PROTEINS: Structure, Function, and Genetics 32:350–361 (1998)

Exploring the Active Site of Herpes Simplex Virus Type-1 Thymidine Kinase by X-Ray Crystallography of Complexes With Aciclovir and Other Ligands

John N. Champness,[1]* Matthew S. Bennett,[1] Frank Wien,[1] Rob Visse,[1] William C. Summers,[2] Piet Herdewijn,[3] Erik de Clercq,[3] Tomasz Ostrowski,[3] Richard L. Jarvest,[4] and Mark R. Sanderson[1]

[1]*Division of Biomedical Sciences, Randall Institute, King's College, London, United Kingdom*
[2]*Department of Therapeutic Radiology, Yale University School of Medicine, New Haven, Connecticut*
[3]*Rega Institute for Medical Research, Katholieke Universiteit Leuven, Leuven, Belgium*
[4]*SmithKline Beecham Pharmaceuticals, New Frontiers Science Park, Essex, United Kingdom*

ABSTRACT **Antiherpes therapies are principally targeted at viral thymidine kinases and utilize nucleoside analogs, the triphosphates of which are inhibitors of viral DNA polymerase or result in toxic effects when incorporated into DNA. The most frequently used drug, aciclovir (Zovirax), is a relatively poor substrate for thymidine kinase and high-resolution structural information on drugs and other molecules binding to the target is therefore important for the design of novel and more potent chemotherapy, both in antiherpes treatment and in gene therapy systems where thymidine kinase is expressed. Here, we report for the first time the binary complexes of HSV-1 thymidine kinase (TK) with the drug molecules aciclovir and penciclovir, determined by X-ray crystallography at 2.37 Å resolution. Moreover, from new data at 2.14 Å resolution, the refined structure of the complex of TK with its substrate deoxythymidine (R = 0.209 for 96% of all data) now reveals much detail concerning substrate and solvent interactions with the enzyme. Structures of the complexes of TK with four halogen-containing substrate analogs have also been solved, to resolutions better than 2.4 Å. The various TK inhibitors broadly fall into three groups which together probe the space of the enzyme active site in a manner that no one molecule does alone, so giving a composite picture of active site interactions that can be exploited in the design of novel compounds. Proteins 32:350–361, 1998.** © 1998 Wiley-Liss, Inc.

Key words: antivirals; Zovirax; drug target; drug binding; enzyme structure; intermolecular interactions

INTRODUCTION

Humans are hosts for at least seven herpesviruses which are responsible for a range of diseases of differing severity. Herpes simplex virus types 1 and 2 (HSV-1 and HSV-2) can cause painful epithelial ulcers in the mouth, on the cornea and genitals, as well as frequently fatal encephalitis. Herpesviruses are highly evolved; HSV-1 codes for at least 70 proteins and has a sophisticated regulatory pathway of early, middle, and late proteins, enabling it to orchestrate lytic production of virus or to enter into a latent state, depending on the cellular environment.[1] HSV-1 thymidine kinase (TK) (ATP: thymidine 5-phosphotransferase; EC 2.7.1.21) is the center of activation of antiviral nucleosides such as aciclovir.[2,3] A similarly acting, but potentially more toxic drug, ganciclovir, is effective in controlling infections with another herpesvirus, cytomegalovirus, as well as HSV-1 and HSV-2. Both drugs attack the replication of the viral genome and virally encoded TK or TK-related enzymes play a part in this process, activating these nucleoside analogs by phosphorylation. The viral TK is more promiscuous than the host-cell TK, which will not phosphorylate these drugs to any appreciable extent. In the case of some analogs, such as aciclovir, the activated drug prevents DNA synthesis by the introduction of a chain-terminating nucleoside at the 3′ end of the growing DNA strand. In the case of some of the simple halogenated nucleosides, such as iododeoxyuridine, the analog is extensively incorporated into DNA, which results in several kinds of DNA malfunctions. These analogs have been used in a virological study of TK mutations[4] and recently have been employed extensively in gene therapy for cancer[5,6] to achieve selective toxicity after introduction of the viral TK

Abbreviations: HSV-1, herpes simplex virus type 1; TK, thymidine kinase; dT, deoxythymidine; IDU, 5-iodo-2′-deoxyuridine; BVDU, (E)-5-(2-bromovinyl)-2′-deoxyuridine; BTDU, 5-(5-bromothien-2-yl)-2′-deoxyuridine; AHIU, 1,5-anhydro-2,3,dideoxy-2-(5-iodouracil-1-yl)-D-*arabino*-hexitol.

