W. F. Furey

The mechanism of enzymatic activation of the otherwise rather unreactive thiamin diphosphate by the 2-oxoacid decarboxylases is summarized. Given the significant number of such enzymes with known 3-dimensional structures, some highly conserved features emerge. Most prominent among these are: (a) a precisely tailored Mg(II) binding site in which the octahedral metal provides two coordination sites to the diphosphate; (b) the 'V' coenzyme conformation that brings the amino nitrogen of the 4′-aminopyrimidine of the coenzyme to nearly hydrogen-bonding distance of the thiazolium C2 atom; (c) highly conserved hydrogen bonds to each of the three nitrogens of the 4′-aminopyrimidine moiety. The working hypothesis based on these conserved features is that deprotonation of the C2H thiazolium position, the first absolutely required step in all thiamin-requiring enzymes, is carried out by the coenzyme itself via intramolecular acid-base catalysis. The semi-conserved acid-base groups at the active center carry out an important rate accelerating function, but their absence is seldom 'fatal' to the enzyme. It is further hypothesized that a major role of the protein is to create a 'low effective polarity' in the active center, thereby stabilizing zwitterionic intermediates and preceding transition states.

With the advent of X-ray crystallographic structure determinations of proteins, and in conjunction with the ability to make substitutions at specific sites of the protein through the capabilities of modern molecular biology, there is an excellent opportunity to evaluate the additional catalytic effects of the protein on coenzyme-dependent enzymatic reactions. Thiamin diphosphate (ThDP, the vitamin B1 coenzyme) is the cofactor responsible for enzymatic decarboxylations of 2-oxo acids [1]. Its structure and function on pyruvate decarboxylase (PDC, EC4.1.1.1—Refs. [2, 3] are shown in Scheme 1. The accepted mechanism involves the intermediacy in the reaction of two unstable zwitterionic intermediates: the C2-carbanion/ylide/carbene and the C2α-carbanion or enamine. The mechanism also invokes three covalent ThDP-bound intermediates: the C2α-lactylThDP (LThDP, a ThDP-substrate adduct); C2α-hydroxyethylidene-ThDP (the enamine produced by the decarboxylation), and the C2α-hydroxyethylThDP (HEThDP, a ThDP-product adduct). In solution, the formation of these unstable zwitterionic intermediates is evidenced by their high pK_as (pK_a is the negative log of the acid dissociation constant of the positively charged stable species to the zwitterionic intermediates). In particular, the currently mentioned value for the pK_a at C2 is 17–19 in water [4],



Scheme 1

Department of Chemistry, Rutgers, the State University, Newark, New Jersey, 07102 USA, and Veterans Administration Biocrystallography Laboratories, Pittsburgh, Pennsylvania, USA. E-mail: frjordan@newark.rutgers.edu.

Abbreviations: ThDP, thiamin diphosphate; PDC, pyruvate decarboxylase from *Saccharomyces cerevisiae* overexpressed in *Escherichia coli*; E91D PDC, the variant with glutamate to aspartate substitution at position E91; LThDP, C2α-lactylThDP; enamine, C2α-hydroxyethylideneThDP; HEThDP, C2α-hydroxyethylThDP; HBThDP, C2α-hydroxybenzylThDP.

it is 15.4 at C2α for the C2α-hydroxybenzylthiazolium salt in water [5, 6], and > 15 for the C2α-hydroxyethylthiazolium salt in DMSO-water mixtures [7]. In pure DMSO, the pK_a is 14 at C2α for the C2α-hydroxyethylthiazolium salt and 12.5 for the C2α-hydroxybenzylthiazolium salt analogue [8]. The enamine intermediate has been characterized extensively in solution [9] and has also been observed on PDC when derived from highly conjugated pyruvate analogs [10–12]. Recently, the enamine has also been observed on the enzyme benzoylformate decarboxylase (BFD; Ref. [13]).

