4. L. H. Weaver, W. R. Kester, B. W. Matthews, *J. Mol. Biol.* **114**, 119 (1977).
5. A. F. Monzingo and B. W. Matthews, *Biochemistry* **23**, 5724 (1984).
6. D. G. Hangauer, A. F. Monzingo, B. W. Matthews, *ibid.*, p. 5730.
7. S. J. Gardell, C. S. Craik, D. Hilverg, M. S. Urdea, W. J. Rutter, *Nature (London)* **317**, 551 (1985).
8. C.-M. Kam, N. Nishino, J. C. Powers, *Biochemistry* **18**, 3032 (1979) (also see references listed therein).
9. D. E. Tronrud *et al.*, *Eur. J. Biochem.* **157**, 261 (1986) (also see references listed therein).
10. P. A. Bartlett and C. K. Marlowe, *Biochemistry* **22**, 4618 (1983).
11. ———, *Science* **235**, 569 (1987).
12. M. F. Schmid *et al.*, *Acta Crystallogr. Sect. A* **37**, 701 (1981).
13. D. E. Tronrud, L. F. TenEyck, B. W. Matthews, *ibid.*, in press.
14. Analysis of the inhibitor ZG^{P}LL will be described in detail elsewhere in connection with the analysis of ZF^{P}LL, another peptide analog that binds in a somewhat different manner.
15. P. A. Bash *et al.*, *Science* **235**, 574 (1987).
16. P. V. Luzzatti, *Acta Crystallogr.* **5**, 802 (1952).
17. M. A. Holmes and B. W. Matthews, *J. Mol. Biol.* **160**, 623 (1982).
18. Figure 4 was generated by P. A. Bartlett and P. A. Bash from the coordinates furnished by B.W.M., using the programs MS (M. Connolly) and UCSF MIDAS at the Computer Graphics Laboratory, University of California, San Francisco (R. Langridge and T. Ferrin).
19. We thank P. A. Bartlett and C. K. Marlowe for the inhibitors and for discussions of the results obtained in Berkeley and in Eugene. We also thank T. Alber for discussions and comments. H.M.H. was supported in part by fellowship DRG-667 of the Damon Runyon–Walter Winchell Cancer Fund. The work was supported in part by grants from the National Science Foundation (PCM8312151), the National Institutes of Health (GM20066), and the M. J. Murdock Charitable Trust.

22 August 1986; accepted 4 December 1986

Calculation of the Relative Change in Binding Free Energy of a Protein-Inhibitor Complex

PAUL A. BASH,* U. CHANDRA SINGH,† FRANK K. BROWN,‡ ROBERT LANGRIDGE, PETER A. KOLLMAN

By means of a thermodynamic perturbation method implemented with molecular dynamics, the relative free energy of binding was calculated for the enzyme thermolysin complexed with a pair of phosphonamidate and phosphonate ester inhibitors. The calculated difference in free energy of binding was 4.21 ± 0.54 kilocalories per mole. This compares well with the experimental value of 4.1 kilocalories per mole. The method is general and can be used to determine a change or "mutation" in any system that can be suitably represented. It is likely to prove useful for protein and drug design.

A NEW METHOD HAS RECENTLY BEEN developed to calculate relative changes in free energy between interacting molecular species. It is based on a thermodynamic perturbation method (*1*) with Monte Carlo (*2*) or molecular dynamics (*3–6*) being used to change or "mutate" one molecule into another. It has been applied to a variety of small molecules to calculate relative changes in free energy of solvation within 1 kcal/mol of experimentally measured values (*2, 5, 6*), and to the conversion of benzamidine to *p*-fluorobenzamidine in trypsin (*7*). The results obtained suggest that the relative free energies of interaction between biological macromolecules and their ligands can be calculated by using a generalization of this perturbation method. We present a computational example to support this belief.

We selected thermolysin for this study because extensive structural and thermodynamic data are available for this protein and its inhibitors (i) Cbz-Gly^P-(NH)-Leu-Leu and (ii) Cbz-Gly^P-(O)-Leu-Leu. X-ray crystal structures of these complexes have been determined (*8*), and binding constants measured (*9*). These constants differ by three orders of magnitude, although the two inhibitors differ only by the replacement of an NH by an O and lead to complexes which are nearly identical [see Fig. 4 in the preceding report (*8*)].