Grant sponsor: the Wellcome Trust.

Rob Visse's present address is Department of Molecular Genetics, Gorlaeus Laboratories, Leiden University, PO Box 9502, NL-2300 RA Leiden, The Netherlands.

*Correspondence to: John N. Champness, Division of Biomedical Sciences, Randall Institute, King's College, 26–29 Drury Lane, London, WC2B 5RL, UK. E-mail: john@helios.rai.kcl.ac.uk

Received 12 December 1997; Accepted 24 February 1998

gene into tumor cells, which render them vulnerable to extermination by viral TK-dependent drugs.

While aciclovir (Zovirax) has enjoyed a good reputation for high selectivity and negligible side effects, it has rather poor efficacy. Though the triphosphate of aciclovir is a good inhibitor of DNA synthesis, because of its strong chain-terminating effect, it is itself a poor substrate for TK. Design of an analog providing good binding (low K_m/k_{cat}) to both viral TK and viral DNA polymerase, while retaining low toxicity and high specificity for the viral TK over the cellular TK, would lead to significantly improved chemotherapy. Furthermore, given that TK-deficient mutants of HSV-1 appear to be relatively avirulent in the context of a normal immune system,[7] other analogs which are not substrates, but rather completely novel inhibitors, might prove to be useful new chemotherapeutic agents.

Recent considerations[8] have emphasized the "holistic" approach necessary for successful drug design, including the need to ensure the bioavailability and low toxicity of candidate compounds, and have underlined the difficulties in successful computer modeling of binding energies, even after lead compounds have been examined by structure-based drug-design experiments. Nevertheless, it remains true that good information on the structure of a drug target is a valuable adjunct to devising novel inhibitory compounds,[9] with the work on human immunodeficiency virus protease[10,11] being a prime example. The potential value of structural knowledge of the target for the successful antiviral drug aciclovir led to low-resolution X-ray structures of HSV-1 TK with bound substrate dT,[12,13] and a 2.2Å structure of the TK/ganciclovir complex.[13] Recently, a low-resolution study of three TK/substrate-analogs complexes has been reported,[14] as has the structure of *S. cerevisiae* thymidylate kinase, an enzyme in the same pathway as TK, though of fungal origin.[15] Information on the inhibitory potency of some TK ligands is available,[2,16–20] which has led to computationally based studies on a range of compounds, both novel and established, being used to predict binding modes and strengths theoretically.[19,21]

Data from the crystalline complex of TK with its substrate dT have now been collected to high resolution (2.14 Å), leading to refinement of the structure and assignment of 262 water molecules. An earlier 2.2 Å resolution study of a TK/ganciclovir complex reported without inclusion of solvent or temperature factors[13] has now also been refined. More importantly, the structures of complexes of TK with the drugs aciclovir and penciclovir have also been solved, as have also complexes with certain other 5-substituted uridines.

MATERIALS AND METHODS

TK Crystallization and Data Collection

Thymidine kinase was expressed as described previously.[22] Sequence analysis of protein from a redissolved crystal indicates the presence of residues 11 to 376 of the complete 376 amino acid peptide. As previously reported,[23] the crystals of HSV-1 TK belong to the orthorhombic space group $C222_1$, with unit cell parameters a = 113.4 Å, b = 116.6 Å, c = 108.2 Å for the TK/dT complex. The largest crystals were obtained with the crystallization method described previously[13] which used the hanging-drop technique. Crystals were formed by vapor diffusion in which 8 µl of protein solution (HSV-1 TK 1.0 mg/ml; 40 mM Tris-Cl, pH 7.5; 3 mM DTT; 0.2 mM deoxythymidine) were mixed with 4 µl of precipitating solution (30% saturated ammonium sulphate; 200 mM Tris-Cl, pH 6.75; 3 mM DTT; 0.2 mM deoxythymidine) and equilibrating against the precipitating solution at 25°C. The crystals are well-ordered, with dimensions typically 0.3 × 0.1 × 0.5 mm^3, but anisotropic diffraction limits the resolution by about 0.5 Å in the b^* direction compared with the plane perpendicular to b^*. The exchange of deoxythymidine bound to the protein for other compounds was made by washing the crystals five times in a 3 ml solution (33% saturated ammonium sulphate; 100 mM Tris-Cl, pH 6.75) containing the compound. Crystals were flash-frozen by briefly transferring the crystals to a cryo-protecting solution (33% saturated ammonium sulphate; 100 mM Tris-Cl, pH 6.75; 25% glycerol; plus compound) before being mounted in the cryostream (Oxford Cryosystems, Oxford, UK). Data (1° oscillation frames) were collected on an RAXIS-II imaging plate system fitted with Yale Mirrors (Molecular Structure Corp., Texas), mounted on an R-AXIS 5.0 KW X-ray generator. All datasets were processed with DENZO/SCALEPACK,[24] and intensities and standard deviations were converted to structure factor moduli Fs and µ(F)s using TRUNCATE from the CCP4 suite.[25]

