Most antiviral therapies exploit the large substrate acceptance of herpes simplex virus type 1 thymidine kinase (TKHSV1) relative to the human isoenzyme. The enzyme selectively phosphorylates nucleoside analogs that can either inhibit viral DNA polymerase or cause toxic effects when incorporated into viral DNA. To relate structural properties of TKHSV1 ligands to their chemical reactivity we have carried out ab initio quantum chemistry calculations within the density functional theory framework in combination with biochemical studies. Calculations have focused on a set of ligands carrying a representative set of the large spectrum of sugar-mimicking moieties and for which structural information of the TKHSV1-ligand complex is available. The $k_{cat}$ values of these ligands have been measured under the same experimental conditions using an UV spectrophotometric assay. The calculations point to the crucial role of electric dipole moment of ligands and its interaction with the negatively charged residue Glu225. A striking correlation is found between the energetics associated with this interaction and the $k_{cat}$ values measured under homogeneous conditions. This finding uncovers a fundamental aspect of the mechanism governing substrate diversity and catalytic turnover and thus represents a significant step toward the rational design of novel and powerful prodrugs for antiviral and TKHSV1-linked suicide gene therapies.

The thymidine kinase from herpes simplex virus type 1 (TKHSV1) salvages thymine into the metabolism of the virus by converting it to thymidine monophosphate (1). In contrast to the human isoenzyme, TKHSV1 acts as phosphorylating agent toward a large variety of nucleoside analogs such as (North)-methanocarba-thymidine (n-MCT), aciclovir (ACV), ganciclovir (GCV), and penciclovir (PCV). These analogs exhibit chemical diversity for the nucleobase as well as for the sugar-like chain moiety (2, 3) (Fig. 1). Substrate diversity of TKHSV1 provides the molecular basis for effective and selective treatment of viral infections. It is also exploited for gene therapy approaches involving TKHSV1 as suicide gene by anticancer intervention (4) or the control of graft versus host disease by allogeneic bone marrow transplantation (5).

The recent determination of the x-ray structure of a large spectrum of ligand-enzyme complexes (6–11) has opened a new avenue for the understanding structure-function relationships as well as of the functional role of amino acids involved in binding. It has emerged that the enzyme accommodates the nucleobase and sugar-like chain moieties in two different pockets (P1 and P2 in Fig. 2), interacting with the ligand by a specific and extensive H-bond network.

In the pocket P1, the nucleobase moiety (either a thymine or a guanine ring) is stabilized by direct H bonds with the highly conserved Gln125 and with Arg176 by means of two ordered water molecules. The pyrimidine ring is further fixed by a peculiar interaction with Tyr172 and Met128 (11) in a sandwich-like orientation (Refs. 12 and 13 and Fig. 2).

In P2, the sugar-like chains interact with the protein via its hydroxyl groups. The 3′-OH (and its mimics) forms specific H bonds with Tyr101 and/or Glu225 (Figs. 2 and 3), whereas 5′-OH forms a direct H bond or water-mediated interactions with Arg163, Glu81, and Arg222 (Figs. 2 and 3). In contrast, the polar C-1′–O′–C-4′ function belonging to dT and the corresponding ether groups of ACV and GCV interact neither with polar charged groups of the protein nor with the solvent; instead, they point toward a hydrophobic region made up of the Trp88, Arg163, Glu83, and Arg222 (Fig. 2). This is precisely the reason why TKHSV1 is able to phosphorylate a plethora of analogs that can either inhibit viral DNA polymerase or cause toxic effects when incorporated into viral DNA. To relate structural properties of TKHSV1 ligands to their chemical reactivity we have carried out ab initio quantum chemistry calculations within the density functional theory framework in combination with biochemical studies. Calculations have focused on a set of ligands carrying a representative set of the large spectrum of sugar-mimicking moieties and for which structural information of the TKHSV1-ligand complex is available. The $k_{cat}$ values of these ligands have been measured under the same experimental conditions using an UV spectrophotometric assay. The calculations point to the crucial role of electric dipole moment of ligands and its interaction with the negatively charged residue Glu225. A striking correlation is found between the energetics associated with this interaction and the $k_{cat}$ values measured under homogeneous conditions. This finding uncovers a fundamental aspect of the mechanism governing substrate diversity and catalytic turnover and thus represents a significant step toward the rational design of novel and powerful prodrugs for antiviral and TKHSV1-linked suicide gene therapies.