The reactions of ThDP include both: non-oxidative reactions such as PDC [1] and BFD [14] producing acetaldehyde and benzaldehyde; and oxidative pathways, such as the pyruvate oxidases (POX) using flavin as the oxidant to produce acetate or acetyl phosphate [15], the family of 2-oxo acid dehydrogenase multienzyme complexes which utilize lipoic acid as the oxidant to produce acylCoA [16]; and pyruvate–ferredoxin oxidoreductase [17] which uses Fe_4S_4 cluster chemistry to produce acetylCoA. It is the intent of this report to summarize some of the features shared by the known structures of ThDP enzymes. Efforts in the authors' labs have so far concentrated on the enzymes carrying out decarboxylations; hence others such as transketolase (TK, Ref. [18]) are not discussed.

The quaternary structures. The known examples of PDC [from yeast [2, 3] and from *Zymomonas mobilis* [19] share a α_4 quaternary structure with pyruvate oxidase [15] and benzoylformate decarboxylase [14]. Most intriguing is the similarity in the extent of α-helix and β-sheets in all of these structures, also accounting for very similar domain structures. In PDC, for example, there are three domains (α, β, γ, starting with the amino terminus) of approximately 185 amino acids each. Typically, the ThDP cofactor resides at the interface of the α and γ domains, while the β domain appears to hold the two other domains in the proper orientation. The β domain also appears to have regulatory roles. The 2-oxo acid dehydrogenase multienzyme complexes exist both as an $\alpha_2\beta_2$ heterotetramer and as a α_2 homodimer. Just in 1999 and 2000, finally there is structural information forthcoming about the first, ThDP-dependent enzymes in these fascinating complexes. Unanticipated by the authors earlier, elucidation of the regulatory pathways on these enzymes has become a major goal of these studies, especially since there is a wealth of different regulatory mechanisms observed.

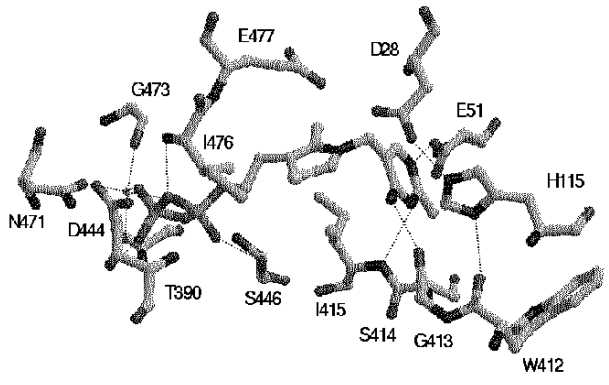


Fig. 1. Active Center of Pyruvate Decarboxylase.

The thiamin diphosphate fold and its functions. Alignment of sequences of ThDP enzymes [20] led to the early observation that there is a short stretch of amino acid residues consisting of a GDG $(X)_{26}$ N(C)N in all of these enzymes (where the 26 amino acid stretch is not highly conserved), and that this stretch could indeed constitute the ThDP fold. This astute suggestion has been amply substantiated in the intervening decade on all of the ThDP enzymes with known structures. Substitution of the D in the GDG triplet is fatal to the enzyme, it cannot function. Even substitution of the glycine at the carboxyl side of the D has greatly deleterious effects on the *E. coli* PDHc-E1 [21].

The function of the fold on PDC appears to be first and foremost to hold on to the Mg(II) ion, also absolutely required by ThDP enzymes. The D of the GDG triplet is one of the six ligands on the octahedrally coordinated ion, while the N of the N(C)N provides one additional ligand. In turn, the Mg(II) forms inner-sphere complexes to oxyanions of both the α and β phosphates of the diphosphate side chain of ThDP. This picture accounts for the absolute requirement for an intact metal binding site.

More unexpected, it was found that substitutions at the glycine at the carboxyl side the of the GDG triplet in *E. coli* PDHc-E1 [21] affects the hysteretic kinetic behavior of the enzyme with respect to ThDP.Mg(II). It was thereupon concluded that the loop comprising this diphosphate.Mg(II) fold is responsible for the behavior. Since the environment around the Mg(II) is so intimately involved in this, we speculated that ligand distortion around the ion induced by the substitutions was responsible for the observations.