Refinement

The earlier models for the TK/dT and TK/ganciclovir complexes[13] were used as the starting points for the refinement of the atomic positions, derivation of temperature factors, and assignment of water positions using X-PLOR.[26,27] For each of the other ligand-bound TK structures, the coordinates of the latest best TK/dT model (less ligand and active-site waters) were used as a starting point, which were then taken through rounds of rigid-body, positional, and temperature-factor refinement, alternating with modeling sessions. X-PLOR was used to calculate the maps, which were fitted using O.[28] Once finally modeled, the complexes were taken through a double round of positional (100 cycles) and temperature refinement (20 cycles), final data from which are included in Table I. It was necessary to apply an anisotropic overall B-value to most datasets prior to refinement, except in the case of the crystals of the aciclovir, ganciclovir, and BTDU complexes. The eight coordinate data-sets have been deposited with the

TABLE I. X-Ray Data Collection and Refinement Statistics for Binary Complexes With Thymidine Kinase†

(a) **Data collection**

	$C222_1$ Unit cell (Å)							Highest resolution shell			
Complex	*a*	*b*	*c*	No. of frames	Resolution (Å)	Unique data	R_{sym}	Data present	I/σ >2.0	Resolution range (Å)	Multiplicity
dT (native)	113.4	116.6	108.2	220	2.14	39,842	0.081	100%	59%	2.25–2.14	7
Gancyclovir*	114.9	118.7	109.1	78	2.20	28,516	0.061	46%		2.28–2.20	
Pencyclovir	113.8	117.7	108.6	86	2.37	28,282	0.112	93%	38%	2.45–2.37	4
Acyclovir	113.3	116.7	108.4	64	2.33	29,945	0.117	97%	44%	2.51–2.33	6
IDU	113.6	116.6	108.4	90	2.15	38,252	0.070	93%	47%	2.25–2.15	3
AHIU	113.6	116.4	108.4	90	2.37	26,808	0.119	96%	46%	2.48–2.37	4
BVDU	113.4	116.9	108.2	136	2.15	36,243	0.118	92%	31%	2.25–2.15	2
BTDU	114.1	116.0	108.7	60	2.34	29,537	0.144	100%	52%	2.52–2.34	4

*X-ray data of Brown et al.[13]

†Statistics from Denzo/Scalepack processing.[24] Rsym = $\Sigma|I_i - \langle I\rangle|/\Sigma I_i$; where I_i is the measured intensity of an individual reflection, and $\langle I\rangle$ the mean of repeated measurements. For the highest resolution shell of data, completeness is given as % of unique reflections measured, together with % of data for which $I > 2\sigma(I)$ and the minimum multiplicity for at least 50% of the data.

(b) **Refinement**

Complex	No. of data	Resolution (Å)	Completeness (σ-min.)	R-value (R-free)	No. of waters	RMS-bond (Å)	Mean B-value (Å²)
dT (native)	38,108	2.14	96% (2.6)	0.209 (0.279)	262	.013	20.3
Gancyclovir	28,392	2.20	74% (1.0)	0.192 (0.291)	209	.015	22.2
Pencyclovir	20,953	2.37	70% (1.2)	0.200 (0.314)	132	.013	20.7
Acyclovir	21,523	2.37	73% (1.2)	0.226 (0.316)	101	.014	21.7
IDU	28,592	2.20	78% (5.0)	0.206 (0.286)	161	.013	16.6
AHIU	25,870	2.37	88% (2.0)	0.226 (0.306)	250	.015	15.3
BVDU	28,510	2.20	78% (2.8)	0.199 (0.290)	249	.012	18.7
BTDU	25,327	2.34	83% (1.2)	0.210 (0.309)	263	.013	19.0

The number of data given are those extending to the resolution stated and for which $\sigma > \sigma$-min; R-free is for a test set of 10% of data not used in driving refinement. RMS-bond (X-PLOR software) is a measure of the discrepancy between refined and standard bond lengths. The mean B-value is for mainchain atoms after two pairs of positional and B-factor refinement rounds. The R-value is a measure of the agreement between calculated and observed X-ray data: $R = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R-free is similarly defined for the set of reflections excluded from driving refinement.