The thymidine kinase from herpes simplex virus type 1 (TKHSV1) salvages thymine into the metabolism of the virus by converting it to thymidine monophosphate (1). In contrast to the human isoenzyme, TKHSV1 acts as phosphorylating agent toward a large variety of nucleoside analogs such as (North)-methanocarba-thymidine (n-MCT), aciclovir (ACV), ganciclovir (GCV), and penciclovir (PCV). These analogs exhibit chemical diversity for the nucleobase as well as for the sugar-like chain moiety (2, 3) (Fig. 1). Substrate diversity of TKHSV1 provides the molecular basis for effective and selective treatment of viral infections. It is also exploited for gene therapy approaches involving TKHSV1 as suicide gene by anticancer intervention (4) or the control of graft versus host disease by allogeneic bone marrow transplantation (5).

The recent determination of the x-ray structure of a large spectrum of ligand-enzyme complexes (6–11) has opened a new avenue for the understanding structure-function relationships as well as of the functional role of amino acids involved in binding. It has emerged that the enzyme accommodates the nucleobase and sugar-like chain moieties in two different pockets (P1 and P2 in Fig. 2), interacting with the ligand by a specific and extensive H-bond network.

In the pocket P1, the nucleobase moiety (either a thymine or a guanine ring) is stabilized by direct H bonds with the highly conserved Gln125 and with Arg176 by means of two ordered water molecules. The pyrimidine ring is further fixed by a peculiar interaction with Tyr172 and Met128 (11) in a sandwich-like orientation (Refs. 12 and 13 and Fig. 2).

In P2, the sugar-like chains interact with the protein via its hydroxyl groups. The 3′-OH (and its mimics) forms specific H bonds with Tyr101 and/or Glu225 (Figs. 2 and 3), whereas 5′-OH forms a direct H bond or water-mediated interactions with Arg163, Glu81, and Arg222 (Figs. 2 and 3). In contrast, the polar C-1′–O′–C-4′ function belonging to dT and the corresponding ether groups of ACV and GCV interact neither with polar charged groups of the protein nor with the solvent; instead, they point toward a hydrophobic region made up of the Trp88, Arg163, Glu83, and Arg222 (Fig. 2). This is precisely the reason why TKHSV1 is able to phosphorylate a plethora of analogs that can either inhibit viral DNA polymerase or cause toxic effects when incorporated into viral DNA. To relate structural properties of TKHSV1 ligands to their chemical reactivity we have carried out ab initio quantum chemistry calculations within the density functional theory framework in combination with biochemical studies. Calculations have focused on a set of ligands carrying a representative set of the large spectrum of sugar-mimicking moieties and for which structural information of the TKHSV1-ligand complex is available. The $k_{cat}$ values of these ligands have been measured under the same experimental conditions using an UV spectrophotometric assay. The calculations point to the crucial role of electric dipole moment of ligands and its interaction with the negatively charged residue Glu225. A striking correlation is found between the energetics associated with this interaction and the $k_{cat}$ values measured under homogeneous conditions. This finding uncovers a fundamental aspect of the mechanism governing substrate diversity and catalytic turnover and thus represents a significant step toward the rational design of novel and powerful prodrugs for antiviral and TKHSV1-linked suicide gene therapies.
(Fig. 5), yet the HBPG molecule, which shares the same binding mode as the structurally related prodrug aciclovir (6), inhibits the enzyme. Furthermore, it is not known why the \( k_{cat} \) values of prodrugs are much smaller than that of the natural substrate (see Table I), although the protein-sugar mimicking H bonding interactions is very similar (Fig. 5). Finally, as discussed above, it is rather surprising that the environment accommodating the inner sugar ring C-1\(^9\)–O-4\(^9\)–C-4\(^9\) group of the natural substrate is totally hydrophobic (Fig. 4).