The coenzyme conformation and its consequences. The free thiamin and the ThDP coenzyme have virtually free rotation of the two aromatic rings with respect to the bridging methylene group, with some preference for the so-called F conformation. Once substituted at the thiazolium C2 atom, a so-called S conformation is induced. It was a considerable surprise that when the first ThDP enzyme structures were published (transketolase, pyruvate oxidase and yeast pyruvate decarboxylase), neither of these conformations was found, rather the bound conformation is the so-called 'V' conformation. This feature has been found in all ThDP structures examined, and by now, there are examples of all major classes of ThDP enzymes. In a detailed examination of the reasons for this, the size of I415, the hydrophobic side chain poised under ThDP in PDC was varied to smaller and smaller groups, leading to diminishing activity [22]. In addition, it was also found that replacement of Ile415 by Met or Leu (as in POX and TK at the corresponding position), also led to at least 10-fold reduction in activity. It was concluded that for each enzyme, the side chain with the best fit (dictated by the backbone structure) has evolved. Computational studies were carried out to test the hypothesis that this side chain by itself supported the V conformation. Surprisingly, in fact, the binding of the diphosphate.Mg(II) and the three conserved hydrogen bonds at the aminopyrimidine ring (see below) also enforces the V conformation. In other words, there must have existed considerable evolutionary pressures to create this conformation on all ThDP enzymes.

One can speculate about the consequences of the V conformation and the hydrophobic side chain poised to support it. The preponderance of the evidence is that Leu, Ile or Met serves as the pivot in all of these enzymes, and they could contribute as much as 2–3 kcal/mol stabilization of the transition states by creating a more hydrophobic active center. A second consequence is revealed by simply measuring the distance between the N4′ atom of the aminopyrimidine ring and the thiazolium C2 atom, the key to catalytic activity [3]. This distance in all of the structures determined so far is 3–3.2 Å, certainly appropriate for intramolecular proton transfer. In fact, whether the proton transfer is direct (a view favored by the authors) or water mediated, this short distance (a result of the V conformation) must have some function, in view of the intrinsic high energy of this ThDP conformer off the enzyme.

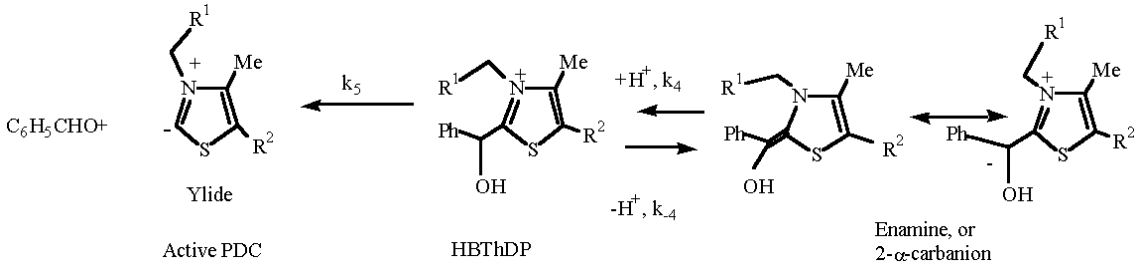
The conserved hydrogen bonds around the pyrimidine ring. There are three conserved hydrogen bonds to N1′, N3′ and N4′$H_{3'}$ (denoting the proton bonded to N4′ on the N3′ side of the pyrimidine ring) of the amino-pyrimidine ring. On PDC, these are from E51COOH, I415 main chain NH and G413 main chain C=O on the protein, respectively. The highly conserved nature of these hydrogen bonds, the regiospecificity of the one (always from the N3′ side) to the amino nitrogen, and of a carboxylic acid at position 51 prompted the authors to suggest that a model that accounts for these interactions is one in which the aminopyrimidine is tautomerized to the N1′H–N4′H imino tautomer [3]. This tautomerization would then endow the N4′-imino nitrogen with sufficient basicity to trigger proton abstraction from the thiazolium C2H, thereby initiating the catalytic cycle. An earlier model from Rutgers suggested such a reversal in the acid-base properties of the aminopyrimidine by showing that N1′-methyl-thiamin and analogues possess a low pK_a at the exocyclic amines in water, some 12.5 units [23]. One could then speculate that the E51 side chain simply stabilizes the proton at N1′, in essence accomplishing the same mission.

