Glycogen phosphorylase inhibitors: A free energy perturbation analysis of glucopyranose spirohydantoin analogues

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · DECEMBER 2005

Impact Factor: 2.63 · DOI: 10.1002/prot.20641 · Source: PubMed

8 AUTHORS, INCLUDING:

Kimberly A Watson
University of Reading
55 PUBLICATIONS **1,397** CITATIONS
SEE PROFILE

Evangelia D Chrysina
National Hellenic Research Foundation
50 PUBLICATIONS **1,029** CITATIONS
SEE PROFILE

Spyros E Zographos
National Hellenic Research Foundation
65 PUBLICATIONS **1,463** CITATIONS
SEE PROFILE

PROTEINS: Structure, Function, and Bioinformatics 61:984–998 (2005)

Glycogen Phosphorylase Inhibitors: A Free Energy Perturbation Analysis of Glucopyranose Spirohydantoin Analogues

G. Archontis,[1]* K. A. Watson,[2] Q. Xie,[1] † G. Andreou,[1] E. D. Chrysina,[3] S. E. Zographos,[3] N. G. Oikonomakos,[3,4] and M. Karplus[5,6]*

[1]*Department of Physics, University of Cyprus, Cyprus*
[2]*Structural Biology Unit, the Biocentre, University of Reading, Whiteknights, Berkshire, United Kingdom*
[3]*Institute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation, Athens, Greece*
[4]*Institute of Biological Research and Biotechnology, National Hellenic Research Foundation, Athens, Greece*
[5]*Laboratoire de Chimie Biophysique, ISIS, Université Louis Pasteur, Strasbourg, France*
[6]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts*

ABSTRACT **GP catalyzes the phosphorylation of glycogen to Glc-1-P. Because of its fundamental role in the metabolism of glycogen, GP has been the target for a systematic structure-assisted design of inhibitory compounds, which could be of value in the therapeutic treatment of type 2 diabetes mellitus. The most potent catalytic-site inhibitor of GP identified to date is spirohydantoin of glucopyranose (hydan). In this work, we employ MD free energy simulations to calculate the relative binding affinities for GP of hydan and two spirohydantoin analogues, methyl-hydan and n-hydan, in which a hydrogen atom is replaced by a methyl- or amino group, respectively. The results are compared with the experimental relative affinities of these ligands, estimated by kinetic measurements of the ligand inhibition constants. The calculated binding affinity for methyl-hydan (relative to hydan) is 3.75 ± 1.4 kcal/mol, in excellent agreement with the experimental value (3.6 ± 0.2 kcal/mol). For n-hydan, the calculated value is 1.0 ± 1.1 kcal/mol, somewhat smaller than the experimental result (2.3 ± 0.1 kcal/mol). A free energy decomposition analysis shows that hydan makes optimum interactions with protein residues and specific water molecules in the catalytic site. In the other two ligands, structural perturbations of the active site by the additional methyl- or amino group reduce the corresponding binding affinities. The computed binding free energies are sensitive to the preference of a specific water molecule for two well-defined positions in the catalytic site. The behavior of this water is analyzed in detail, and the free energy profile for the translocation of the water between the two positions is evaluated. The results provide insights into the role of water molecules in modulating ligand binding affinities. A comparison of the interactions between a set of ligands and their surrounding groups in X-ray structures is often used in the interpretation of binding free energy differences and in guiding the design of new ligands. For the systems in this work, such an approach fails to estimate the order of relative binding strengths, in contrast to the rigorous free energy treatment. Proteins 2005;61:984–998.**

Key words: molecular dynamics simulations; free energy calculations; protein–ligand interactions; glycogen phosphorylase; spirohydantoin analogues

INTRODUCTION

GP catalyzes the degradative phosphorylation of glycogen to Glc-1-P, which is used in muscle tissue to provide

Abbreviations: GP, glycogen phosphorylase; Glc-1-P, glucose-1-phosphate; hydan, (2R,3R,4S,5R,6R)-3,4,5-trihydroxy-2-hydroxymethyl-7,9-diaza-1-oxa-spiro[4,5]decane-8,10-dione; MDFE, molecular dynamics free energy; methyl-hydan, (2R,3R,4S,5R,6R)-3,4,5-trihydroxy-2-hydroxymethyl-9-methyl-7,9-diaza-1-oxa-spiro[4,5]decane-8,10-dione; MMFF, Merck Molecular Force Field; MM/PBSA, molecular mechanics/Poisson–Boltzmann surface area; n-hydan, (2R,3R,4S,5R, 6R)-9-amino-3,4,5-trihydroxy-2-hydroxymethyl-7,9-diaza-1-oxa-spiro[4,5]decane-8,10-dione; PB, Poisson–Boltzmann; PP1, protein phosphatase 1; TIP3P, three-point transferable intermolecular potential.

Grant sponsor: University of Cyprus, through the program "From Strong Interactions to Molecular Recognition; Theoretical and Computational Studies" (G. Archontis and Q. Xie). Grant sponsor: A. G. Leventis Foundation, through the program "Regulation of Glycogen Metabolism and Oxidative Repair of DNA by Molecular Association; Insights From Simulations and Experiment" (G. Archontis). Grant sponsor: Cyprus Research Promotion Foundation, through the grant "Molecular Recognition and Rational Drug Design of New Drugs," awarded in the "First Program for the Scientific and Technological Cooperation between Greece–Cyprus" (G. Archontis and N. G. Oikonomakos). Grant sponsor: Greek General Secretariat for Research and Technology; Grant numbers: PENED-204/2001 (to E. D. Chrysina) and ENTER-EP6/2001 (to N. G. Oikonomakos). Grant sponsor: Lister Institute of Preventive Medicine (personal research fellowship to K. A. Watson). Grant sponsor: National Institutes of Health (for the work done at Harvard).

*Correspondence to: G. Archontis, Department of Physics, University of Cyprus, PO20537, Cy1678, Cyprus. E-mail: archonti@ucy.ac.cy and M. Karplus, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138. E-mail: marci@tammy.harvard.edu

†Now in the Concord Consortium, Concord, MA 01742

Received 15 January 2005; Accepted 4 May 2005

Published online 21 October 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20641

the energy needed to sustain muscle contraction; in the liver, Glc-1-P is converted to glucose and provides energy for other tissues.[1] Glucose acts as a regulator of glycogen metabolism in the liver in two ways. First, it binds weakly to the catalytic site of GP (with an inhibition constant of 1.7 and 7.4 m*M* for α- and β-D-glucose, respectively[2]), and stabilizes the T state (the less active state) of the enzyme. In addition, glucose binds to a phosphorylated form of the enzyme (GPa), which normally exists in the active (R) state. Glucose binding induces a conformational transition of GPa from the R state to the inactive T state, which in turn is dephosphorylated to produce the GPb form. The reaction of dephosphorylation that inactivates the enzyme is catalyzed by PP1, an enzyme that is regulated in response to insulin. The conversion of GPa to GPb relieves the allosteric inhibition that GPa exerts on the glycogen-associated PP1, which converts glycogen synthase D to the I form, thus allowing the phosphatase to stimulate the synthesis of glycogen (see Oikonomakos[1] and references therein).

Thus, design of glucose-based inhibitors of GPb could help shift the balance from glycogen degradation to glycogen synthesis, and lead to compounds that suppress the level of glucose in the blood. This would be of therapeutic value for the treatment of the non-insulin-dependent form of diabetes mellitus.

Systematic work involving structure-assisted design, synthesis, kinetic characterization, and X-ray structure determination has led to a database of over 80 well-characterized glucose-analogue catalytic site inhibitors of GPb,[1,3–10] as well as inhibitors that bind to other regulatory sites[11–15] (for a comprehensive review of the efforts to design inhibitory compounds of GP, see Oikonomakos[1] and references therein). Common kinetic and structural features of these compounds are the strong selectivity for GPb, the stabilization of the T-state enzyme upon binding, and competitive inhibition with respect to the substrate Glc-1-P.[4]

One of the most potent catalytic-site inhibitors of GPb is the spirohydantoin of glucopyranose (hydan), with a K_i value that is approximately 550 times lower than the corresponding value for the native ligand α-D-glucose.[1,4–5] The chemical structure of this compound is shown in Figure 1. A high-resolution (1.8 Å) crystal structure of the GP-hydan complex has been determined by Gregoriou et al.[6] at a temperature of 100 K. Moreover, a number of spirohydantoin analogues have been synthesized, and their binding to GP has been studied by kinetic and crystallographic methods.[4,7] The crystallographic structures of the various complexes have been determined at room temperature at a resolution of 2.3–2.4 Å.

The spirohydantoin analogue inhibitors bind in approximately the same position in the GP active site (see companion article).[7] Examination of the structures of the complexes suggests that the stronger binding of these inhibitors, compared to glucose, results at least in part from improved interactions with the protein catalytic site and particularly with the main-chain oxygen of His377.[6,7] A comparison between the water B-factors in the free and

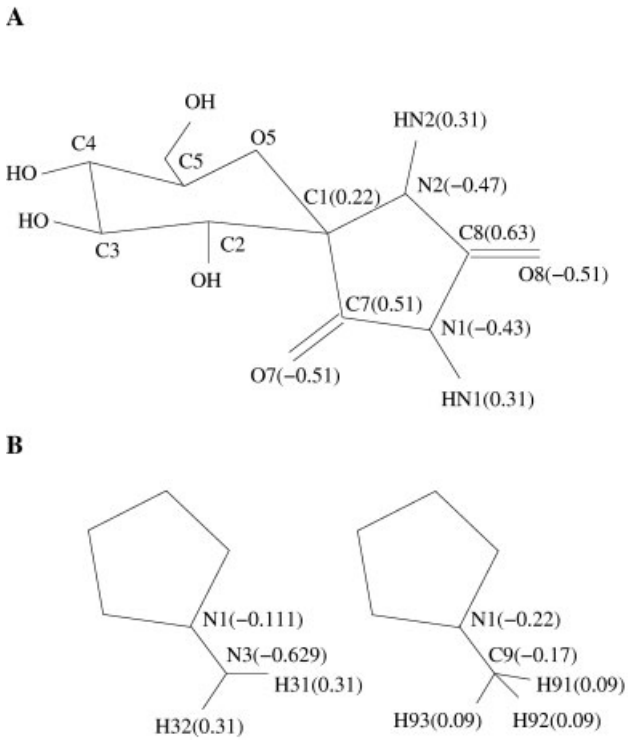

Fig. 1. Chemical structures of the hydantoin analogues studied in this work. (**A**) Hydan. The atomic charges of the five-member spirohydantoin moiety are indicated in parentheses. (**B**) N-Hydan (left) and Methyl-Hydan (right). Only the spirohydantoin moieties are shown, for clarity. The atomic charges (where different from hydan) are indicated in parentheses.

complexed enzyme suggests that the inhibitor binding also stabilizes the water network in the protein active site.[1,6] Furthermore, the rigid stereochemistry of the spirohydantoin group implies that the conformational entropy loss upon binding of hydan (and its analogues) is small, which contributes to its enhanced affinity relative to other ligands.[5] However, there are no detailed simulation studies of these systems at present. To supplement the insights from these structural studies, we have performed free energy difference simulations for spyrohydantoin and two of its analogues. The results are reported in this article.