We calculated Gibbs (G) free energy changes according to Eq. 1 (*1*)

$$G(\lambda + d\lambda) - G(\lambda) = -k_B T \ln < \exp\left(-\frac{H(\lambda + d\lambda) - H(\lambda)}{k_B T}\right) >_\lambda \quad (1)$$

where k_B is Boltzmanns' constant, T is the absolute temperature, $<>_\lambda$ indicates the ensemble or time average at intermediate positions along the conversion pathway characterized by the coupling parameter λ, and $H(\lambda)$ and $H(\lambda + d\lambda)$ are the Hamiltonians in the states λ and $\lambda + d\lambda$. In the application below, H is taken to be the interaction energy of the inhibitor with its surroundings. Molecular dynamics with a nonbonded cutoff of 8 Å was used to calculate the time average of the Hamiltonian differences from which ΔG between states A($\lambda = 0$) and B($\lambda = 1$) was evaluated by

$$\Delta G = \sum_{\lambda=0}^{\lambda=1} G(\lambda + \Delta\lambda) - G(\lambda) \quad (2)$$

The characterization of states A and B, and calculation of interaction energies, was done with an empirical force field, using standard AMBER force field values (*10*) except for charges on the inhibitor. Ab initio quantum mechanics, with the STO-3G* basis set, was used to determine electrostatic potential based charges (in atomic units) for both inhibitors (*11*). The main change was at the perturbation site where the amide NH ($q_N = -0.705$, $q_H = 0.227$) group on the leucine attached to the phosphate is converted into an oxygen ($q_O = -0.451$), while the charge on nearby atoms changed by less than 0.1. The ability to transform a system between two such states by using perturbation methods and molecular dynamics has been incorporated into the molecular simulation program AMBER (*12*) in a general way (*13*). Perturbation calculations are carried out by specifying the parameters in the beginning and end states, the number of intervals or "windows" between the states, and the length of time for equilibration and data collection in each window. The program automatically carries out the transformation between the two states during which intermediate energy values for each window are accumulated, stored, and reported. The program and methodology are described in more detail elsewhere (*6*).

Two calculations are required. The first determines the solvation free energy difference between the two inhibitors (ΔG_{solv}), and the second determines the free energy difference in the two inhibitor enzyme complexes (ΔG_{bind}). The difference ($\Delta G_{bind} - \Delta G_{solv}$) is equal to the difference in aqueous solution binding free energies of the two inhibitors, $\Delta\Delta G_{bind}$, which was found by Bartlett and Marlowe (*9*) to be 4.1 kcal/mol.

The phosphonamidate inhibitor, in its x-ray conformation taken from the complex, is placed in a box of 782 waters [Jorgensen TIP3P potential (*14*)] generated by a Monte Carlo simulation, and the system is equilibrated for 4 picoseconds at constant temperature and pressure under periodic boundary conditions (*15*) with SHAKE (*16*) being used to constrain all bond lengths to their equilibrium values. The perturbation is carried out in a series of 20 windows with a $\Delta\lambda = 0.05$. At each λ, 500 steps of equilibration (0.002-picosecond

Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143.

*Present address: Chemistry Department, Harvard University, Cambridge, MA 02138.
†Present address: Scripps Clinic and Research Foundation, La Jolla, CA 92037.
‡Present address: Smith Kline and French Laboratories, Philadelphia, PA 19101.

Table 1. Results of free energy evaluations in kilocalories per mole. A represents the amide states (phosphonamidate or *N*-methylphosphoramidate) and B their associated esters. (+) and (−) refer to free energies in the $\lambda + \Delta\lambda$ and $\lambda - \Delta\lambda$ directions. ΔG_{ave} is determined by averaging the absolute values for A→B and B→A, whereas $\Delta G_{bind} - \Delta G_{solv}$ is obtained by averaging $\Delta G_{bind} - \Delta G_{solv}$ (A→B) and $\Delta G_{solv} - \Delta G_{bind}$ (B→A). The standard deviations taken about these calculated values only provide an indication of the statistical fluctuations inherent in the methodology as applied herein and are independent of approximations in the energy function or errors in crystallographic coordinates.