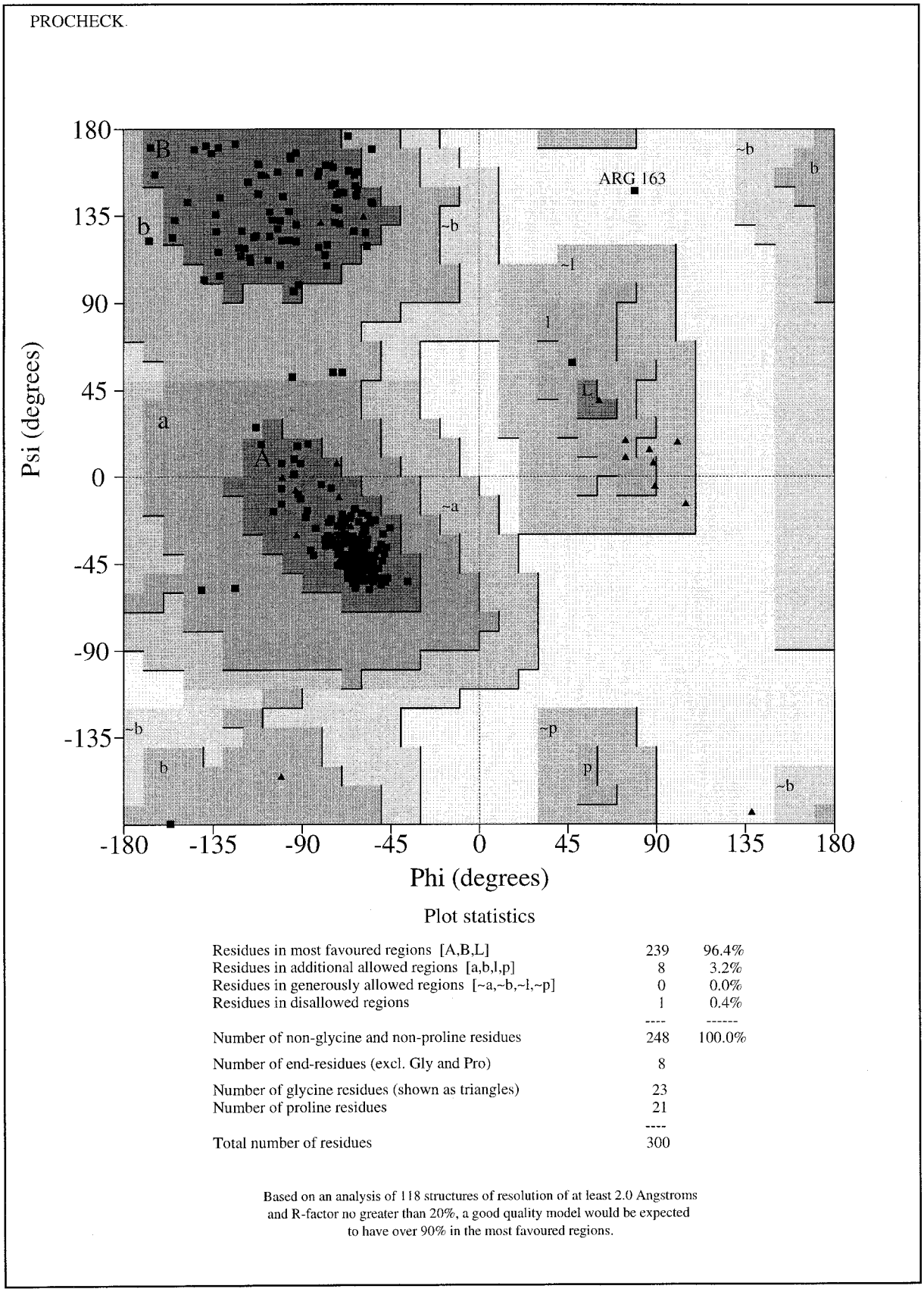
Protein Data Bank at Brookhaven, MA, as 1KIM and 1KI2 to 1KI8.