In past years, several simulation methods have been developed for the calculation of absolute and relative binding affinities of ligands to macromolecules.[16–33] The MDFE simulation method is one of these. This method represents the entire, or part of the biomolecular system and surrounding solvent in atomic detail,[18,22–24] and provides the most accurate computational method for the evaluation of relative binding affinities of different ligands to a protein.[16–18] More approximate methods, which are of primary interest for surveying a large series of possible ligands, also have been used. They include MDFE simulations with an implicit (generalized Born) representation of the solvent,[28] simulations combined with postprocessing using the PB[22] or MM/PBSA,[29–30] and calculations that utilize empirical, linear response theory–based formulas.[16,31–33]

In this study we employ MDFE simulations to evaluate the relative binding free energies of the complexes between GP and three spirohydantoin compounds; hydan and the analogues methyl-hydan and n-hydan. The two analogues differ from hydan in that the hydrogen atom connected to atom N1 of the planar hydantoin ring (see Fig. 1) is replaced by a methyl- and amino group, respectively. The structure of the complexes between GP and these ligands has been determined at room temperature by X-ray crystallography.[7] Kinetic measurements of the corresponding inhibition constants show that hydan is the most potent inhibitor, with a K_i of 3.1 μM.[5] The methyl- and n-hydan compounds have K_i values of 1200 μM and 146 μM,[7] yielding respective binding affinities relative to hydan of 3.6 kcal/mol and 2.3 kcal/mol for the two compounds. The calculated binding affinity of methyl-hydan (relative to hydan) is 3.75($\pm$1.4) kcal/mol, in very good agreement with the experimental value. For n-hydan, the calculated value is 1.0($\pm$1.1) kcal/mol, somewhat smaller than the experimental result.

A free energy decomposition analysis gives insights concerning the microscopic origin of the differences in binding affinities of the three ligands. The best ligand (hydan) makes optimum interactions with protein residues and water molecules in the catalytic site. In the other two ligands, the added methyl or amino group have steric interactions with the catalytic site that reduce the binding affinity. Interestingly, the calculations indicate that a specific water molecule can occupy two well-defined positions in the catalytic site. The behavior of this water is analyzed in detail, and the free energy profile for the translocation of the water between the two positions is calculated to confirm the analysis.

In the next section we describe the system and present the methods employed in the simulations and the free energy computations. The results are given in the following section. We describe the dynamical behavior of the different complexes and present the free energy analysis of the relative binding affinities of the three ligands. We also present the free energy component results and the free energy profile of the special water molecule. The final section discusses the results and summarizes the conclusions.

METHODS

The Microscopic Model

A stochastic boundary simulation system[34] was used for the protein–ligand complexes. It consists of the catalytic site region of one monomer of GP, one ligand molecule, and explicit water molecules. Initial protein heavy atom coordinates were taken from the crystal structure of the protein–hydan complex, determined at 2.36 Å resolution at a temperature of 300 K[7] (accession code: 1GGN). An approximately spherical region of radius 20 Å was selected, centered on atom C1 of the hydan ligand. The model included 74 water molecules that were visible in the crystal structure. Initial hydrogen atom coordinates were placed with the HBUILD facility of CHARMM.[35] Additional water molecules were placed by overlaying a pre-equilibrated 24 Å sphere, and deleting water that extended beyond the distance of 20 Å, or overlapped with heavy atoms of the protein–ligand complex; a distance of 2.8 Å between the O atom of a water molecule and the protein heavy atoms was used. This overlaying procedure was repeated two times during the equilibration, yielding a total of 121 water molecules, in addition to the crystallographic waters. The additional overlays are needed because the water tends to shrink due to the electrostriction arising from interaction with charged residues. The final protein–hydan complex model consisted of 4418 atoms.

The complexes with the other ligands contained the same protein atoms and the same number of water molecules as the hydan complex. The crystallographic structures of these complexes are nearly identical with that of the hydan complex.[7]

The solution system consisted of a ligand immersed in a 20 Å sphere of 1244 explicit water molecules. The ligands were kept at the center of the solution sphere by applying a weak harmonic restraint on the C1 atom. This restraint has a very small effect on the ligand dynamics. Since the restraint term does not appear in the Hamiltonian derivative, it does not contribute to the transformation free energy. Also, the free energy correction for removing the restraint should be the same at the two end points of the transformation; hence, it should not contribute at all to the free energy.

Force Field and Simulation Method

All calculations were performed with the CHARMM program,[36] version c28b1. Atomic charges, van der Waals, and force field parameters for the protein were taken from the CHARMM22 all-atom parameter set.[37] The water molecules were represented by a modified TIP3P model.[38] The sugar portion of the spirohydantoin ligands was modeled using the CHARMM22 glucose parameters. To parameterize the charges of the atoms that form the spirohydantoin ring of the ligands we employed the MMFF option of CHARMM.[39] The MMFF-derived charges for the peptide bond–like CO and NH groups of the rings, the NH_2 group of the n-hydan, and the CH_3 group of the methyl-hydan were adjusted to values consistent with the ones used for analogous groups in the CHARMM22 parameterization (respectively: peptide CO/NH group, Asn side-chain, and sp3 carbons with nonpolar hydrogens[37]). The charges used are indicated in Figure 1. The van der Waals parameters were taken from the CHARMM22 parameters of analogous atoms.

Electrostatic interactions were treated without truncation by use of a multipole expansion approximation (extended electrostatics)[40] for groups more than 12 Å apart. The van der Waals interactions were switched off at interatom distances beyond 12 Å. A buffer region was defined at distances between 15 Å from the center of the sphere and the end of the system (20 Å). The protein and water atoms in the buffer region obeyed Langevin dynamics, and were subjected to random and frictional forces that correspond to a thermal bath at 293 K.[41] Protein heavy atoms in the buffer region were restrained to their equilib-

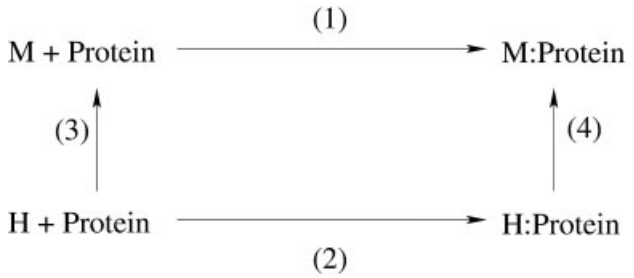

Fig. 2. Thermodynamic cycle for the relative-binding calculations in this work (see Methods section). Each transformation [e.g., H → M between dual-topology ligands containing a fully grown hydan moiety (H end) and a fully grown methyl-hydan moiety (M end)] is carried out in the protein complex (4) and in a solution sphere (3). Experiments measure the free energies of binding [processes (1) and (2)]. Because the free energy is a state function, the experimental relative free energy of binding $\Delta A_1 - \Delta A_2 = \Delta A_4 - \Delta A_3$. Thus, the double free energy difference $\Delta\Delta A_{43} = \Delta A_4 - \Delta A_3$ (obtained from the simulations) yields the relative free energy of binding (ΔA_{12}) of the M versus the H ligand to the protein.

rium crystallographic positions by mild harmonic restraints. The strength of the restraining potentials was adjusted to reproduce atomic fluctuations, consistent with the crystallographic B-factors. The water was subjected to a stochastic boundary potential.[34] The lengths of bonds containing hydrogens and the internal geometry of the water molecules was constrained with the SHAKE algorithm,[42] as implemented in CHARMM.

Thermodynamic Cycle and Mutation Protocol

The spirohydantoin ligands considered in this work differ by a single chemical group (see Fig. 1).

To evaluate the relative binding affinity of two ligands, we simulate the alchemical transformation of one ligand into the other, both in the protein complex and in its unbound, solvated state. Each transformation is associated with a free energy difference, which is computed by using the MD trajectory and a statistical mechanical formula involving averages over the trajectory. The relative binding affinity is equal to the difference in the two free energies. This calculation is explained with the help of the thermodynamic cycle, presented in Figure 2; see the caption for an explanation.

To simulate the transformation of one ligand to the other, we use the "dual topology" scheme, in which a part of the ligand containing the changing groups is duplicated.[25,43] Each hybrid ligand contained a single copy of the sugar moiety, and two separate copies of the planar, five-member spirohydantoin ring, linked to each other on atom C1 (see Fig. 1). Each copy consisted of atoms C7, O7, N1, C8, O8, N2, HN2, and either the initial- or the final-state chemical group, attached to atom N1.

In each free energy simulation the total system was subdivided into three "blocks" (B. Tidor, unpublished data). The first block [block 1 in Eq. (1)] contains the part of the system (water, protein, and the invariant portion of the hybrid ligand) that do not change in the mutation. The other two blocks (2 and 3) contain, respectively, the two copies with the initial and final state chemical groups. The function describing intra- and interatomic interaction energies was of the form

$$H(\lambda) = H^e + H^b + H^{nb}_{11} + (1-\lambda)H^{nb}_{12} + (1-\lambda)H^{nb}_{22} + \lambda H^{nb}_{13} + \lambda H^{nb}_{33}. \quad (1)$$

In Eq. (1) we denote by H^{nb}_{xy} the nonbonded (dispersion and electrostatic) interaction terms between blocks x and y. These terms were scaled linearly by a parameter λ. By varying λ from 0 to 1, the potential energy function changes from the functional form that corresponds to the initial state to the one of the final state. The term H^e includes the stochastic boundary potential used to retain the water in the simulation sphere, and harmonic-restraint potential terms.[44] The term H^b includes all bonded energy terms of the force field (bond stretching, angle bending, torsional and improper-dihedral terms) for the protein and ligand; they are not scaled in the alchemical transformation of the ligand.[43]

The two end states correspond to fictitious molecules in which the atoms of blocks 2 and 3 coexist, but only the nonbonded interactions between block 3 and 1 (initial state), or block 2 and 1 (final state) are switched off. Thus, in the initial (final) state, the atoms of block 3 (block 2) constitute ideal-gas moieties; the atoms within each moiety interact with each other and with the glucose ring (block 1) with bonded-energy terms. Due to the very small flexibility of the spirohydantoin fused-ring structure, the motion of each ideal-gas moiety is very limited, both in the protein complex and in solution. Thus, end point corrections associated with turning off the bonded-energy terms of atoms in blocks 2 and 3 are expected to approximately cancel in the double free energy difference between protein and solution, including contributions from dihedral angle terms between atoms in block 1 and atoms in the two ideal-gas moieties. For a detailed discussion of these end point corrections, we refer the readers to Boresch and Karplus[45,46] (i.e., see pages 107 and 115,[45] and pages 130 and 135[46]).