System	ΔG(A→B)	ΔG(B→A)	ΔG_{ave}	$\Delta G_{bind} - \Delta G_{solv}$
		Protein/inhibitor		
Phosphonamidate solution (ΔG_{solv})	3.58 (+); −3.58 (−)	−3.30 (+); 3.28 (−)	3.44 ± 0.17	
Protein-inhibitor x-ray (ΔG_{bind})	7.32 (+); −7.33 (−)	−7.93 (+); 7.99 (−)	7.64 ± 0.37	4.21 ± 0.54
Protein-inhibitor model (ΔG_{bind})	7.87 (+); −7.96 (−)	−7.71 (+); 7.71 (−)	7.81 ± 0.14	4.38 ± 0.06
Experimental				4.10
		Phosphates		
Trimethylphosphate solution (ΔG_{solv})	0.27 (+); −0.20 (+)	−0.12 (−); 0.30 (−)	0.22 ± 0.08	
Experimental			0 to 1	

time steps) are followed by 1000 steps of data collection (0.001-picosecond time steps). The interaction energy of the inhibitor with the solvent at values of its parameters $(\lambda + \Delta\lambda)$ and $(\lambda - \Delta\lambda)$ are evaluated and designated as (+) and (−), respectively. This provides an internal check on the method since the sum of these two energy changes over all the intervals should be the same. The resultant free energy changes are 3.58 (+) and −3.58 (−). Using the final coordinates from this simulation, we carried out a second run for the B → A transition with resultant free energy changes of −3.30 (+) and 3.28 (−). This inverse transformation was done to test the dependency of the result on the starting configuration. Although the changes from A → B and B → A occurred over a period of only 40 picoseconds, our experience with other such transformations (*6*) suggests the time span is adequate.

The initial state for the complex is the x-ray structure of the phosphonamidate (A) in thermolysin (*8*). The enzyme is too large (more than 4000 atoms) to conveniently include all its atoms in the simulation, so we allow residues that contain atoms within 15 Å of the inhibitors' atoms to move. Other atoms are included in the determination of forces, but are fixed at their starting locations. To include the effects of solvation, we add a spherical "cap" of 168 water molecules within 15 Å of the center of the Cbz group in the inhibitor, and harmonic radial forces restrain any water leaving this 15 Å boundary (*17*). The resultant free energy changes are 7.32 kcal/mol (+) and −7.33 kcal/mol (−) for A → B, and −7.93 kcal/mol (+) and 7.99 kcal/mol (−) for B → A.

The average for the free energy changes of the inhibitor in the enzyme and in solution for both A → B and B → A is 4.21 ± 0.54 kcal/mol, and the experimental value for this pair is 4.1 kcal/mol. If one considers the complexity of the calculation, the quality of agreement is encouraging. Table 1 summarizes calculated and experimental free energy changes and the meaning of the error limits.

It is clear that our approach is very powerful when the differences in the two states are small enough that conformational space can be adequately sampled. To give an indication of the robustness of the method, the utility of molecular modeling and computer graphics combined with numerical calculations, and the sensitivity of the calculation to errors in the x-ray structure, we began with coordinates for thermolysin complexed with the irreversible inhibitor $ClCH_2CO$–DL–(N–OH)Leu-OCH_3 (*18*), and, using MIDAS (*19*), model-built the phosphonamidate inhibitor into its active site. The perturbation for this model-built amide to ester was done as described above, with a free energy change of 7.81 ± 0.14 kcal/mol, close to the 7.6 value for this process when the actual x-ray coordinates were used.

There are no direct experimental data to verify the individual correctness of these relative solvation and binding calculations; however, there are data on related neutral phosphate esters (*20*). We applied the above methods to the conversion of trimethyl phosphate to its associated amide with a resultant relative free energy change of 0.28 kcal/mol, in good agreement with the experimental value of 0 to 1 kcal/mol for tripropyl phosphate relative to dibutyl *N*-methylphosphoramidate (*20*).

A simple interpretation of the net difference of binding is as follows. In the active site, the NH interacts favorably with the carbonyl oxygen of Ala 113. The ester oxygen is forced to interact unfavorably with the carbonyl because the remaining strong interactions of the phosphate and hydrophobic residues of the inhibitor with the enzyme keep the ester oxygen (O) in the same place as the amide (NH). This results in a net free energy difference of the interaction between inhibitor and enzyme of 7.5 kcal/mol. The relative solvation free energy reduces this to 4 kcal/mol. In this way our calculations bring out two important features that have not always been adequately considered in comparing protein inhibitors: (i) the importance of both attractive and repulsive interactions and (ii) the importance of relative solvation effects in inhibitor binding. The difference in relative solvation free energies between neutral phosphoramidates and anionic phosphonamidates with respect to their esters is unexpected and deserves further study. However, the fact that the calculations reproduce experimental values for the neutral system (*20*) lends credibility to the calculated solvation free energy difference for the anions and was useful in the interpretation of the experimental binding energy (*9*).