RESULTS AND DISCUSSION

High-Resolution Structure of the TK/dT Complex

The structure of TK complexed with dT has now been refined to 2.14 Å with data collected from a crystal maintained under cryo-cooled conditions; the R-value for the structure is 0.210 (Table I), in which both monomer chains have been refined independently using 38,108 data for which $F > 2.6*\sigma(F)$ (96% of the unique set), and B-values for individual atoms have been assigned (averaging 20.3 and 22.2 for mainchain and non-mainchain atoms, respectively). The topology of the α-β structure having a core five-stranded parallel-β sheet surrounded by 14 helices has been described previously.[13] A Ramachandran plot of the peptide structure is presented in Figure 1; as previously noted,[14] the one violation of normal peptide conformation is a residue involved in catalysis, Arg-163. The enzyme is present as a dimer in the asymmetric unit of the crystal. The molecules of the dimer have a local twofold axis and interface via the relatively flat surfaces formed by three of the longer helices in each monomer. Nonpolar sidechain interactions predominate here and water molecules are largely excluded: instead, sidechains from two of the interface helices of each monomer make van der Waals contact across the twofold axis; 262 water molecules have been assigned overall. Of these, 150 can be regarded as particularly reliable, having been independently assigned to noncrystallographic symmetry-related, mostly internal, regions of the two monomers.

The N-terminal pole of the parallel β-sheet, containing the N-terminus of the polypeptide, forms part of the molecular surface; this is a region where there are minor chain-breaks in the model owing to poor or nonexistent electron density. The residues concerned are the first 35 (11–45) at the N-terminus of the expressed protein, 70–76 (73–76 in molecule II), 148–153, and 375–376 (C-terminus). A more substantial missing region, 264–279 (263–277 in molecule II), occurs in another region of the molecular surface. The C-terminal end of the β-sheet is the



Plot statistics

Residues in most favoured regions [A,B,L]	239	96.4%
Residues in additional allowed regions [a,b,l,p]	8	3.2%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	1	0.4%
Number of non-glycine and non-proline residues	248	100.0%
Number of end-residues (excl. Gly and Pro)	8	
Number of glycine residues (shown as triangles)	23	
Number of proline residues	21	
Total number of residues	300	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig. 1. Ramachandran plot for thymidine kinase molecule I.

location for the active site, the site otherwise being surrounded by sidechains of residues from four of the helices. This region is occupied in the binary complex structure by deoxythymidine (dT), one of the enzyme's substrates, which is located in a deep pocket of the active site, and is stabilized by both nonpolar van der Waals interactions and direct or water-mediated hydrogen bonds. The other region of the site, near a part of the molecular surface which also contains the C-terminus of the polypeptide, has been identified from the earlier 3.0 Å study[12] as the site of binding of the second substrate, ATP; in our study, a sulphate ion is located here, enfolded in a classic GXGKT (residues 59–63) phosphate binding loop[29] that connects the N-terminal β-strand to the first helix. Nearby is a sequence (212–226) rich in lysine and arginine residues which appear to be able to form a flap that encloses the site. A small part of the structure of the flap, residues 220–223, is rather ill-defined in electron-density maps for molecule II. The definition of these residues is better for molecule I, where the flap is in a different intermolecular environment.

The interactions between dT and the enzyme are detailed in Figure 2a,b, which highlight the hydrogen bonding and nonpolar interactions, respectively. The deoxyribose makes hydrogen bond interactions via its 3′-OH with Tyr-101 and Glu-225, via its 5′-OH with Glu-83, and via its 5′-OH and a water molecule with the sulphate and Arg-163; a second water in this region extends the hydrogen bonding to the sidechain of Asp-162. The thymine base makes van der Waals contacts with Met-128 and Ile-100 on one side of its plane and with Tyr-172 on the other, and makes pairwise hydrogen-bond interaction via its 4-carbonyl and 3-NH groups with the amide of Gln-125. Its 2-carbonyl group is hydrogen bonded to two water molecules, which in turn interact with the guanidinium group of Arg-176, the hydroxyl group of Tyr-101, and the sidechain carbonyl of Gln-125. The 5-methyl group is in van der Waals contact with sidechains of Trp-88, Tyr-132, Arg-163, and is quite close also to Tyr-172 and the conserved Ala-167. The conformation of the sidechain of Tyr-132 is not constant from molecule I to molecule II, leading to slightly different locations for the hydroxyl group. In molecule I, the hydroxyl is linked via a water-mediated hydrogen bond to the carbonyl of Asp-162. In molecule II this hydrogen bond is direct, while the altered location of the aromatic group allows a water to form a hydrogen bond bridge from the carbonyl group of Pro-164 to the amino-group of Ala-168.