Two alchemical transformations were studied. In the transformation H → M, a dual topology ligand with one hydan and one methyl-hydan moiety was employed; the states with a fully grown hydan (methyl-hydan) moiety are denoted, respectively, by H and M. In the second transformation (H → N), the ligand had one hydan and one n-hydan moiety.

To test the convergence of the free energy calculations, the mutation H → N was also carried out along a two-step pathway. In the first step, the charges of atoms N1, HN1, C8, and O8 in the spirohydantoin ring of the hydan ligand were switched off, yielding an intermediate, single-topology ligand that is denoted by H′ below. In the second step, the H′ → N transformation was carried out with a dual-topology ligand containing the H′ and N moiety.

Free Energy Calculations

The free energy change ΔA was calculated by the thermodynamic integration formula[25]

$$\Delta A = \int_0^1 d\lambda \left\langle \frac{\partial \Delta H}{\partial \lambda} \right\rangle_\lambda = \Sigma_i (\lambda_{i+1} - \lambda_i) \left\langle \frac{\partial \Delta H}{\partial \lambda} \right\rangle_{\lambda_i}. \quad (2)$$

Simulations were conducted at λ values of 0.002, 0.01, 0.02, 0.06, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.94, 0.98, 0.99, and 0.998. For each value of λ we performed an equilibration, followed by a data collection run. For various windows, the equilibration was 20–100 ps and the production 60–400 ps. The typical length of a complete free energy run was 800–1600 ps.

The free energy change was calculated by use of Eq. (2). In the region 0.06–0.94, a trapezoeidal method was used. The end points $\lambda \to 0, 1$ correspond to the creation or elimination of a group of atoms. When the appearing (or disappearing) potential of the particle has a Lennard–Jones repulsive functional form $\sim r^{-12}$ and is varied in a linear manner, as in this work, the free energy derivative diverges with a power law dependence $\lambda^{-0.75}$.[47,48] This is observed here in some end points; it is most apparent in the cases where the created group is exposed to the environment. For example, in the H → M transformation, the derivative is several hundred kcal/mol at the λ = 0.002 point (appearing methyl group), but only a few kcal/mol at the final λ = 0.998 point, where the disappearing H atom is shielded by the existing, almost fully grown methyl group. In the diverging end points the derivatives are fitted to a power law of the form $\lambda^{-0.75}$; the corresponding free energy change is obtained by the integral of the derivative. To obtain an accurate dependence of the free energy derivative on λ at the diverging end points, we perform simulations at several points with λ values close to 0 (or 1), as explained above; see also the Results section.

The free energy transformations H → H′ corresponded to switching off (direction H → H′), or on (direction H′ → H) the partial charges of atoms N1, HN1, C8, and O8 in the hydan ligand. During the course of this transformation, the charges were scaled linearly by a coefficient λ with values λ = 0.05, 0.10, … , 0.90, 0.95.

Free Energy and Energy Component Analysis

The free energies obtained by the thermodynamic integration method [Eq. (2)] were decomposed into contributions from different interaction energy terms (i.e., electrostatic and van der Waals) and individual residues and water molecules. This decomposition was achieved by expressing the variable part of the potential energy function ΔH as a sum of underlying energy components. The integral becomes a sum of terms, each containing a specific energy component.[26,27,49]

It is also possible to identify, in part, the contribution of different components to the energy change in transforming one ligand to another. We write the total energy change in the transformation as

$$\Delta E = \langle H_0 + \Delta H \rangle_1 - \langle H_0 \rangle_0 = \sum_{res} \langle \Delta H_{\mathrm{res}} \rangle_1 + [\langle H_0 \rangle_1 - \langle H_0 \rangle_0] = \sum_{res} \Delta E_{\mathrm{res}} + [\langle H_0 \rangle_1 - \langle H_0 \rangle_0]. \quad (3)$$

In the above equation, H_0 and $H_0 + \Delta H$ are, respectively, the total energy functions of the initial (0) and final (1) states, and $\langle \ldots \rangle_{0(1)}$ denote expectation averages evaluated over the initial(final) state. The first term on the right-hand side of the last equality is a sum over the *direct* contributions to the total energy change from different residues, $\Delta E_{res} \equiv \langle \Delta H_{res} \rangle_1$. In addition to its explicit appearance in the terms ΔE_{res}, each residue also contributes indirectly to the mean energy $\langle H_0 \rangle_1 - \langle H_0 \rangle_0$ through the Boltzmann factor that affects all expectation values.

For Eq. (3) to be useful, it is best to choose the end point "1," where the terms ΔE_{res} are evaluated, so that divergences are avoided (e.g., the complex between GP and the M or N ligands, where the system is equilibrated in the presence of an additional, fully grown methyl- or amino group). Moreover, with this selection, the terms ΔE_{res} express the contribution to the total energy change from the interaction between specific residues and the inserted groups.

Free Energy Profile of Water X4

The computed binding free energies are sensitive to the preference of a specific water molecule to occupy two well-defined positions in the catalytic site. Following the nomenclature of Gregoriou et al.,[6] this water is denoted as X4 in what follows. In the first position ("position 1"), it forms a hydrogen bond with Asp283; in the second position ("position 2"), it interacts with Glu88. In the simulations of the hydan complex, position 1 is thermodynamically dominant, whereas in the other two complexes, X4 is observed most of the time at position 2. In the free energy transformation runs H → M and H → N, X4 switches from position 1 (H end) to 2 (M or N end). In the crystallographic structures, water molecules are also observed in both positions, depending on the experimental temperature and the complex (see the Results section).

The proximity of X4 to residues Asp283 and Glu88 can be adjusted by restraining the distances $r_{O_w C_\gamma}$ and $r_{O_w C_\delta}$, between the oxygen atom O_w of X4 and the atoms C_γ (Asp283) and C_δ (Glu88), respectively. In position 1(2), the distances $r_{O_w C_\gamma}$ and $r_{O_w C_\delta}$ have mean values of ≈ 3.75 Å (5.0 Å) and 4.6 Å (3.48 Å) respectively, observed in the simulations of the hydan complex. To calculate the free energy profile of the X4 motion between the two positions, we performed 30 simulations in which the two distances were varied in a concerted manner along a path linking the two positions, with reference distances changing from 3.67–5.08 Å ($r_{O_w C_\gamma}$) and 4.68–3.4 Å ($r_{O_w C_\delta}$). Between neighboring windows, the reference distances differed by 0.04–0.05 Å. The employed restraints corresponded to a harmonic potential with a force constant set to 200 kcal/mol; with this force constant, the simulations yield mean-square fluctuations of the restrained distances on the order of ≈ 0.07 Å. Thus, there was sufficient overlap of water positions between successive windows to obtain a meaningful potential of mean force. For each restrained distance we performed 10 ps of equilibration and 40 ps of production dynamics; the full profile corresponded to a 1.5 ns simulation.

At position i, the restraining potential had the form $V_{\mathrm{res}}(\vec{r}_1,\vec{r}_2;d_1(i),d_2(i)) \equiv k(|\vec{r}_i| - d_1(i))^2 + k(|\vec{r}_2| - d_2(i))^2$

where $\vec{r}_1$ and $\vec{r}_2$ were the instantaneous distance vectors $\vec{r}_{OwC_\gamma}$ and $\vec{r}_{OwC_\delta}$, and $[d_1(i), d_2(i)]$ was the pair of reference distances employed at simulation i. At each window i, the biased probability $\rho_b(|\vec{r}_1|,|\vec{r}_2|;i)$ was calculated directly from the simulation, as a function of the distances $|\vec{r}_1|$ *and* $|\vec{r}_2|$; the total unbiased profile $\rho(|\vec{r}_1|,|\vec{r}_2|)$ was obtained after merging the biased profiles of all windows and removing the effect of distance-restraints by the WHAM method,[50] as implemented in the CHARMM program.[51] The free energy is linked to the unbiased probability by the expression $F(|\vec{r}_1|,|\vec{r}_2|) = -k_BT \ln\rho(|\vec{r}_1|,|\vec{r}_2|)$. Using the values $(|\vec{r}_1|,|\vec{r}_2|) \equiv [d_1(i),d_2(i)]$ in the free energy expression, we obtain the free energy profiles shown in the Results section.

RESULTS

End State Simulations

The dynamical behavior of the bound ligands, the protein residues, and the water in the catalytic site were studied by MD simulations of all complexes. For each complex, a 2-ns simulation, as well as several shorter simulations (several hundred picoseconds each) with different initial conditions were conducted, and the resulting interactions were compared with those of the native ligand.

The biological ligand (glucose) binds at the catalytic site of GP. Glucose is well solvated in water, forming hydrogen-bonding interactions with water molecules through its peripheral hydroxyl groups; in the catalytic site, these hydroxyl groups are hydrogen-bonded to protein residues.[2,52] There is only one charged group (Glu672) involved in polar contacts, and no aromatic/glucopyranose interactions, such as the ones that characterize the high-affinity periplasmic monosaccharide binding proteins. These aspects of the binding interactions explain the 10^4 lower affinity of α-D-glucose for GP (K_i = 1.7 mM), relative to that of the high-affinity periplasmic monosaccharide binding proteins.[3,52]

Interactions of the ligand sugar moiety

The sugar moieties of the different ligands bind approximately in the same position and orientation as glucose in the GP: glucose complex.[7] A typical MD geometry of the GP:hydan catalytic site depicting the ligand sugar moiety and the surrounding protein residues and water molecules is shown in Figure 3(A). Very similar interactions are observed in all ligand complexes. The hydroxyl group 2-OH forms stable hydrogen bonds with the side-chain of Asn284, and with a water molecule [X2 in Fig. 3(A)]. The hydroxyl group 3-OH forms stable hydrogen bonds with residues Glu672 and Ser674, and interacts with the main-chain amide group of Gly675. The 4-OH group interacts with Gly675 and forms a stable hydrogen bond with a water molecule [X3 in Fig. 3(A)]. The hydroxyl group 6-OH forms a stable hydrogen bond with His 377, and a weaker interaction with the Asn484 side-chain. All these hydrogen bonds are also observed in the crystal structures of the GP complexes with the hydantoin ligands and glucose.

Interactions of the spirohydantoin moieties

The hydantoin rings of the various ligands have additional interactions with amino acids and water molecules. These interactions may be partly responsible for the enhanced GP binding and inhibition strength of the hydantoins, as compared with glucose; we discuss this further below. The rings are observed in very similar positions and orientations, both in the crystallographic structures and in the simulations of the different complexes. However, due to the chemical differences between the ligands, the rings form somewhat different interactions with the surrounding protein and water, as discussed below.