This theoretical study and the experimental work described in the preceding reports represent a synergistic application of experiment and theory toward understanding enzyme-ligand interaction. Although such multifaceted approaches have been rare, they will become of increasing importance for the rational design of drugs and in protein engineering.

REFERENCES AND NOTES

1. R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
2. W. L. Jorgensen and C. Ravimohan, *ibid.* **83**, 3050 (1986).
3. J. P. M. Postma, H. J. C. Berendsen, J. R. Haak, *Faraday Symp.* **17**, 55 (1982).
4. B. L. Tembe and J. A. McCammon, *Comput. Chem.* **8**, 281 (1984).
5. T. Lybrand, J. A. McCammon, G. Wipff, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 833 (1986).
6. U. C. Singh, F. Brown, P. A. Bash, P. Kollman, *J. Am. Chem. Soc.*, in press.
7. C. F. Wong and J. A. McCammon, *ibid.* **108**, 3830 (1986).
8. D. E. Tronrud, H. M. Holden, B. W. Matthews, *Science* **235**, 571 (1987) and references listed therein.
9. P. A. Bartlett and C. K. Marlowe, *ibid.*, p. 569.
10. S. J. Weiner *et al.*, *J. Am. Chem. Soc.* **106**, 765 (1984).
11. U. C. Singh and P. A. Kollman, *J. Comp. Chem.* **5**, 129 (1984).
12. P. K. Weiner and P. A. Kollman, *ibid.* **2**, 287 (1981).
13. U. C. Singh, P. K. Weiner, J. W. Caldwell, P. A. Kollman, AMBER (UCSF), version 3.0 (Department of Pharmaceutical Chemistry, University of California, San Francisco, 1986).
14. W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
15. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *ibid.* **81**, 3684 (1984).
16. W. F. van Gunsteren and H. J. C. Berendsen, *Mol. Phys.* **34**, 1311 (1977).
17. C. L. Brooks, A. Brunger, M. Karplus, *Biopolymers* **24**, 943 (1985).
18. M. A. Holmes, D. E. Tronrud, B. W. Matthews, *Biochemistry* **22**, 236 (1983).

19. T. E. Ferrin, C. Huang, L. Jarvis, R. Langridge, *J. Mol. Graph.* **2**, 55 (1984).
20. R. Wolfenden and R. Williams, *J. Am. Chem. Soc.* **105**, 1028 (1983).
21. Supercomputer time was provided on the Cray-XMP/48 at the San Diego Supercomputer Center (SDSC), funded through the NSF supercomputer initiative, through grants awarded to P.A.B., R.L., and P. A. Bartlett (Berkeley) and NSF grant DMB-8419883 to P.A.K. Support from the NIH to P.A.K. (GM-29072) and to R.L. via the UCSF Computer Graphics Laboratory (NIH RR-1081) are acknowledged. We thank R. Hilderbrandt of SDSC for assistance in using the facilities, C. K. Marlowe for suggesting the thermolysin system for study, B. W. Matthews for the x-ray structures used in this study, and H. Berendsen, W. van Gunsteren, and P. A. Bartlett for useful discussions.

22 August 1986; accepted 4 December 1986

The Mitochondrial Genotype Can Influence Nuclear Gene Expression in Yeast

VANDANA S. PARIKH, MELISSA M. MORGAN, ROSE SCOTT, L. SCOTT CLEMENTS, RONALD A. BUTOW*

Isochromosomal, respiratory-deficient yeast strains, such as a mit$^-$, a hypersuppressive petite, and a petite lacking mitochondrial DNA, are phenotypically identical in spite of differences in their mitochondrial genomes. Subtractive hybridizations of complementary DNA's to polyadenylated RNA isolated from derepressed cultures of these strains reveal the presence of nuclear-encoded transcripts whose abundance varies not only between them and their respiratory-competent parent, but among the respiratory-deficient strains themselves. Transcripts of some nuclear-encoded mitochondrial proteins, like cytochrome c and the α and β subunits of the mitochondrial adenosine triphosphatase, whose abundance is affected by glucose or heme, do not vary. In the absence of major metabolic variables, yeast cells seem to respond to the quality and quantity of mitochondrial DNA and modulate the levels of nuclear-encoded RNA's, perhaps as a means of intergenomic regulation.