Structure of TK/Ganciclovir Complex at 2.2 Å Resolution

Refinement of the complex of TK with ganciclovir was done using the previously collected room-temperature data which led to the earlier model lacking water molecules or temperature factors.[13] The R-value for all 28,392 unique data collected (74% complete) is now 0.192 with 209 waters and average atomic B-factor of 22.2 and 23.8 for mainchain and non-mainchain atoms, respectively. The regions of missing or poor density in the map are approximately the same as for the TK/dT complex described above, except that it was possible to model residues 148–153 in each molecule. The quality of the difference map calculated without ganciclovir added to the model is good for molecule I, the shape of ganciclovir being clear at 3.5*σ contouring. It is rather noisier at the 2.5*σ contouring required to delineate the presence of drug in molecule II. Ganciclovir is an analog of guanosine having a 1,3-dihydroxy-2-propoxymethyl group in place of the deoxyribose moiety (Fig. 3). This acyclic group nevertheless lies in the same volume as that occupied by the deoxyribose of dT in the TK/dT complex, with the two hydroxyl groups occupying positions analogous to the 5′-OH and 3′-OH of dT. The hydroxyls of the acyclic group interact respectively with Arg-163 on the one hand and with Tyr-101 and His-58 on the other (the plane of the His-58 imidazole being slightly changed from the dT-bound conformation). The conformation of the acyclic group is appropriate for the prochiral nature of ganciclovir, which is ultimately phosphorylated as the S-enantiomer of ganciclovir triphosphate.[30] Pursuing the comparison with substrate binding, the guanine moiety lies in a somewhat different position from that of thymine in the TK/dT complex, albeit locating in an approximately common geometric plane. To accompany this difference, the sidechain amide of the Gln-125 sidechain is rotated by 180° and interacts via pairwise hydrogen bonds with 1-NH and 6-carbonyl groups of the drug. The latter group, by contrast with the 4-carbonyl of dT, appears not to be in an optimal location to make further hydrogen bonds, lying as it does 3.5 Å from the Arg-176 guanidinium group, too long for a direct hydrogen bond, too close for a water-mediated bond. The 2-amino group of ganciclovir interacts with a water molecule lying in the vicinity of Ala-168 and Tyr-132 but not mediating further bonds.

Perpendicular to the plane, van der Waals interactions are made with Tyr-172 and Met-231 on one side, and with Met-128 and Ile-100 on the other.

Structure of TK Complexes with the Drug Molecules Aciclovir and Penciclovir

Complexes of these drugs (Fig. 3) with TK were solved to 2.37 Å resolution. Both sets of data were collected from cryo-cooled crystals. For each compound, as for ganciclovir, the difference maps for molecules I and II display contrasting qualities, the interpretation for molecule II being less clear than for molecule I. This is particularly marked in the aciclovir case, where, while density at 3*σ is present for both molecules, it is only for molecule I that it is clearly interpretable in terms of meaningful molecu-