A typical structure of the hydantoin moiety and the surrounding protein residues and water molecules in the GP:hydan complex is shown in Figure 3(B). Typical structures of the hydantoin moieties in the GP:methyl- and GP:n-hydan complexes are shown in Figure 4.

As in the crystal structures,[5–7] all ligands form an interaction via their N2 atom with the main-chain carbonyl oxygen of His377. This interaction does not exist in the complex between GP and glucose, and may contribute to the enhanced potency of hydan compared to glucose. In the simulations of all complexes, the N2–His377 interaction is stable, with an average length of ≈3 Å. An exception is the 2-ns GP:hydan simulation, where the interaction is observed in the first 700 ps; in this segment, the average distance N2–His377 O is ≈3.5 Å (i.e., the interaction is somewhat weaker). After 700 ps, and for the remainder of the simulation, the main-chain CO group of His377 rotates and interacts with the His459 NE2 atom and water molecule. Presumably, the duration of the simulation is not sufficient for the sampling of the possible CO orientations with the correct probability. As we show below, His377 provides a very small contribution to the computed free energy changes; thus, insufficient sampling of the His377 CO orientations does not affect significantly the accuracy of the free energy results.

Oxygen O7 makes a stable, water-mediated interaction with the carboxyl side-chain of Asp283, and a weak, direct or water-mediated interaction with the amide group of Gly135. These interactions are observed in the corresponding crystal structures. Oxygen O8 interacts with the backbone nitrogen of Asn284 and a water molecule in all simulations. In the 2-ns simulations, a total of 6–8 different water molecules are observed to make a hydrogen bond with O8, one at a time. The longest residence of a water molecule interacting with O8 is about 700–750 ps, depending on the complex. In this most stable arrangement, the water molecule participates in a three-water network that links atom O8 with the side-chain of Asp339 (GP:hydan), or Glu385 (GP:n-hydan).

The N1 atom of hydan forms a stable interaction with the side-chain of Asp283 in the GP:hydan simulations [Fig. 3(B)]. An analogous stable interaction is observed in the GP:n-hydan complex simulations, between the N3 atom of the n-hydan ligand and Asp283 [Fig. 4(B)]. In the GP:methyl-hydan complex, atom N1 is connected to a methyl group; the carbon atom of this group (C9) is at an average distance of 3.5 Å from Asp283 Oδ1 [Fig. 4(A)]. Asp283 also

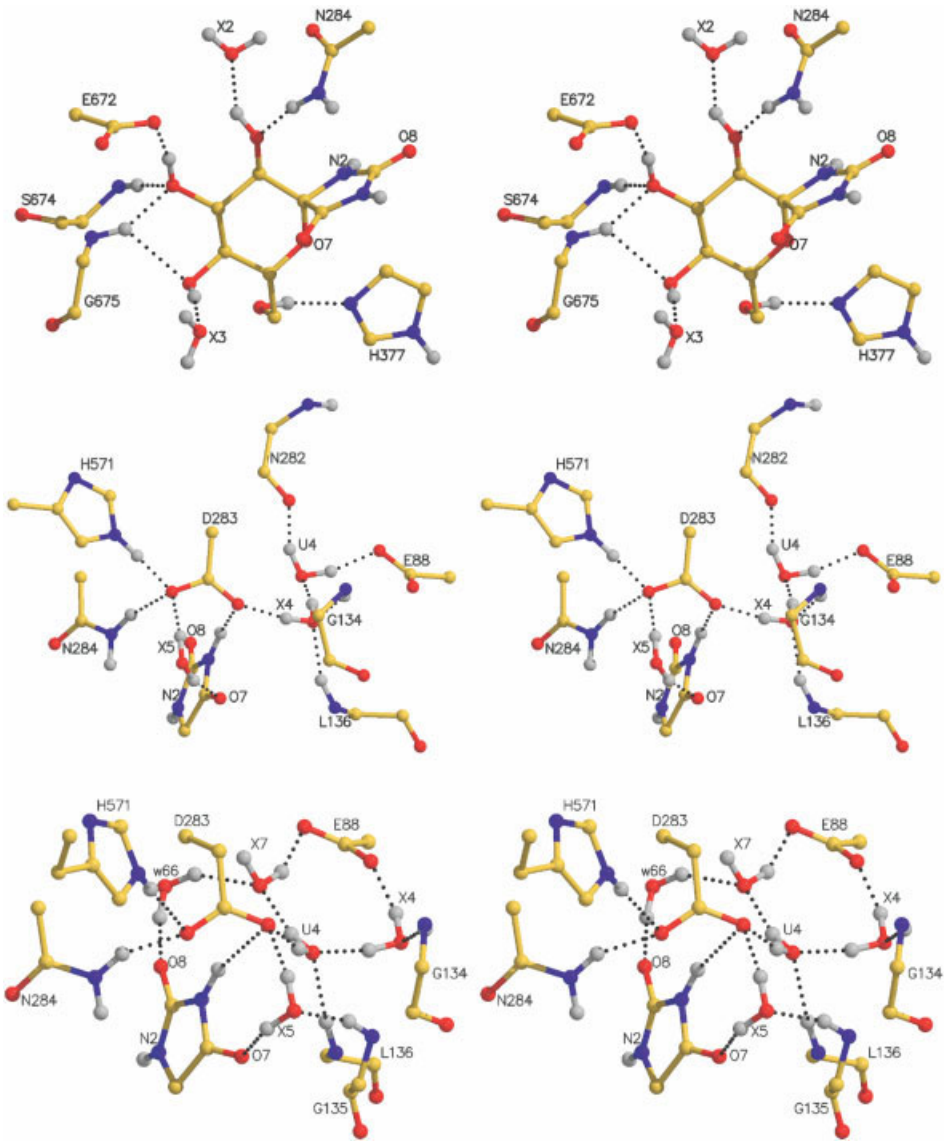

Fig. 3. Typical structures of the hydan-complex catalytic site, observed in the MD simulations. (**A**) Top panel. Protein and water configuration around the sugar moiety of hydan. (**B**) Middle panel. Protein and water configuration around the spirohydantoin moiety. In the structure shown, X4 is near D283 and interacts with D283, U4, G135, and L136. This is referred to as "position 1" in the text. The sugar moiety has been omitted for clarity. (**C**) Bottom panel. Protein and water configuration around the spirohydantoin moiety, observed in the later stages of a 2-ns GP:hydan complex simulation. Compared to the structure shown in Figure 1(B), X4 has moved by ~1.3 Å away from D283 and forms a hydrogen bond with E88. This is referred to as "position 2" in the text. The sugar moiety has been omitted for clarity, as in (B).

forms direct interactions with residues His571 and Asn284, and water-mediated interactions with hydan atom O7, Leu136, Asn284, and Glu88.

Water residue X4 is important because it has a significant effect on the free energy results, as we show below. This water does not interact strongly with the ligands, but is in the vicinity of the mutating group. In the simulations of the various complexes, it is observed in two different positions. In the first position ("position 1"), depicted in Figure 3(B), X4 acts as a hydrogen-bond donor to the side-chain carboxylate of Asp283, and to another water molecule [U4 in Fig. 3(B)]. The oxygen of X4 interacts also with the backbone amide groups of Leu136 and Gly134. This position is thermodynamically dominant in the GP:hydan complex (see below). In the second position ("position 2"), shown in Figure 4, X4 shifts away from the ligand by ≈1.3 Å, breaks its interaction with Asp283, and forms a new hydrogen bond with the carboxylate of Glu88. In the simulations of the GP:methyl- and n-hydan complexes, as well as the M and N sides of the alchemical transformations, X4 is observed in position 2.

In the shorter (several hundred picoseconds) runs of the GP:hydan complex, X4 is observed most of the time (except for a few ≈1 ps time intervals) in position 1 [Fig. 3(B)]. In

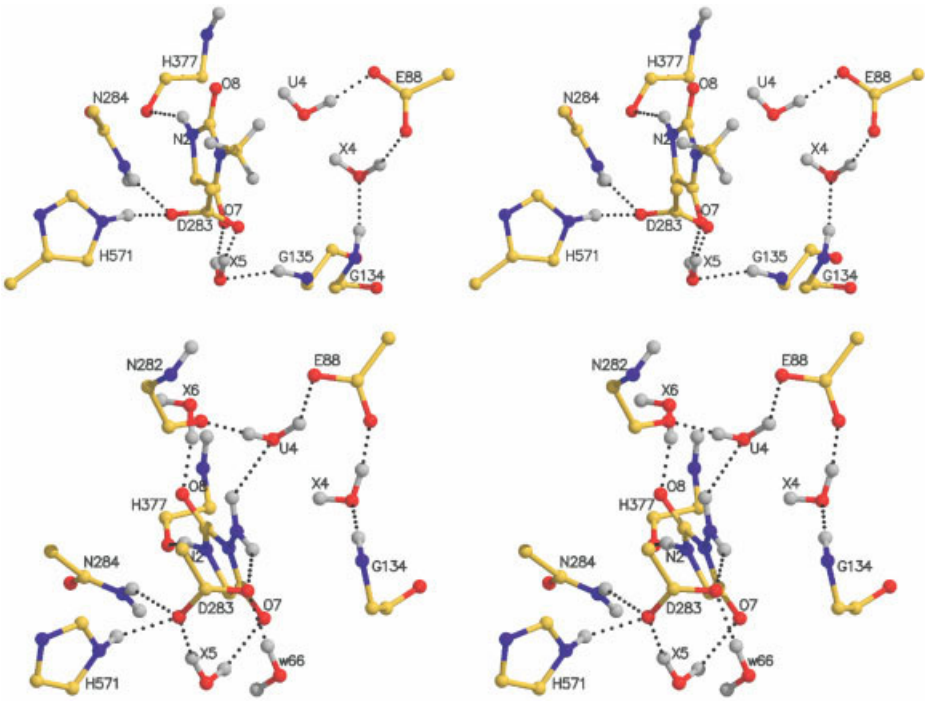

Fig. 4. Typical structures of the methyl- and n-hydan complex catalytic sites, showing the ligand spirohydantoin moiety and its interactions with surrounding protein residues and water molecules. (**A**) Top panel. GP:methyl-hydan complex; (**B**) Bottom panel. GP:n-hydan complex. The sugar moiety of the ligands has been omitted for clarity.

the 2-ns simulation, X4 is in position 1 for the initial 900 ps. At that point, it moves into position 2. Immediately thereafter, position 1 is filled by U4, and a third water molecule (X7) moves into the location typically occupied by U4. This arrangement, with both positions 1 and 2 filled by water, is shown in Figure 3(C); it is stable for the remainder of the simulation. A high-resolution crystal structure of the GP:hydan complex was determined at 100 K by Gregoriou et al.[6] In this structure, an analogous arrangement is observed experimentally [cf. Fig. 3(C) of our current work and Fig. 6(B) of Gregoriou et al.[6]).