The E51 residue has been substituted in two laboratories, indicating that any substitution at E51 results in an at least 100-fold loss in specific activity [24, 25]. Interestingly, however, the E51Q, E51D, E51N and E51A variants are still active enzymes [25]. In a related experiment, it was shown that while N1′-methylthiaminium diphosphate is still bound to the E51A variant (even better than is ThDP), there is no activity produced as measured by steady-state kinetics. These results were interpreted to mean that the state of ionization in the E51COOH...N1′ couple changes during the catalytic cycle, the state of ionization at E51 cycles between the E51COOH and $E51COO^-$ forms. At the same time, the state of tautomerization also cycles between the aminopyrimidine and iminopyrimidine tautomeric forms (both being neutral). However, the aminopyrimidinium form with the positive charge fixed is functionally inactive. As to why the enzyme is still active even in the absence of E51, may be due to the natural pK_a for the N1′H-protonation of thiamin and ThDP, near 5.0. Even in the absence of the E51 there can still form some iminopyrimidine tautomer, perhaps 100–1000 times less than in the wild-type enzyme.

Non-conserved and semi-conserved acid-base groups proximal to ThDP. A systematic variation of the acid-base groups in the active center has been undertaken for several of the enzymes, both yeast and *Zymomonas mobilis* PDC, transketolase and BFD. While these groups are conserved in the two PDCs, they vary even in the functionally closely related BFD and certainly in TK. In yeast PDC, these are residues H114, H115, D28 and E477 [26]. Their systematic substitutions lead to decrease in steady-state parameters of 100-1000-fold. The pH dependence of the steady-state parameters is on the whole are rather uninformative with some rare exceptions: (a) for the D28A variant, there is a slight alkaline shift of the acid-limb of the k_{cat}-pH profile, suggesting something about the state of ionization of this group and the H115 to which it is hydrogen bonded; (b) for the H114F substitution only, there is evidence for participation in the substrate-activation mechanism (see next section).

According to conventional interpretations of steady-state enzyme kinetic properties, these four residues have a function mostly in transition state stabilization, both in some step(s) through decarboxylation and in some step(s) commencing with decarboxylation and culminating in product release. To gain further insight to the function of these residues, an elucidation of the carboligase side reaction leading to acetoin and acetolactate with the wild type and variant PDC was undertaken. Such studies provide insight to steps starting with decarboxylation. It is quite evident from the results that D28 and E477 both have major impacts on steps post-decarboxylation.

The variety of regulatory mechanisms. In addition to the hysteretic kinetics exhibited by the binding of ThDP.Mg(II) referred to above, both yeast PDC [27] and *E. coli* PDHc-E1 exhibit sigmoidal v_0 vs. $[S]$ behavior of varying magnitudes. Much more work has been carried out on this subject on PDC, and all species of this enzyme except for the one from *Zymomonas mobilis* appear to behave the same way. The magnitude of the Hill coefficient, as a measure of sigmoidicity is ca. 2.0 for the wild-type yeast PDC at pH 6.0 and diminishes especially as the acidity of the solution increases. The working hypothesis in the authors' labs was founded on the observation by others that cysteine modification reduced the sigmoidicity [28]. In a series of papers on the subject [29–35], we explored the hypothesis that of the four cysteines in yeast PDC (C69, C152, C221, and C222), C221, the only one readily accessible from the surface of the protein, is the site where substrate activation is triggered. In perhaps the most convincing experiment to date, substitution of Cys221 to Ser or Ala abolished the cooperativity. Since C221 is ca. 20 Å from the active center C2 atom, we undertook a program to map the information transfer pathway. So far, we have identified the pathway C221 → H92 → E91 → W412 and onto the pyrimidine imino NH (via the G413 backbone C=O) and most likely H114 as well. Of the active center residues discussed above, only H114 appears so far to have participation in any manner. With the impending disclosure of the PDHc-E1 and related E1 structures, the various reg-



Scheme 2

ulatory pathways of that enzyme will be further explored. The authors believe that the systematic identification of these regulatory mechanisms is important so that a better understanding of enzyme regulation in general will result.

Contributions to catalysis and overall mechanistic conclusions. As can be seen from the above short summary, there are several contributions from the protein to help account for the 10^{12}–10^{13}-fold rate accelerations that some have attributed to the protein over and above that provided by ThDP itself [36].