In this paper we address these fundamental issues by performing a combined quantum chemical and biochemical investigation of sugar and sugar-like moieties of substrates and inhibitors of TK HSV1. The theoretical methodology is based on gradient-corrected density functional theory. This approach allows an accurate description of electrostatic and hydrogen-bonded complexes (12, 14–17), such as those investigated here. The biochemical study involves the measurement of the catalytic constant (\( k_{cat} \)) of dT and prodrugs using a modified UV spectrophotometric assay based on pyruvate kinase and lactate dehydrogenase (3). This assay does not require the use of radioactively labeled compounds. This provided a set of homogeneous data, which is necessary for a proper comparison with the quantum chemical calculation. As we will show, the interaction between the electric dipole of the sugar mimicking moiety and the negative charge of the conserved residue Glu\(^{225}\) correlates with the \( k_{cat} \) measurements and thus is a critical factor for the phosphorylation reaction.

**EXPERIMENTAL PROCEDURES**

**Biochemical Essays**

MATERIALS—PreScission protease was purchased from Amersham Pharmacia Biotech. AZT and reagents for enzyme assays were obtained from Sigma. Strain BL21 (Amersham Pharmacia Biotech) served as the expression host. n-MCT and PCV were kindly provided by Dr. Marquez (National Institutes of Health, Bethesda, MD) and Dr. Johannsen (Forschungszentrum Rossendorf Institute of Bioinorganic and Radiopharmaceutical Chemistry Research Center, Rossendorf, Germany), respectively. 5-Trifluoromethyl anhydrohexitoluridine (AHTMU), AHIU analogs, and HBPG were gifts of Dr. De Clercq (Rega Institute Katholieke Universiteit Leuven, Belgium) and Dr. Wright (University of...
Massachusetts, Worcester, MA), respectively. ACV and GCV were purchased from Glaxo-Wellcome and Roche, respectively. dT was obtained from Fluka.

Expression and Purification of TKHSV1—TKHSV1, was expressed as glutathione S-transferase fusion protein in competent Escherichia coli BL21 using the vector pGEX-6P-2-TK (9). The protein was purified by glutathione-affinity chromatography followed by on-column PreScission protease cleavage using a previously described protocol (9). Purification was monitored by SDS-polyacrylamide gel electrophoresis and led to a >90% pure TKHSV1, which was directly used for $k_{cat}$ determination. Total protein concentration was measured using the Bio-Rad protein assay.

Spectrophotometric Assay for $k_{cat}$ Determination—An UV spectroscopic test based on a lactate dehydrogenase-pyruvate kinase coupled assay (3) was employed to monitor ADP formation during substrate phosphorylation. The concentrations of ATP and the ligands were 5 and 1 mM, respectively. These concentrations are at saturation for ATP. For the sake of clarity, the superimposed GCV and PCV sharing the same binding mode of ACV (6) are not shown. The H bond network of ACV is shown as dotted lines as a representative for all guanine derivatives. The concentrations of ATP and the ligands were known from the literature and allows the determination of $k_{cat}$ for the natural substrate and the prodrug share the same binding mode with the 5′-OH mimicking group pointing toward Glu83. AHTMU and AZT (6) have the same binding mode, and they are not shown for clarity. The H bond network of dT is shown for clarity. The H bond network of dT is shown as dotted lines for clarity. The H bond network of dT is shown as dotted lines for clarity.

Quantum Chemical Calculations

Our structural models for quantum chemical calculation on the sugar moiety of dT, ACV, GCV, PCV, n-MCT, and HBPG-TKHSV1 were constructed from the correspondent crystal structures (Protein Data Bank codes 1kim, 2ki5, 1ki2, 1ki3, 1e2k, and 1q1i) (6, 8, 9). The resolution of the structures ranges from 1.7 to 2.4 Å.