Free Energy Simulations

We describe and analyze the results of the free energy transformations in what follows. The computed free energy changes are summarized in Table I.

H → M

To calculate the relative binding affinity of methyl-hydan with respect to hydan, we performed three transformations linking the states H and M in the protein complex, and two transformations in solution. Two of the protein free energy perturbation runs (denoted as "run 1" and "run 2" in Table I) were performed in the direction H → M, and the third ("run 3") in the opposite direction. The beginning structures for the "forward" runs 1 and 2 corresponded to end-structures of independent several hundred picosecond MD simulations of the hydan complex. The beginning structure of run 3 was taken at the end of a several hundred picosecond MD simulation of the methyl-hydan complex.

The results listed in Table I show that the mutation H → M is unfavorable, both in the protein environment and in solution. In the protein, it is associated with a free energy of 8.6 ± 1.4 kcal/mol. In solution, the corresponding free energy change is 4.85 ± 0.05 kcal/mol, yielding a double free energy difference of 3.75 ± 1.4 kcal/mol. This compares very well with the difference of 3.6 kcal/mol, estimated from the experimental K_i values of the two ligands.

To gain insight into the interactions that contribute to the calculated free energies we examine the van der Waals and electrostatic free energy components. As shown in Table I, both components destabilize methyl-hydan relative to hydan in the protein complex and in solution. The van der Waals term is dominant, providing ~80% of the total free energy value. In the protein complex, the large van der Waals value arises from strong steric interactions of the introduced methyl group with a nearby protein residue (Asp283) and a specific water molecule (X4; see below and Table II).

Water X4 merits special consideration. In all free energy transformations H → M, it is at position 1 at the H end and interacts with Asp283 [Fig. 3(B)]; at the M end, it is at position 2, interacting with E88 [Fig. 4(A)]. The λ values at which X4 switches position differ among runs, contributing to the spread of the obtained free energy values in the protein complex (±1.4 kcal/mol). The average of the three runs (8.6 kcal/mol) corresponds

TABLE I. Summary of Free Energy Runs[a]

		Run	*Vw*[c]	*Elec*[c]	Total
Protein					
H → M[b]		1	8.7	1.8	10.5
		2	6.5	1.5	8.0
	*[d]	3	5.7	1.7	7.4
Average			7.0 ± 1.3	1.7 ± 0.1	8.6 ± 1.4
		2′[e]	4.1	1.5	5.6
Solution		1	3.8	1.1	4.9
	*	2	3.7	1.1	4.8
Average			3.75 ± 0.05	1.1	4.85 ± 0.05
Difference[f]		Exp:	3.6		3.75 ± 1.4
Protein					
H → N[b]		1	4.6	−3.1	1.5
	*	2	3.2	−1.6	1.6
Average			3.9 ± 0.7	−2.35 ± 0.75	1.55 ± 0.05
H → H′[b]		1	—	6.9	6.9
	*	2	—	6.6	6.6
Average				6.75 ± 0.15	6.75 ± 0.15
H′ → N[b]		1	1.6	−5.1	−3.5
	*	2	1.6	−7.2	−5.6
		3	0.17	−6.2	−6.0
Average			1.1 ± 0.7	−6.2 ± 0.9	−5.0 ± 1.1
Sum[g]			1.1 ± 0.7	0.55 ± 1.0	1.75 ± 1.1[g]
Total Average (using boldface values)					1.65 ± 1.1
Solution					
		1	2.0	−1.5	0.5
	*	2	2.3	−1.5	0.8
Average			2.15 ± 0.15	−1.5	0.65 ± 0.15
Difference[f]		Exp:	2.30		1.0 ± 1.1

[a]All quantities in kcal/mol.
[b]The notation is explained in the Methods section (see "Thermodynamic Cycle and Mutation Protocol").
[c]Free energy components.
[d]An asterisk denotes a run performed in the opposite direction. However, all reported values correspond to the directions indicated by the arrows.
[e]In run 2′, X4 is in the position close to Glu88 throughout the transformation (see text).
[f]Double free energy difference protein–solution.
[g]Result of the composite pathway, H → H′ → N, calculated by adding the underlined values.

to the free energy of the mutation H → M, with X4 switching between positions.

In the first window (λ = 0.002) of run 2 (direction H → M), X4 is observed in both positions, with occupancies lasting for several hundred picoseconds. In all other windows of the same run, X4 occupies position 2. By combining the portion of the first window, where X4 is in position 1, with the other windows of the run, we obtain an estimate for the free energy change in the mutation H → M, with X4 switching between its two positions (as in runs 1 and 3). The result is included as run 2 in Table I. Alternatively, if we combine the portion of the first window, where X4 is at position 2, with the other windows of this run, we obtain an estimate for the free energy change, subject to the restriction that X4 is near Glu88 (at position 2) *throughout* the transformation. This result has been included in Table I as run 2′. We note that the difference between the result of run 2′ (5.6 kcal/mol) and the average of runs 1–3 (8.6 kcal/mol) is significant (3.0 kcal/mol). This difference corresponds to the free energy penalty associated with moving a water molecule from position 1 to position 2 in the GP: hydan complex. To obtain an independent estimate of the relative stability of the two X4 positions, we also evaluated the free energy profile for the translocation of X4 from position 1 to position 2 (see below).

The steric interference with the methyl group of the ligand is very significant in the windows at which the methyl group nonbonded energy terms are multiplied by a small λ value, and yields a very large contribution to the van der Waals free energy derivative (≈150 kcal/mol or 30% of the total derivative at λ = 0.002; data not shown). This large value is expected, since the total van der Waals derivative varies as $\lambda^{-0.75}$ at this end point (end point catastrophy[47,48]). The total average van der Waals derivative is ≈500 and 300 kcal/mol when X4 is respectively in positions 1 and 2, and the X4 contribu-

TABLE II. Free Energy Components (in kcal/mol) in the Protein Mutations: Each Component Refers to the Interaction of the Listed Residue With the Variable Part of the Ligand[a,b]

H → M

	Vw				*Elec*			
Residue	Run 1	Run 2(2′)	Run 3	[c]ΔE_{res}^{VW}	Run 1	Run 2	Run 3	ΔE_{res}^{elec}
Asp283	8.0	7.97(6.01)	6.57	0.21	2.59	2.5	2.47	1.75
Asn284	−0.39	−0.75	−0.44	−0.94	0.30	0.07	0.10	0.15
Gly134	−0.15	−0.21	−0.20	−0.16	−0.10	−0.03	−0.12	−0.10
Gly135	−0.42	−0.61	−0.72	−0.71	−0.05	0.10	−0.04	0.09
Leu136	−0.52	−0.19	−0.47	−0.67	−0.26	−0.11	−0.23	−0.31
His377	−0.22	−0.04	0.20	0.09	−1.0	−0.36	−0.87	−0.29
X4	4.25	3.17(0.81)	2.77	−0.13	0.14	0.05	0.10	0.06
U4	0.59	1.07	0.65	0.41	−0.13	−0.12	−0.15	−0.06
X5	−0.17	−0.01	−0.17	−0.08	−0.70	−0.51	−0.53	−0.46

H → N

	Vw			*Elec*		
Residue	Run 1	Run 2	[d]ΔE_{res}^{VW}	Run 1	Run 2	[d]ΔE_{res}^{elec}
Asp283	5.97	4.69	0.70	0.17	0.35	0.96
Asn284	−0.79	−0.79	−3.42	0.10	0.15	0.31
Gly135	−0.75	−0.61	−2.30	−0.01	0.05	0.28
Leu136	−0.58	−0.42	−1.01	−0.74	−0.68	−1.10
His377	0.17	0.01	−0.04	−0.87	−0.62	−0.86
X4	2.24	2.19	−0.15	0.82	0.40	0.22
U4	0.26	0.81	0.46	−0.74	−0.49	−1.31
X5	−0.05	0.02	0.28	−0.73	−0.77	−1.83

[a]The total free energy components are listed in Table I.
[b]The runs are explained in Table I. For the transformation H → M, selected components, corresponding to run 2′ are listed, when significantly different from the corresponding components of run 2.
[c]Averages evaluated at the methyl-hydan end state.
[d]Averages evaluated at the n-hydan end state.

tion to this derivative is ≈160 and 60 kcal/mol, respectively.

In Figure 5, we plot the total van der Waals derivative for the various runs. The diverging end points (Fig. 5, bottom) are also shown in a log-log plot. This plot shows that the power law ($\langle Vw \rangle \sim \lambda^{-0.75}$), expected at the end point for a disappearing or appearing Lennard–Jones particle,[47,48] is observed between $\lambda = 0$ and ∼0.06. The contribution to the total free energy in this range is obtained by use of the power law; in the rest (0.06–1), a trapezoeidal estimation of the free energy is used. At the other end point (where the hydan proton is eliminated), we do not observe an end point catastrophy. This happens because the eliminated hydan group is shielded by the (almost) fully grown methyl-group of methyl-hydan.

Group decomposition: To gain more insight into the interactions that contribute to the computed free energies, we list in the upper panel of Table II the most important protein residue and water contributions to the van der Waals and electrostatic free energy components of the transformation H → M.

The two largest free energy components are due to Asp283 and water X4. These residues interact sterically with the methyl group that is created in the alchemical mutation H → M. Other important contributions arise from residues Asn284, Gly134, Gly135, and Leu136. These residues are located in the vicinity of the introduced group, but they are sufficiently far to avoid significant steric repulsions. They have better van der Waals interactions with the ligand at the methyl-hydan end state, and give favorable contributions to the free energy. X5 is associated with an electrostatic component that favors the M state. In the H end, this water mediates an interaction between the ligand atom O7 and Asp283. When X4 moves away from Asp283 (in the course of the H → M transformation), X5 approaches Asp283 and improves at the same time its interaction with O7.

It is interesting to compare the residue free energy components with the corresponding energy terms [see Eq. (3)]. The contribution of a particular residue to the van der Waals energy change in the process H → M is approximately equal to the average van der Waals energy of interaction between this residue and the created methyl group. This energy is listed in column $\Delta E_{\mathrm{res}}^{VW}$ of Table II. As mentioned before [see Eq. (3) and the following discussion], each residue contributes also indirectly to the change in the average energy of the unperturbed part of the system.