Our current working hypothesis assigns a significant fraction of this rate acceleration to the special bound conformation and hydrogen-bonding environment around the aminopyrimidine ring. These features are conserved in all ThDP enzymes examined so far. We believe that the purpose for this conservation is to assist with the deprotonation at C2 to form the highly reactive ylide/carbanion. Are the relevant pK_as balanced for rapid proton transfer? The pK_as for the conserved Glu across from the N1′ atom and at N1′ are likely to be similar, while the pK_a for ionization of the amino group at $N4'H_{5'}$ (denoting the proton bonded to N4' on the C5′ side of the pyrimidine ring), once the N1′ atom is protonated is ca. 12.5 [23], and for ionization of C2H is 17–19 [4]. The special environment of the V coenzyme conformation present in all of these enzymes [22] assures that for a reasonable distance between N4′ and C2, proton transfer would take place at a rate exceeding the turnover number for many such enzymes, i. e., 60–100 s^{-1}. This is important in view of the C-13 NMR report on PDC, indicating that C2H of the bound ThDP is undissociated at pH 6.0 [37].

How do ThDP enzymes solve this high pK_a problem? The following experiment provides some of the answers to this riddle. When the E91D variant of yeast apo-pyruvate decarboxylase (EC 4.1.1.1) was exposed to C2α-hydroxybenzylThDP (HBThDP), this putative intermediate was partitioned on PDC between release of the benzaldehyde product (evidenced by regeneration of active enzyme), and dissociation of the proton at C2α to form the enamine/C2α-carbanion intermediate (evidenced by the appearance of the visible spectrum of the intermediate) (Scheme 2). While the pK_a for this dissociation is $\sim$15.4 in water, formation of the enamine at pH 6.0 on the enzyme indicates a greater than 9 unit pK_a suppression by the enzyme environment [38]. The fluorescence emission properties of thiochrome diphosphate, a fluorescent ThDP analog and a competitive inhibitor for PDC, when PDC-bound resemble that observed in 1-pentanol and 1-hexanol, suggesting an apparent dielectric constant of 13–15 for the PDC active center. Such a low effective dielectric constant could account for much of the observed >9 unit pK_a suppression for ionization at the C2α position. The dramatic stabilization of this (and presumably other) zwitterionic intermediate(s) is sufficient to account for as much as a 10^9-fold rate acceleration on PDC, providing the bulk of the rate acceleration by the protein over and above that afforded by the coenzyme, perhaps on all ThDP-dependent 2-oxo acid decarboxylases.

Acknowledgements. Supported at Rutgers by NIH-GM-50380, the Rutgers Busch Biomedical Fund, the Rutgers Board of Governors Fund, Roche Diagnostics Inc., Indianapolis, IN and NSF-BIR94/13198; at Pittsburgh by NIH-GM-48195.

References

1. Jordan, F., Nemeria, N., Guo, F., Baburina, I., Gao, Y., Kahyaoglu, A., Li, H., Wang, J., Yi, J., Guest, J.R., and Furey, W. (1998) *Biochim. Biophys. Acta* **1385**, 287–306.
2. Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., and Jordan, F. (1993) *Biochemistry* **32**, 6165–6170.
3. Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D., and Jordan, F. (1996) *J. Mol. Biol.* **256**, 590–600.
4. Washabaugh, M.W. and Jencks, W.P. (1988) *Biochemistry* **27**, 5044–5053.
5. Barletta, G., Huskey, W.P., and Jordan, F. (1992) *J. Am. Chem. Soc.* **114**, 7607–7608.
6. Barletta, G., Huskey, W.P., and Jordan, F. (1997) *J. Am. Chem. Soc.* **119**, 2356–2362.
7. Zou, Y. (1999) Ph.D. Dissertation, Rutgers University Graduate Faculty at Newark.
8. Bordwell, F.G., Satish, A.V., Rios, C.B., Chung, A., and Jordan, F. (1990) *J. Am. Chem. Soc.* **112**, 792–797.
9. Jordan, F., Kudzin, Z.H., and Rios, C.B. (1987) *J. Am .Chem. Soc.* **109**, 4415–4417.
10. Kuo, D.J., and Jordan, F., (1983) *J. Biol. Chem.* **258**, 134l5–13417.
11. Zeng, X., Chung, A., Haran, M., and Jordan, F. (1991) *J. Am. Chem. Soc.* **113**, 5842–5849.
12. Menon-Rudolph, S., Nishikawa, S., Zeng, X., and Jordan, F. (1992) *J. Am. Chem. Soc.* **114**, 10110–10112.
13. Sergienko, Wang, Polovnikova, Hasson, McLeish, Kenyon, and Jordan, submitted.
14. Hasson, M.S., Muscate, A., McLeish, M.J., Polovnikova, L.S., Gerlt, J.A., Kenyon, G.L., Petsko, G.A., and Ringe, D. (1998) *Biochemistry* **37**, 9918–9930.
15. Muller, Y.A. and Schulz, G.E. (1993) *Science*, **259**, 965–967.
16. Reed, L. (1974) *Acc. Chem. Res.* **7**, 40–46.
17. Chabriere, E., Charon, M., Volbeda, A., Pieulle, L., Hatchikian, E., and Fontecilla-Camps, J. (1999) *Nature Struct. Biol.* **6**, 182–190