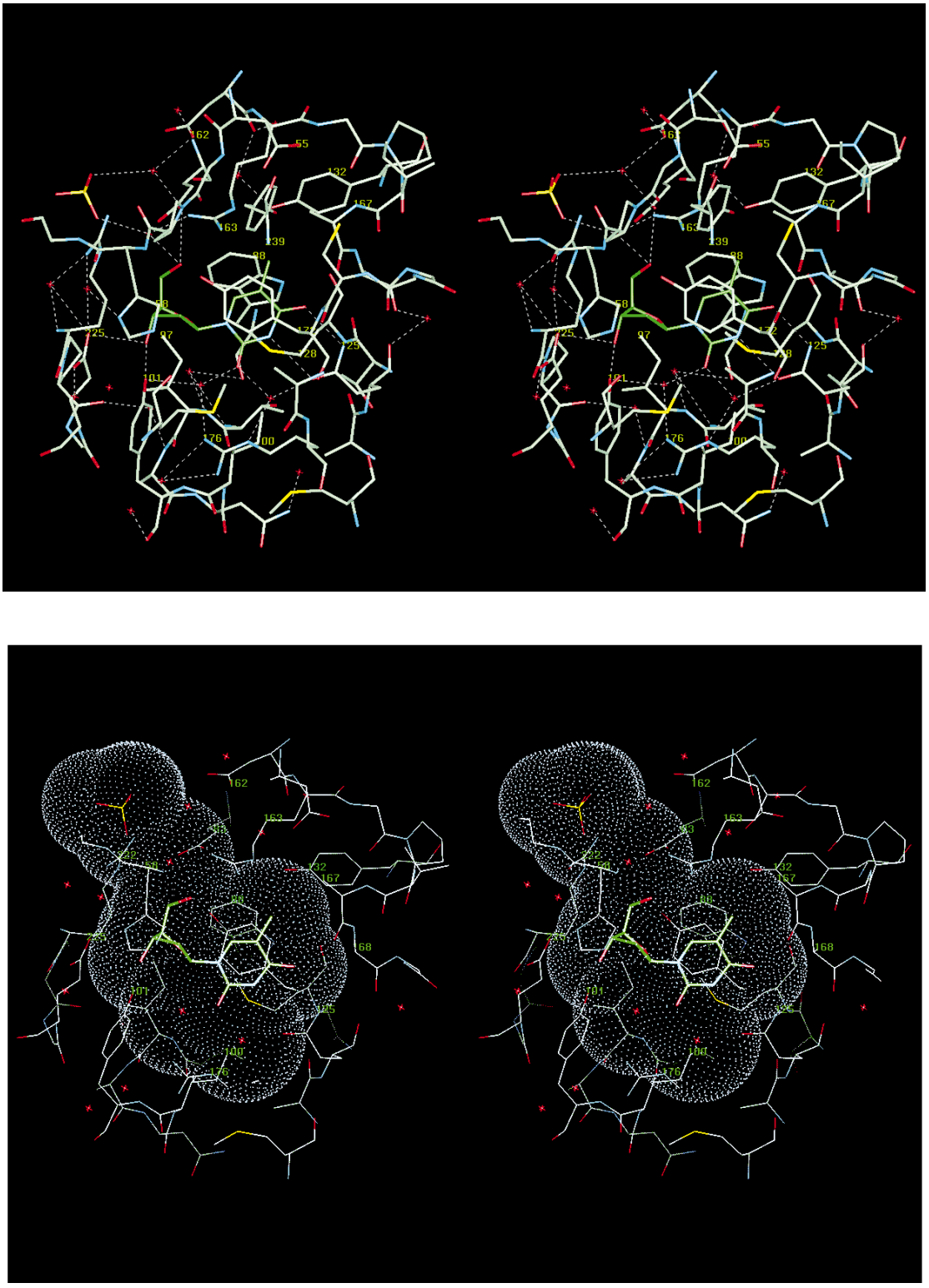
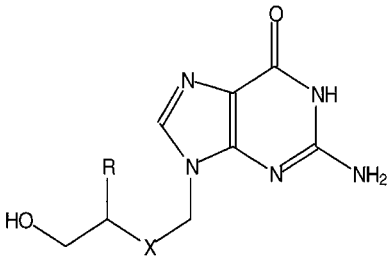
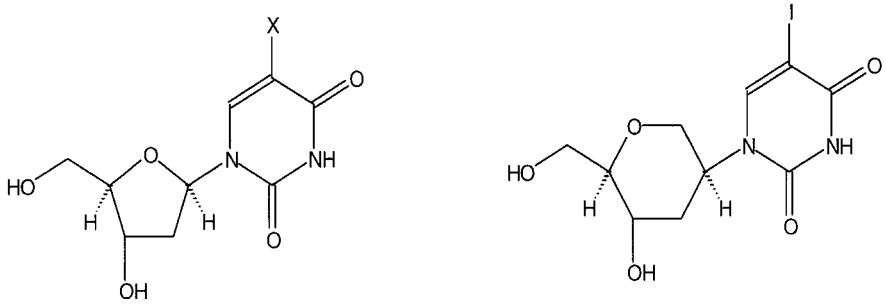


Fig. 2. **Top:** Stereo view of the binding of deoxythymidine binding to TK, showing active-site residues and intermolecular hydrogen bonding. **Bottom:** Extra-radius van der Waals surface of deoxythymidine (QUANTA software) showing residues in close contact with the ligand.

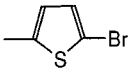


Ganciclovir	X = -O-	R = CH_2OH
Penciclovir	X = $-CH_2-$	R = CH_2OH
Aciclovir	X = -O-	R = H

Acyclic analogues of deoxyguanosine



BVDU X = -CH=CHBr (E-configuration)

BTDU X = 

IDU X = I

AHIU

Fig. 3. Structures of the acyclic deoxyguanosine analogues **(top)** and 5-substituted uracil analogues **(bottom)** complexed with TK.

lar structure and interactions. Given TK's relatively poor affinity for aciclovir, compared, for example, with penciclovir,[2] it could be that molecule II density represents an average from more than one binding mode.