The contributions to the van der Waals energy change from residues Asp283 (0.21 kcal/mol) and X4 (−0.13 kcal/mol) are significantly smaller than the corresponding van der Waals *free energy* components. This is because the energetic terms are calculated at the final (M) state, when the system has relaxed in the presence of the additional methyl group and has optimized its interactions with it. In

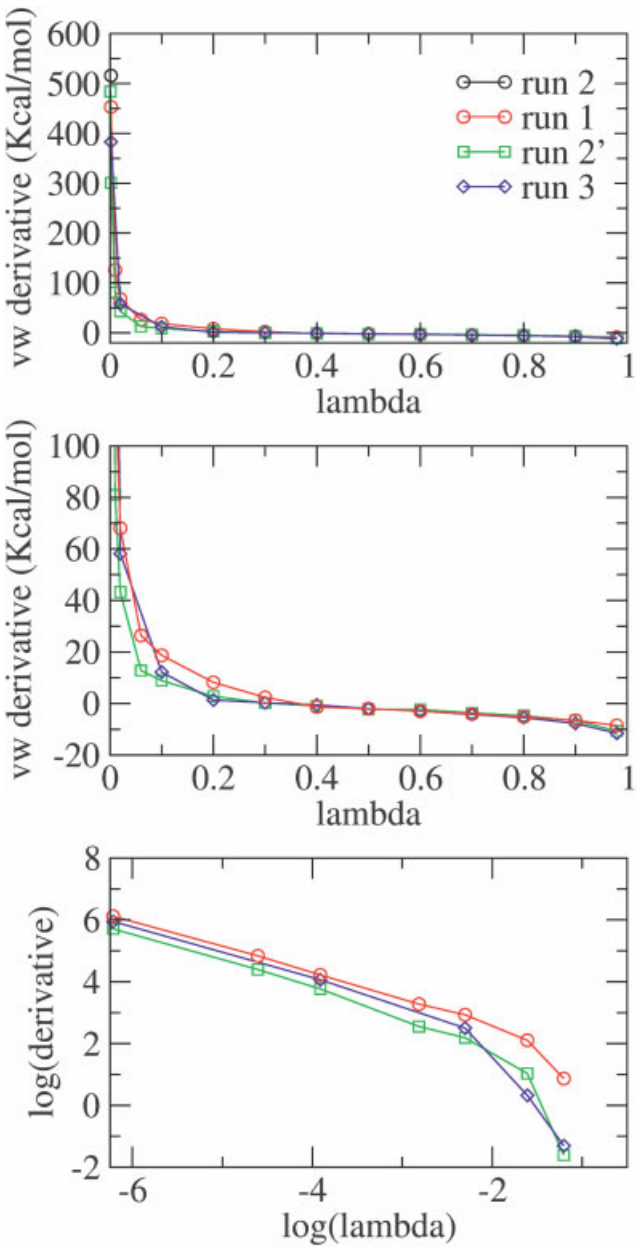

Fig. 5. Plots of the van der Waals free energy derivative for the transformation H → M as a function of the parameter λ. The methyl group appears at λ → 0. **Top:** Total derivative. **Middle:** Same as the upper plot, but in the region λ = 0.06–1.0. **Bottom:** Log-log plot of the total van der Waals derivative at the end λ → 0. The power law is observed up to λ ∼ 0.06 (logλ ≈ −2.81). Run 2 corresponds to a single point at λ=0.002, and is included only in the top plot. For an explanation of the various runs, see text and Table I.

contrast, the free energy components are integrals of energy terms over the entire transformation.

Other residues (e.g., Asn284, Gly135, and Leu136) have more favorable van der Waals interactions with the additional methyl group, in better agreement with the corresponding free energy components. Presumably this agreement is due to the fact that relaxation is less important for these residues.

It is worth noting that an assessment of the GP:methylhydan recognition from the interactions between the methyl group and the surrounding residues would lead one to conclude that removal of the methyl group is not favored. The interactions formed by a specific ligand or side-chain group at a particular end state structure are often used to describe the probable effects of a mutation in theoretical studies of biomolecular complexes.[16,31–33] Since the free energy components incorporate information on the behavior of the system *throughout* the transformation and the total free energy is an integral over the transformation, they are more instructive. This point is analyzed further in the Concluding Discussion section. Alternatively, a linear-response approximation to the free energy components would involve averaging the corresponding energy components over the two end states.[31] In our case, evaluation of the methyl group interactions at the GP:hydan end state will give very large values.

H → N

The mutation H → N was studied by two different protocols in the protein complex, in order to obtain independent estimates of the free energy change. Since the free energy is a state function and the protocols employ the same end states, the results should be identical. Indeed, as we show below, the obtained values are in close agreement, respectively 1.55 kcal/mol and 1.65 kcal/mol in favor of the H state.

The first protocol employed a "dual-topology" ligand for the entire transformation; with this protocol, a part of the ligand containing the changing groups was duplicated (see "Thermodynamic Cycle and Mutation Protocol" in the Methods section). The second protocol employed an intermediate state (H′ in Table I), with the topology of hydan and zero charge on the atoms N1, HN1, C8, and O8. We first discuss the dual-topology protocol.

With the dual-topology H → N protocol we performed two runs in the protein and two in solution (one in the forward and one in the backward direction). In the protein complex, the N-state is more unstable than H by 1.55 ± 0.05 kcal/mol. This value is significantly smaller than the change (8.6 kcal/mol) in the H → M transformation, due to two factors. First, the van der Waals free energy component contributes to the free energy only 3.9 kcal/mol, compared to 7.0 kcal/mol for the H → M transformation. Furthermore, the electrostatic component disfavors the H state by 2.35 kcal/mol, in contrast to the H → M transformation, where the corresponding component favors the H end point by 1.7 kcal/mol. The difference between the values of the two protein transformations is partially cancelled by a similar difference in the solution transformations (see below).

By use of the intermediate state H′, the mutation H → N can be decomposed into two steps. Step H → H′ is performed by elimination (or creation) of the partial charges on atoms N1, HN1, C8, and O8 of the hydan ligand (see Fig. 1); step H′ → N employs a dual-topology ligand carrying an H′ and an N moiety, and corresponds to switching off(on) the nonbonded interactions on the H′(N)

moiety. The total mutation H → N is a sum of these two steps.

The uncharging mutation H → H′ was performed in the forward and backward direction. The elimination of charges on the hydan moiety destroys ligand interactions with surrounding groups, yielding an unfavorable free energy of 6.75 ± 0.15 kcal/mol. This is projected entirely on the electrostatic component.

The H′ → N mutation was also performed in both directions (two runs in the forward and one in the reverse). The van der Waals free energy component (1.1 kcal/mol) is smaller than in the H → N transformation (3.9 kcal/mol). This is due to the fact that the H′ portion of the dual-topology ligand forms weaker electrostatic interactions with Asp283, and the steric interactions between the N portion of the ligand and the residues Asp283 and X4 are also reduced. A similar behavior of the van der Waals free energy component was observed in the free energy transformations $Cl^- \rightarrow Br^-$, and $Cl^0 \rightarrow Br^0$ in water.[47] The electrostatic contribution on the other hand is large (−6.2 kcal/mol) and favors the N state; this is expected since the H′ state has an uncharged group of atoms; these atoms acquire charge and form hydrogen bonds with protein and water residues at the N end state.

When the two steps H → H′ and H′ → N are added, the resulting total free energy change disfavors the GP:n-hydan complex by 1.75 ± 1.1 kcal/mol. Averaging over the results of the pathways with and without the H′ intermediate, we obtain a total free energy change of 1.65 ± 1.1 kcal/mol, disfavoring the N state.

In solution, the van der Waals and electrostatic free energy components have similar signs as in the protein complex, yielding a total free energy that is mildly destabilizing for n-hydan (by 0.65 ± 0.15 kcal/mol). Using the protein and solution results, we predict that the hydan complex is more stable by 1.0 ± 1.1 kcal/mol, in fair agreement with the estimate of 2.3 kcal/mol from the experimentally measured K_i values.[7]

Group decomposition: A residue free energy decomposition of the H → N transformation is listed in the lower panel of Table II, along with the average interaction energy terms $\langle \Delta H \rangle_1$, evaluated at the n-hydan end of the transformation (state 1). We discuss first the van der Waals components.

As for H → M, the largest positive van der Waals free energy components correspond to Asp283 and X4. The computed values are significantly smaller than in the H → M case. This indicates that the NH_2 group of the n-hydan ligand has reduced steric interactions compared to the methyl group of m-hydan, and can be easier accomodated in the catalytic site.

The contributions of various residues to the total van der Waals energy change are also listed in column 4 (lower panel) of Table II; they are evaluated at the N state. As before, most energetic components differ significantly from the free energy components of the same residues. The Asp283 and X4 values are much smaller than the corresponding free energy values, reflecting the fact that at the N state the system has relaxed and has optimized its interactions with the additional NH_2 group. Residues Asn284, Gly135, and Leu136 contribute large, negative values to the van der Waals energy, whereas the corresponding free energy components are smaller. The energy terms are increased (in absolute value) with respect to the corresponding values at the M state (Table II, upper panel), indicating that the substitution of a CH_3 by an NH_2 group improves the van der Waals interactions with these residues.

The electrostatic free energy components are also included in Table II, lower panel. The Asp283 electrostatic free energy component is small; more distant residues such as His377, Leu136, X5, and U4 have more negative electrostatic interactions with the n-hydan ligand, yielding a total negative electrostatic free energy component. In most cases, these terms are more negative with respect to the M state. Thus, improved electrostatic interactions contribute to the higher relative affinity of GP for state N compared to M.

Relative stability of X4 positions in the hydan and methyl-hydan complex

As described above, X4 is observed in two different positions in simulations. In the hydan complex and the H side of the alchemical transformations, position 1 is preferred [Fig. 3(A)]; in the methyl- and n-hydan complexes, and the M or N sides of the alchemical transformations, X4 stays almost exclusively at position 2 (Fig. 4).

In the crystallographic structures, water molecules are also observed in either position, depending on the experimental temperature and the complex. In a GP:hydan complex crystal structure determined at room temperature,[7] there is a single water at position 1. This position is also preferred in the MD simulations of the hydan complex. In a hydan-complex crystal structure determined at 100 K, both positions 1 and 2 are occupied [see Fig. 6(B) of Gregoriou et al.,[6] where the waters in the two positions are as X4 and X7, respectively]. These two waters, along with Asp283, Gly134, and Glu88, form a hydrogen bond network that is very similar to the network observed in the last 1-ns segment of our 2-ns hydan-complex MD simulation. In the n- and methyl-hydan crystal structures, there is also only one visible water at position 1 [X4 in Fig. 1(a, b, and d) in Watson et al.[7]). On the other hand, in a complex with a spirohydantoin derivative that contains an OH group attached to N1 (compound 3 in Watson et al.[7]), position "1" is empty and position "2" is occupied [X7 in Fig. 1(c) of Watson et al.[7]). The same is observed in the crystal structure of the native, unliganded Gpb protein.