THE SYNTHESIS OF A NUMBER OF mitochondrial proteins in *Saccharomyces cerevisiae* can be affected by the levels of glucose (catabolite repression), oxygen, and heme (*1*). These effects are most evident by large changes in the cell's respiratory capacity that accompany these "metabolic" variables. Recent studies have focused on the molecular basis for regulation of a few mitochondrial proteins encoded in the nucleus, like cytochrome c (the product of the *CYC1* gene) (*2*), the α and β subunits of the mitochondrial F1 adenosine triphosphatase (F1 ATPase) (*3*), and some subunits of cytochrome oxidase (*4, 5*). The most detailed molecular information available on the regulation of expression of these genes is that for *CYC1*: control of cytochrome c synthesis is largely transcriptional and involves cis-acting upstream activation sites (UAS) that are analogous in some ways to higher eukaryotic enhancers (*2*); these sites are believed to control *CYC1* expression in response to glucose and heme levels through interactions with trans-acting transcriptional factors (*6*).

Since the biogenesis of functional mitochondria requires a large contribution from the nucleus, many nuclear genes might be regulated, and some, perhaps, independent of the metabolic variables noted above. To examine these questions, we have used complementary DNA (cDNA) subtractions to detect differentially expressed transcripts among isonuclear, derepressed yeast strains that differ only by the presence or absence of the kind of mitochondrial DNA (mtDNA) they contain. We show that some nuclear DNA sequences are differentially expressed between the respiratory competent cells (ρ^+) and isonuclear, respiratory-deficient derivatives, and others are differentially expressed among the respiratory-deficient cells themselves. These results suggest a hitherto undescribed type of nuclear-mitochondrial interaction whereby nuclear DNA sequences can respond to the state of the mitochondrial genome.

Experiments were carried out with four isonuclear yeast (*Saccharomyces cerevisiae*) strains all grown on raffinose, a fermentable but nonrepressing sugar. Three of the strains are respiratory-deficient with different lesions in their mitochondrial genomes; all are derived from COP161, a respiratory-competent ρ^+. One strain is a mit$^-$ (E69) containing a deletion in the *oxi2* gene (*7*); another strain is a hypersuppressive cytoplasmic petite (ρ^-), designated HS40, in which all but about 700 base pairs (bp) of the wild-type mitochondrial genome are deleted; the other strain is a ρ^0 petite that completely lacks mtDNA. Although mit$^-$ E69 is unable to respire because it lacks subunit III of cytochrome oxidase, it can still transcribe and translate its stable mitochondrial genome. The mitochondrial genome of HS40 ρ^- contains an *ori/rep* sequence, one of several closely related sequences believed to function as preferred origins of mtDNA replication (*8*); like other ρ^- petites, HS40 lacks mitochondrial protein synthesis, but it can transcribe sequences on its truncated genome. Thus, although these strains are phenotypically identical with respect to respiratory function, they contain different mtDNA lesions that give rise to the same respiratory-deficient phenotype.

To determine whether the steady-state levels of any cytoplasmic RNA's might be differentially regulated in these respiratory-deficient cells, cDNA subtractions were carried out between mit$^-$ E69 and the ρ^0 petite, and between the HS40ρ^- and ρ^0 petites. The former pair represents, in principle, the largest difference between cells with defective mitochondrial genomes; the latter pair represents a more stringent test for transcripts that might be differentially regulated among respiratory-deficient cells since both cell types are genotypically petite. In each case, single-stranded cDNA's were prepared by reverse transcription of polyadenylated [poly(A)$^+$] RNA fractions of cytoplasmic RNA isolated from cultures in the mid-logarithmic phase of growth, and hybridized with the heterologous poly(A)$^+$ RNA fraction to a $R_o t$ of about 4000. Under these conditions the cDNA's of low abundance transcripts that are not differentially expressed should be quantitatively driven into hybrid. The single-stranded cDNA's were isolated by hydroxyapatite chromatography and cloned into the Pst I site of plasmid pBR322 after second-strand synthesis and GC-tailing (*9*).

The cloned cDNA inserts were initially screened by differential colony hybridization with ^{32}P-labeled single-stranded cDNA probes; these were prepared by reverse transcription of poly(A)$^+$ RNA fractions isolated from the ρ^+, and each of the respiratory-deficient strains. Plasmid DNA's isolated from positive clones were then screened by dot-blot hybridization with the same single-stranded cDNA probes. False positives and cDNA's derived from mtDNA were screened out, and selected cDNA clones were analyzed by Northern blotting of poly(A)$^+$ RNA. The input RNA's were normalized to the amount of actin message present in each preparation. Actin, which is an essential product in yeast required for

Department of Biochemistry, Division of Molecular Biology, University of Texas Health Science Center at Dallas, 5323 Harry Hines Boulevard, Dallas, TX 75235.

*To whom all correspondence should be sent.