18. Lindquist, Y., Schneider, G., Ermler, U., and Sundstrom, M. (1992) *EMBO Journal*, **11** 3273–3279.
19. Dobritzsch, D., König, S., Schneider, G., and Lu, G. (1998) *J. Biol. Chem.* **273**, 20196–20204.
20. Hawkins, C.F., Borges, A., and Perham, R.N. (1989) *FEBS Lett.* **255**, 77–82.
21. Yi, J., Nemeria, N.S., McNally, A., Jordan, F., Machado, R.S., and Guest, J.R. (1996) *J. Biol. Chem.* **271**, 33192–33200.
22. Guo, F., Zhang, D., Kahyaoglu, A., Farid, R.S., and Jordan, F. (1998) *Biochemistry.* **37**, 13379–13391.
23. Jordan, F., Mariam, Y. H (1978) *J. Am. Chem. Soc.* **100**, 2534–2541.
24. Killenberg-Jabs, M., König, S., Eberhardt, I., Hohmann, S., and Hübner, G. (1997) *Biochemistry.* **36**, 1900–1905.
25. Gao, Yuhong (2000) Ph.D. Dissertation, Rutgers University, Newark, NJ 07102.
26. Liu, Sergienko, Guo, Wang, and Jordan, submitted.
27. G. Hübner, G., R. Weidhase, R., and Schellenberger, A. (1978) *Eur. J. Biochem.* **92** 175–181.
28. Hübner, G., König, S., and Schellenberger, A. (1988) *Biomed. Biochim. Acta* **47**, 9–18.
29. Zeng, X., Farrenkopf, B., Hohmann, S., Dyda, F., Furey, W., and Jordan, F. (1993) *Biochemistry* **32**, 2704–2709.
30. Baburina, I., Gao, Y., Hu, Z., Jordan, F., Hohmann, and Furey, W. (1994) *Biochemistry* **33**, 5630–5635.
31. Baburina, I., Moore, D.J., Volkov, A., Kahyaoglu, A., Jordan, F., and Mendselsohn, R. (1996) *Biochemistry* **35**, 10249–10255.
32. Baburina, I., Li, H., Bennion, B., Furey, W., and Jordan, F. (1998) *Biochemistry* **37**, 1235–1244.
33. Baburina, I., Dikdan, G., Guo, F., Tous, G.I., Root, B., and Jordan, F. (1998) *Biochemistry* **37**, 1245–1255.
34. Li, H., Furey, W., and Jordan, F. (1999) *Biochemistry* **38**, 9992–10003.
35. Li, H. and Jordan, F. (1999) *Biochemistry* **38**, 10004–10012.
36. Alvarez, F.J., Ermer, J., Hübner, G., Schellenberger, A., and Schowen, R.L. (1995) *J. Am. Chem. Soc.*, **117**, 1678–1683.
37. Kern, D., Kern, G., Neef, H., Tittmann, K., Killenberg-Jabs, M., Wikner, C., Schneider, G., and Hübner, G. (1997) *Science* **275**, 67–70.
38. Jordan, F., Li, H., and Brown, A. (1999) *Biochemistry* **38**, 6369–6373.