In the H → M transformations, we obtained estimates of the free energy with X4 either moving from one position to the other (8.6 ± 1.4 kcal/mol), or staying at position 2 throughout the run (5.6 kcal/mol; run 2′ in Table I). This suggests that position 1 is more stable by ≈3.0 kcal/mol. The uncertainty in this value is at least as large as 1.4 kcal/mol, since the estimate of 5.6 kcal/mol (transformation with X4 is position 2) was obtained from only one run (2′).

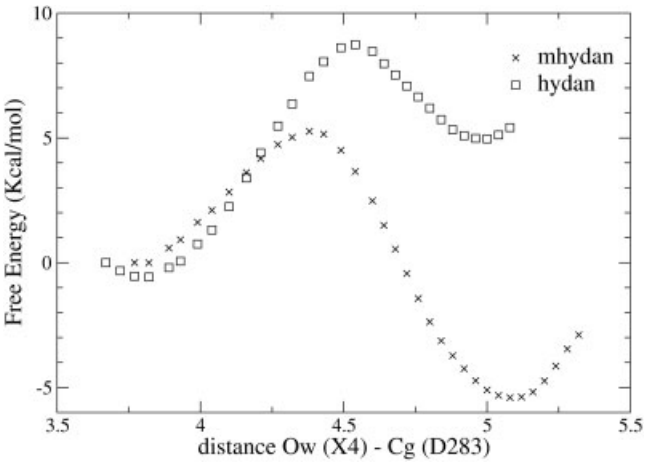

Fig. 6. Free energy profiles for the translocation of X4 from position 1 (near Asp283) to position 2 (near Glu88) in the GP:hydan and GP:methyl-hydan complex (see text). The profiles have been shifted to 0.0 at the point close to D283.

To obtain an independent estimate of the relative stability of the two positions, we calculated the free energy profiles for the translocation of X4 from position 1 to position 2 in the hydan- and methyl-hydan complex. The translocation of X4 was achieved by umbrella sampling calculations along 30 windows, in which the distances between the oxygen atom O_w of X4 and the atoms C_γ (Asp283) and C_δ (Glu88) were varied in a concerted manner (see the Methods section). The resulting profiles are shown in Figure 6.

In the hydan complex, position 1 is calculated to be more stable by ≈5.5 kcal/mol. This is somewhat larger than the value (3.0 kcal/mol) estimated above. Conversely, in the methyl-hydan complex, position 2 is more stable by 5.5 kcal/mol. Thus, in the hydan(methyl-hydan) complex, X4 is expected to spend most of the time in position 1(2).

To interpret the differences between the two profiles, we calculated the average total interaction energy of X4 with its surrounding residues Asp283, Glu88, U4, Gly134, Gly135, Leu136, and the ligand, using the trajectories of the free energy profile simulations. The energetic profile, and its decomposition into van der Waals and electrostatic contributions is plotted as a function of the distance O_w(X4) − C_γ (Asp283) in Figure 7. In both complexes, the total energy profile has a shape similar to the corresponding free energy profile, with a barrier at intermediate distances and low values at the distances of positions 1 and 2. The shape of the energetic profile is determined by the electrostatic interaction energy. The interactions of X4 with Asp283 (position 1) and Glu88 (position 2) stabilize the profile at the extreme positions, whereas the loss of electrostatic interactions at intermediate distances creates the large energetic barrier. In the hydan complex, the energetic profile has a lower minimum at position 1 [small values of O_w(X4) − C_γ (Asp283)]. Inspection of the contributions from individual residues shows that position 1 is more stable energetically than position 2 due to better interactions with residue Leu136 and water U4 (data not shown). In the methyl-hydan complex, position 2 is more stable energetically due to the electrostatic energy term. The main contribution to this stabilization is due to interactions with residues Glu88 and Gly134. Residues Leu136 and U4 favor position 1, but to a smaller degree than in the hydan complex.

The above analysis shows that the stabilization of position 2 in the simulations of the methyl-hydan complex is due to electrostatic interactions, particularly with Glu88. Inspection of the structures from the restrained-distance simulations (X4 restrained at position 2) shows that in the methyl-hydan complex the average distance between the X4 hydrogens and the two side-chain carboxylate oxygens of Glu88 is smaller. At the same time, the distance between atoms Asp283 C_γ and Glu88 C_δ is somewhat larger (by ≈0.25 Å) in the methyl-hydan complex; this could contribute to the improved interactions between X4 and Glu88 in the latter complex.

CONCLUDING DISCUSSION

In this work we have employed MD free energy simulations to calculate the relative binding free energies of the complexes between GP and three spirohydantoin ligands: hydan, methyl-hydan, and n-hydan. Methyl-hydan had been synthesized as a potentially better inhibitor than hydan, based on the expectation that the addition of a nonpolar group would increase the binding strength; n-hydan had been chosen because it was thought that improved binding would result from possible additional hydrogen bonds between the NH_2 group and the protein. However, both methyl-hydan and n-hydan were found to bind more weakly to GP than hydan itself; the calculated differences are, respectively, 3.75 ± 1.4 and 1.0 ± 1.1 kcal/mol, as compared with the measured values of 3.65 and 2.30 kcal/mol.

The free energy decomposition analysis from the simulations showed that the introduction of the methyl group in the GP catalytic site is much more unfavorable (by 8.6 ± 1.4 kcal/mol) than the same introduction in the unbound (solvated) hydan. A specific protein residue (Asp283) and a water molecule (X4) make the most significant contribution to the unfavorable H → M free energy. Both residues interact sterically with the methyl group at the initial stages of its introduction in the simulation, yielding large destabilizing van der Waals free energy components. As in the methyl-hydan case, the introduction of the additional NH_2 group in the complex was associated with a larger free energy penalty than the same introduction in the solvated ligand. The free energy decomposition showed that the van der Waals free energy component was also positive (but smaller than in the H → M case) with the main contributions again arising from Asp283 and X4. As in the H → M case, this implies that the introduction of the NH_2 group in the protein interior is disfavored due to lack of space, although to a smaller extent than in the case of the methyl group. The electrostatic free energy component of the transformation H → N was negative, indicating that n-hydan does form improved electrostatic interactions with the GP catalytic-site, as compared with hydan. How-

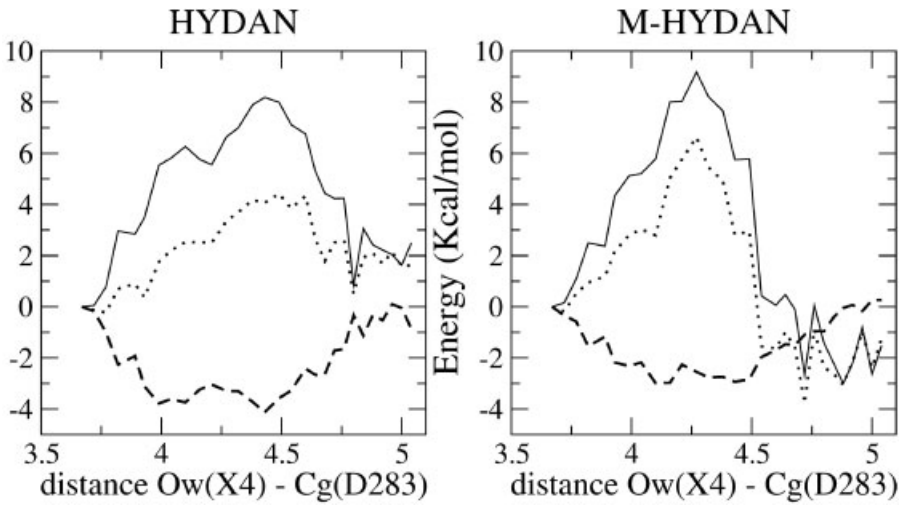

Fig. 7. Change in the average interaction energy between X4 and surrounding residues, in the translocation of X4 from position 1 (near Asp283) to position 2 (near Glu88) in the GP:hydan and GP:methyl-hydan complex (see Fig. 6 and text). The energy curves have been shifted to 0.0 at the point close to D283: electrostatic energy (solid line); van der Waals energy (dashed line); and total energy (dotted line).

ever, this stabilizing contribution is not sufficient to render n-hydan a better inhibitor than hydan.

A comparison of the interactions between a set of ligands and their surrounding groups observed in X-ray structures is often used in the interpretation of binding free energy differences, and in suggesting chemical modifications of a ligand to produce better binding. Often such an approach can yield correct predictions (e.g., when the system responds linearly to a new interaction[16,31–33]). However, consideration of the ligand–environment interactions alone may not be sufficient for an accurate estimation of relative binding strength, or even of the order of the binding strengths. Also, it is important to note that the change in solubility (i.e., the alchemical transformation of the ligand in solution, which is often neglected in structural analyses) can play an important role in the observed binding.[23] This complexity is present in the studies described here, as indicated by the results of the H → M and H → N free energy analyses. In the transformation H → M, the methyl group of M forms favorable van der Waals interactions with Asn284, Gly135, and Leu136 (see column 5 in Table II, upper panel); these interactions are to a large extent canceled by the unfavorable electrostatic interaction with Asp283 (column 9). Based on the energy of the interactions alone, it could be concluded that ligands H and M are almost equally stable in the binding site, in contrast to the free energy prediction of a much more favorable H state. In the transformation H → N, favorable interactions between the N-state NH_2 group and the surrounding protein residues (particularly, van der Waals interactions with Asn284 and Gly135; see Table II lower panel, columns 4 and 7) would support the conclusion that the N state is more stable than H, also in contrast to the free energy result.

For the series of ligands studied here, the insertion of a group at position N1 is not favored, primarily due to lack of space in the protein. Even though biomolecular systems have the ability to accomodate mutations by structural reorganizations, rendering generalizations uncertain,[22] it is likely to prove difficult to obtain improved inhibitors by chemical modifications at N1. Other substitutions on the hydantoin ring are being exploited with promising results.[4]

ACKNOWLEDGMENTS

Figures 4 and 5 were prepared with MOLSCRIPT[53] and Raster3D.[54] We wish to express our gratitude to Prof. Dame Louise Johnson for her continuing interest and for useful discussions concerning this study. We would like to thank the referees, for helpful comments. All simulations were performed on a Linux cluster at the University of Cyprus.