The structure of AHTMU-TKHSV1 has not been solved yet, whereas the structure of the complex AHIU-TKHSV1 is known (Protein Data Bank code 1k16) (8). AHIU is chemically and structurally extremely similar to AHTMU (Fig. 1); the two analogs differ only for the group at position 5 of the nucleobase ring (a iodine in AHIU and a CF3 in AHTMU). All available structural data show that this substitution does not affect the binding orientation of the sugar mimicking moiety (8). The initial configuration of AHTMU is therefore built from this x-ray structure.

The x-ray structure of the AZT-TKHSV1 complex is not available, and therefore one has to resort to theoretical structural models. Here we used molecular dynamics-based models reported earlier (18).

The complexes included the sugar-like moieties of the ligands and part of side chains of all of the groups directly interacting with it or forming water-mediated hydrogen bonds (Tyr163, Arg165, Glu83, Glu225, and Arg228). Trp96, Ile97, and Met128 side chains were not included because (i) they would have dramatically increased the size of our model complexes, (ii) they do not contribute significantly to the electrostatic interactions (the focus of the present work) because they form only weak hydrophobic contacts with the sugar, and (iii) the position of these groups is essentially the same in all of the complexes investigated here (6, 8, 9) and therefore their contribution is expected to be rather constant. Hydrogen atoms were added assuming standard bond lengths and bond angles. In Fig. 3 the quantum mechanical model for dT, the natural substrate, is shown in detail. The overall charge of our complexes is 0.

The calculations were performed within the framework of density functional theory (19) in its Kohn and Sham (20) formulation. In this approach the use of gradient corrections is crucial for correctly describe hydrogen bond interactions. Here we used the prescription of Becke and Lee (41, 43) because it has been shown to provide an accurate description of water and hydrogen bonding in biological systems (14, 15, 21, 22). The basis set consisted of a plane wave basis set up to a cut-off of 70 Rydberg, the interactions between valence electrons and ionic cores being described by pseudopotentials of the Martin-Trouiller type (23). Only the Γ point was used. The dT-TK complex was inserted in orthorhombic box with edges $14 \times 15 \times 16$ Å. Similar box sizes were used for all of the other complexes.

In all circumstances the separation between periodic images was at least 6 Å. However, the Coulombic interaction between images was screened using the approach of Barnett and Landman (24). Geometry optimization was performed using the direct inversion iterative subspace (25). C N

![Fig. 5. Active site superposition of representative ligands of the guanine (a) and thymine (b) series.](http://www.jbc.org/)
atoms were constrained to crystallographic positions to take into account the reduced mobility of residues caused by protein environment. A similar procedure has been reported elsewhere (14).

The electronic structure was described with the geometrical analysis of centers of maximally localized Wannier functions (WFC) (26–28). This analysis is useful for investigating the polarity of chemical bonds, because WFC shifts are related quantitatively to the differences $D_x$ of Pauling electronegativities (29) with respect to the non-interacting compounds (22); an increase of $D_x$ corresponds to an increase of polarization of the electronic density in the X-Y bond toward the most electronegative atom X. The effects of the protein electric field were estimated through a comparison of Wannier center shifts for the system in vacuum or in the presence of a protein electric field. The procedure was identical to that of Piana and Carloni (14).

Interaction energies between the substrate and the active site are calculated within the central dipole approximation (30) where the dipole is the result of the quantum calculation. Interaction energy was calculated as $E_{\text{charge-dipole}} = (\xi \cdot \mu R)/(4 \pi \epsilon_0 R^3)$, where $R$ is the distance vector between the centers of charges of the sugar moiety and the residue of charge $\xi$ and $\mu$ is the dipole moment for the sugar moiety.

Even if the expansion is limited to large distances between two interacting groups, this model can provide quantitative results for a comparison between systems that have similar geometries. We have reported the energetics in arbitrary units.

## RESULTS

### Catalytic Activity of TK HSV1

The catalytic activity of TK HSV1 with the compounds listed in Fig. 1 was measured under the same conditions. Table I shows that the measured $k_{cat}$ of dT agrees with those reported in literature (31, 32). Also the $k_{cat}$ of ACV corresponds to that measured in two independent laboratories (33, 34). In the literature a value of 0.015 s$^{-1}$ is also reported for ACV (31), but, considering the newly reported values, this seems to be an outlier. The $k_{cat}$ values of GCV and PCV are also congruent.