REFERENCES

1. Oikonomakos NG. Glycogen phosphorylase as a molecular target for type 2 diabetes therapy. Curr Protein Pept Sci 2002;3:561–586.
2. Martin, JL, Veluraja K, Ross K, Johnson LN, Fleet GWJ, Ramsden NG, Bruce I, Orchard M, Oikonomakos NG, Papageorgiou AC, Leonidas DD, Tsitura HS. Glucose analogue inhibitors of glycogen phosphorylase: the design of potential drugs for diabetes. Biochemistry 1991;30:10101–10116.
3. Watson KA, Mitchell EP, Johnson LN, Son JC, Bichard CJF, Orchard M, Fleet GWJ, Oikonomakos NG, Leonidas DD, Kontou M, Papageorgiou AC. Design of inhibitors of glycogen phosphorylase: a study of α- and β-C-glucosides and 1-thio-beta-D-glucose compounds. Biochemistry 1994;33:5745–5748.
4. Oikonomakos NG, Skamnaki VT, Osz E, Szilagyi L, Somsak L, Docsa T, Toth B, Gergely P. Kinetic and crystallographic studies of glucopyranosylidene spirothiohydantoin binding to glycogen phosphorylase b. Bioorg Med Chem 2002;10:261–268.
5. Bichard CJF, Mitchell EP, Wormald MR, Watson KA, Johnson LN, Zographos SE, Koutra D, Oikonomakos NG, Fleet GWJ. Potent inhibition of glycogen phosphorylase by a spirohydantoin of glucopyranose: first pyranose analogues of hydantocidin. Tetrahedron Lett 1995;36:2145–2148.
6. Gregoriou M, Noble MEM, Watson KA, Garman EF, Krulle EM, de la Fuente C, Fleet GWJ, Oikonomakos, NG, Johnson LN. The structure of a glycogen phosphorylase glucopyranose spirohydantoin at 1.8 Å resolution and 100 K: the role of water structure and its contribution to binding. Protein Sci 1998;7:915–927.
7. Watson KA, Chrysina ED, Tsitsanou KE, Zographos SE, Archontis G, Fleet GWJ, Oikonomakos NG. Kinetic and crystallographic studies of glucopyranose spirohydantoin and glucopyranosylamine analogues inhibitors of glycogen phosphorylase. Proteins. Forthcoming.

8. Oikonomakos NG, Kosmopoulou M, Zographos SE, Leonidas DD, Chrysina ED, Somsák L, Nagy V, Praly JP, Docsa T, Tóth A, Gergely P. Binding of *N*-acetyl-*N*′-β-D-glucopyranosyl urea and *N*-benzoyl-*N*′-β-D-glucopyranosyl urea to glycogen phosphorylase b: kinetic and crystallographic studies. Eur J Biochem 2002;269:1684–1696.
9. Chrysina ED, Oikonomakos NG, Zographos SE, Kosmopoulou MN, Bischler N, Leonidas DD, Kovács L, Docsa T, Gergely P, Somsák L. Crystallographic studies on α- and β-D-glucopyranosyl formamide analogues, inhibitors of glycogen phosphorylase. Biocatal Biotransfor 2003;21:233–242.
10. Chrysina ED, Kosmopoulou MN, Kardakaris R, Bischler N, Leonidas DD, Kannan T, Loganathan D, Oikonomakos NG. Binding of β-D-glucopyranosyl bismethoxyphosphoramidate to glycogen phosphorylase b: kinetic and crystallographic studies. Bioorg Med Chem 2005;13:765–772.
11. Chrysina ED, Kosmopoulou MN, Tiraidis C, Kardakaris R, Bischler N, Leonidas DD, Hadady Z, Somsak L, Dicsa T, Gergely P, Oikonomakos NG. Kinetic and crystallographic studies on 2-(β-D-glucopyranosyl)-5-methyl-1,3,4-oxadiazole, -benzothiazole, and -benzimidazole, inhibitors of muscle glycogen phosphorylase b: evidence for a new binding site. Protein Sci 2005;14:873–888.
12. Zographos SE, Oikonomakos NG, Tsitsanou KE, Leonidas DD, Chrysina ED, Skamnaki VT, Bischoff H, Goldmann S, Watson KA, Johnson LN. The structure of glycogen phosphorylase b with an alkyldihydropyridine-dicarboxylic acid compound, a novel and potent inhibitor. Structure 1997;5:1413–1425.
13. Oikonomakos NG, Tsitsanou KE, Zographos SE, Skamnaki VT, Goldmann S, Bischoff H. Allosteric inhibition of glycogen phosphorylase a by the potential antidiabetic drug 3-isopropyl 4-(2-chlorophenyl)-1,4-dihydro-1-ethyl-2-methyl-pyridine-3,5,6-tricarboxylate. Protein Sci 1999;8:1930–1945.
14. Oikonomakos NG, Skamnaki VT, Tsitsanou KE,.Gavalas NG, Johnson LN. A new allosteric site in glycogen phosphorylase b as a target for drug interactions. Structure 2000;8:575–584.
15. Oikonomakos NG, Schnier JB, Zographos SE, Skamnaki VT, Tsitsanou KE, Johnson LN. Flavopiridol inhibits glycogen phosphorylase by binding at the inhibitor site. J Biol Chem 2000;275:34566–34573.
16. Jorgensen WL. The many roles of computation in drug discovery. Science 2004;303:1813–1818.
17. Lazaridis T. Binding affinity and specificity from computational studies. Curr Org Chem 2002;6:1319–1332.
18. Simonson T, Archontis G, and Karplus M. Free-energy simulations come of age: protein–ligand recognition. Acc Chem Res 2002;35:430–437.
19. Boresch S, Tettinger F, Leitgeb M, Karplus M. Absolute binding free energies: a quantitative approach for their calculation. J Phys Chem B 2003;107:9535–9551.
20. Gilson MK, Given JA, Bush BL, McCammon JA. The statistical–thermodynamic basis for computation of binding affinities: a critical review. Biophys J 1997;72:1047–1069.
21. Straatsma TP, McCammon JA. Theoretical calculations of relative affinities of binding. Methods Enzymol 1991;202:497–511.
22. Archontis G, Simonson T, Karplus M. Binding free energies and free energy components from molecular dynamics and Poisson–Boltzmann calculations: application to amino acid recognition by aspartyl-tRNA synthetase. J Mol Biol 2001;306:307–327.
23. Michielin O, Karplus M. Binding free energy differences in a TCR-peptide-MHC complex induced by a peptide mutation: a simulation analysis. J Mol Biol 2002;324:547–569.
24. Price MP, Jorgensen WL. Analysis of binding affinities for Celecoxib analogues with COX-1 and COX-2 from combined docking and Monte Carlo simulations and insight into the COX-2/COX-1 selectivity. J Am Chem Soc 2000;122:9455–9466.
25. Tidor B, Karplus M. Simulation analysis of the stability mutant R96H of T4 lysozyme. Biochemistry 1991;30:3217–3228.
26. Gao J, Kuczera K, Tidor B, Karplus M. Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. Science 1989;244:1069–1072.
27. Archontis G, Karplus M. Cumulant expansion of the free energy: application to free energy derivatives and component analysis. J Chem Phys 1996;105:11246–11260.
28. Carlsson J, Simonson T, Case DA. Proton binding to proteins: pK_a calculations with explicit and implicit solvent models. J Am Chem Soc 2004;126:4167–4180.
29. Srinivasan J, Cheatham III TE, Cieplak P, Kollman PA. Continuum solvent studies of the stability of DNA, RNA and phosphoramidate-DNA helices. J Am Chem Soc 1998;120:9401–9409.
30. Swanson JMJ, Henchman RH, McCammon JA. Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. Biophys J 2003;86:67–74.
31. Aqvist J, Medina C, Samuelson JE. A new method to predict binding affinity in computer-aided drug design. Protein Eng 1994;7:385–391.
32. Florian J, Goodman MF, Warshel A. Theoretical investigation of the binding free energies and key substrate-recognition components of the replication fidelity of human DNA polymerase. J Phys Chem B 2002;106:5739–5753.
33. Wesolowski SS, Jorgensen WL. Estimation of binding affinities for celecoxib analogues with COX-2 via Monte Carlo-extended linear response. Bioorg Med Chem Lett 2002;12:267–270.
34. Brooks CL III, Karplus M. Deformable stochastic boundaries in molecular dynamics. J Chem Phys 1983;79:6312–6325.
35. Brünger A, Karplus M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction analysis. Proteins 1988;4:148–156.
36. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4:187–217.
37. Mackerell AD, Bashford D, Bellott M, Dunbrack R, Evanseck J, Field M, Fischer S, Gao J, Guo H, Ha S, Joseph D, Kuchnir L, Kuczera K, Lau F, Mattos C, Michnick S, Ngo T, Nguyen D, Prodhom B, Reiher W, Roux B, Smith J, Stote R, Straub J, Watanabe M, Wiorkewicz-Kuczera J, Yin D, Karplus M. An all-atom empirical potential for molecular modelling and dynamics study of proteins. J Phys Chem B 1998;102:3586–3616.
38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
39. Halgren TA. Merck Molecular Force Field: I. Basis, form, scope, parameterization, and performance of MFF94. J Comp Chem 1996;17:490–519.
40. Stote R, States D, Karplus M. On the treatment of electrostatic interactions in biomolecular simulation. J Chim Phys 1991:88;2419–2433.
41. Brooks CL III, Brünger AT, Karplus M. Active site dynamics in protein molecules: a stochastic boundary molecular-dynamics approach. Biopolymers 1985;24:843–865.
42. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–341.
43. Boresch S, Karplus M. The meaning of component analysis: decomposition of the free energy in terms of specific interactions. J Mol Biol 1995;254:801–807.
44. Brooks CL III. The influence of long-range force truncation on the thermodynamics of aqueous ionic solutions. J Chem Phys 1987;86:5156–5162.
45. Boresch S, Karplus M. The role of bonded terms in free energy simulations: 1. Theoretical analysis. J Phys Chem A 1999;103:103–118.
46. Boresch S, Karplus M. The role of bonded terms in free energy simulations: 2. Calculation of their influence on free-energy differences of solvation. J Phys Chem. A 1999;103:119–136.
47. Simonson T. Free energy of particle insertion: an exact analysis of the origin singularity for simple liquids. Mol Phys 1993;80:441–447.
48. Resat H, Mezei M. Studies on free energy calculations: II. A theoretical approach to molecular solvation. J Chem Phys 1994;101:6126–6140.
49. Boresch S, Archontis G, Karplus M. Free energy simulations: the meaning of the individual contributions from a component analysis. Proteins 1994;20:25–33.
50. Ferrenberg AM, Swendsen RH. Optimized Monte Carlo data analysis. Phys Rev Lett 1989;63:1195–1198.
51. Souaille M, Roux B. Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. Comput Phys Commun 2001;135:40–57.
52. Martin JL, Johnson LN, Withers SG. Comparison of the binding of glucose and glucose-1-phosphate derivatives to T-state glycogen phosphorylase b. Biochemistry 1990;29:10745–10757.
53. Kraulis P. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 1991;24:946–950.
54. Merrit EA, Bacon DJ. Raster3D: photorealistic molecular graphics. Methods Enzymol 1997;277:505–524.