![Fig. 6. Electric dipoles of HSV-1 TK sugar-like chains.](image-url)
and sugar moiety dipoles. Because the chemical bond between the nucleobase and the sugar interaction was additive, even if different for thymine and guanine.

The chemical bond between the nucleobase and the sugar interaction was additive, even if different for thymine and guanine.

The chemical bond between the nucleobase and the sugar interaction was additive, even if different for thymine and guanine. The calculated stabilization energy resulting from the electrostatic interaction is the maximum for the natural substrate (Table III). The dipoles of AZT and the inhibitor HBPG do not point toward the negative residue and as a consequence the interaction energy is repulsive.

In an attempt at correlating electrostatic interactions with kinetics data, we first subdivide the compounds into two major classes. The first class includes the compounds bearing a thymine or a thymine-substituted base (namely dT, AHTMU, n-MCT, and AZT); the second class includes the compounds bearing guanine (ACV, HBPG, GCV, and PCV).

The subdivision is necessary as the computational model did not include the base and consequently the electrostatic calculations do not consider the contributions of the nucleobases (12), whose dipole moments are different but additive. In Fig. 7, the correlation between the interaction energy dipole-Glu$^{225}$ and the catalytic constant of the studied ligands appears clearly. The slope of the linear fit depends on the type of sugar mimicking moiety of the ligand. The additivity of the sugar-like chains and the nucleobase dipoles is confirmed by a rigid shift between the linear fit for the guanine and thymine derivatives (Fig. 7).

**Quantum Chemistry Calculations**

Overall, the geometry of the complexes are in very good agreement with the experimental structural data as shown by the low root mean squared deviation values between experimental and optimized structures (Table II). In particular Table III shows that the structural parameters of the geometry optimized systems are in good agreement with experimental data. Thus our first principle approach appears to reproduce very reliably the structural properties of the adducts. Geometry optimization permits also us to establish the H bond network between the substrate and the enzyme that is crucial in determining the electrostatic properties of the adducts and that is only indirectly deducible from the crystal structure where hydrogen atoms are missing.

**Enzyme-Sugar Interactions**—One of the most important physical properties of polar molecules such as ribose and its analogs is given by their electric dipole moment. However, the dipole of the sugar moiety of the natural substrate aligns strikingly to the negatively charged Glu$^{225}$ group. Interestingly, the dipoles of most of the prodrugs investigated here also point to these residues. The calculated stabilization energy resulting from the electrostatic interaction is the maximum for the natural substrate (Table III). The dipoles of AZT and the inhibitor HBPG do not point toward the negative residue and as a consequence the interaction energy is repulsive.

In an attempt at correlating electrostatic interactions with kinetics data, we first subdivide the compounds into two major classes. The first class includes the compounds bearing a thymine or a thymine-substituted base (namely dT, AHTMU, n-MCT, and AZT); the second class includes the compounds bearing guanine (ACV, HBPG, GCV, and PCV).

The subdivision is necessary as the computational model did not include the base and consequently the electrostatic calculations do not consider the contributions of the nucleobases (12), whose dipole moments are different but additive. In Fig. 7, the correlation between the interaction energy dipole-Glu$^{225}$ and the catalytic constant of the studied ligands appears clearly. The slope of the linear fit depends on the type of sugar mimicking moiety of the ligand. The additivity of the sugar-like chains and the nucleobase dipoles is confirmed by a rigid shift between the linear fit for the guanine and thymine derivatives (Fig. 7).

**The Roles of Arg$^{222}$, Arg$^{163}$, and Glu$^{83}$ Interacting Groups**—Although there is no direct correlation between the $k_{cat}$ and the Arg$^{163}$-dipole interaction, the calculated interaction between Arg$^{163}$ and the sugar is important, consistent with previously published mutagenesis experiments (35). From the crystallographic and mutagenesis studies, Arg$^{222}$ and Glu$^{83}$ are essential elements of the kinase machinery; Arg$^{222}$ forms the anion hole with Arg$^{270}$ to make the phosphate atom more electrophilic, and Glu$^{83}$ behaves as base on the O-5’ atom, causing an increase in nucleophily. Thus it is not surprising that the Arg$^{222}$- and Glu$^{83}$-dipole interactions appear not to have any direct correlation with $k_{cat}$ (data not shown), because their influence is based on a totally different mechanisms.

**Sugar-TKHSV1 Hydrophobic Interactions**—Our calculations have pointed to the crucial role of the electric dipole moment of the sugar moieties. We now address the following question. In the natural substrate, is the polar C-1′-O-4′-C-4′ (which faces a hydrophobic pocket; Fig. 4c) important for the correct orientation of the dipole? To answer this question, we calculate the change of electric dipole associated with replacement of O-4′ with the CH$_3$ apolar group (Fig. 4b). Our calculations show that the resulting dipole is both smaller and different in orientation relative to the sugar moiety (Fig. 4, b and c). We conclude that the polar function is essential for a correct alignment of the dipole to the Glu$^{225}$ charge.

**Protein Environment Effects**—The field of the protein may be very important for the chemistry of the active site of this and other enzymes (36). Here we estimate the effect of the protein frame by comparing the electronic structure of the complexes in a vacuum with those in the presence of the protein. A convenient representation of the electronic structure is given by the Wannier centers (26–28) in the presence and the absence of the protein electrostatic potential are represented as yellow and green spheres, respectively.

---

2 The contribution of the nucleobase of the nucleotides to the dipole interaction was additive, even if different for thymine and guanine. Because the chemical bond between the nucleobase and the sugar moiety is nonpolar, we have verified the additivity of the nucleobase and sugar moiety dipoles.
interaction is included in the model we have chosen for quantum mechanics calculations.

**DISCUSSION**

Deciphering the binding mode of prodrugs to the enzyme and their catalytic turnover may help to rationally design more and potent prodrugs for antiviral therapy to overcome the problems of resistance or to be used in enzyme-prodrug gene therapy as well as improve variants of TKHSV1 for gene therapeutic approaches (37).

The mechanisms ruling binding affinities toward TKHSV1 have been previously presented (8, 9, 12, 38). However, the intriguing question posed by the recently published crystal structure of the inhibitor HBPG (6, 9) as to what is the structural basis for the different properties between inhibitor and substrate sharing the same binding mode remained open. This issue prompted us to perform a combined biochemical and quantum chemical study. Two key parameters have been considered: on the one hand binding affinity and on the other hand catalytic turnover.

Because not all $k_{cat}$ values were available and some conflicting kinetics results had been reported in the literature (Table I), the catalytic activity of TKHSV1 toward all ligands has been measured under the same experimental conditions. In this way, a homogeneous set of measured $k_{cat}$ values have been obtained and compared with the calculations.

The calculated sugar dipole points to Glu225, and the resulting stabilization energy correlates strikingly with the catalytic activity (Fig. 7). These results are consistent with a previously proposed stabilization energy correlates strikingly with the catalytic activity (32).

Indeed, replacement of the O-4 gen bond and, more interestingly, a dramatic drop in catalytic turnover may help to rationally design more and potent prodrugs for antiviral therapy to overcome the problems of resistance or to be used in enzyme-prodrug gene therapy as well as improve variants of TKHSV1 for gene therapeutic approaches (37).
The Rational of Catalytic Activity of Herpes Simplex Virus Thymidine Kinase: A
COMBINED BIOCHEMICAL AND QUANTUM CHEMICAL STUDY
Marialore Sulpizi, Pierre Schelling, Gerd Folkers, Paolo Carloni and Leonardo Scapozza

J. Biol. Chem. 2001, 276:21692-21697.
doi: 10.1074/jbc.M010223200 originally published online March 21, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M010223200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 38 references, 7 of which can be accessed free at
http://www.jbc.org/content/276/24/21692.full.html#ref-list-1