For molecule I of aciclovir and both molecules of penciclovir, difference density maps show a binding mode (Fig. 4a,b) quite similar to that of ganciclovir, the guanine moiety of each lying in a location similar to that occupied by the guanine of ganciclovir, with hydrogen bond pairing being made with Gln-125 via the 1-NH and 6-carbonyl groups; also, Gln-125 is close enough for a possible further hydrogen bond, with the 2-amino group. For both aciclovir and penciclovir, the 6-carbonyl group appears close enough to the guanidinium group of Arg-176 to form a hydrogen bond. Furthermore, in the case of penciclovir molecule II binding, the water molecule that accepts a hydrogen bond from the 2-NH_2 group also donates one to N3 of the ligand.

The acyclic group of aciclovir has, of course, only one potential hydrogen bonding group, the hydroxyl, lying close to the position occupied by the 5′-OH of dT, and interacts with the sidechains of Arg-163 and Glu-83 and a nearby water. The interactions of the acyclic hydroxyl groups of penciclovir molecule II are similar to those of ganciclovir, matching the substrate deoxyribose hydroxyl binding. By contrast with the ganciclovir or dT binding modes, in molecule I of the penciclovir complex, while one acyclic hydroxyl parallels the binding mode of the hydroxyl group of aciclovir, the other is located so as to make a hydrogen bond with Arg-222, rather than mimicking the 3′-OH of deoxyribose by interacting with Tyr-101. Modeling penciclovir in both molecules I and II with the pro-(S) hydroxymethyl group close to Arg-163 and the region of the sulphate (which marks the position of the β-phosphate) gave a good fit to the difference-Fourier density, although the possibility of the alternative conformation, having the pro-(R) group close to the phosphate site, could not be excluded. Better resolution data will be needed to confirm this binding mode, which is consistent with ^{13}C-NMR experiments.[31]

Structures of TK Complexes with Thymidine Analogs

Four 5-substituted uracil derivatives in complex with TK have been examined. Three of these are analogs of dT, having direct substitutions on the 5-position of uracil (Fig. 3): 5-bromovinyl-, 5-bromothienyl-, and 5-iodouridine (BVDU, BTDU,[32] IDU). The fourth compound is an analog of 5-iodouridine, with an anhydrous hexitol group in place of deoxyribose (AHIU[16] Fig. 3). The three deoxyuridine analogs all show features of binding similar to deoxythymidine with respect to the location and interactions of the deoxyribose group. Moreover, the location of the uracil moiety, and interactions of the 3-NH and 4-carbonyl groups with Gln-125, are broadly similar. The bulkier 5-substituents of BVDU and BTDU, the bromovinyl and bromothienyl groups, occupy the deep space available in the neighborhood of residues Trp-88, Tyr-132, Arg-163, and Ala-167. This accommodation is at the expense of a relocation of the sidechain of Tyr-132, the residue that shows distinct conformations in molecules I and II for the TK/dT structure. In the presence of these dT analogs, Tyr-132 of molecule I now has an orientation similar to that for molecule II, shifted away from the ligand to make room for the substituent group, with a consequent alteration to the hydrogen bonding of its sidechain hydroxyl, now donating to the carbonyl group of Asp-162. Positive difference density in the difference Fourier map (Fig. 5a) for the BVDU complex indicates this movement of Tyr-132, as well as the presence of the ligand and active-site waters. The conformation of the planar 5-substituent is such that the torsion angle defined by C4-C5-C5A-C5B is +40°, close to the quasi-coplanarity of the 5-substituent with the uracil reported for the compound from its small-molecule crystal structure determination.[19] For BTDU in complex with TK, the torsion angle C4-C5-C5A-S is −30°. Thus, the thienyl sulphur is within 3 Å of the uracil O4 atom, a proximity and quasi-coplanarity also reported from crystallographic study of BTDU alone;[19] it is also in van der Waals contact with the Cβ atom of Ala-168. The bromine is in van der Waals contact with the carbonyl group of His-164 and with sidechain atoms of Tyr-132.

The compounds IDU and AHIU both show binding modes in which the 5-substituted iodines occupy the same region of the active site, as do the vinyl and thienyl moieties of BVDU and BTDU, but penetrate less deeply, and Tyr-132 of molecule I remains in native conformation. BVDU, BTDU, and IDU have binding modes for the ribose moieties similar to dT, and the anhydrous hexitol of AHIU overlaps this ribose binding mode closely, with its 3′-hydroxyl analog hydrogen bonding to His-58, Tyr-101, and Glu-225 in both molecules. Its 5′-hydroxyl analog hydrogen-bonds to Glu-83. A consequence of the similarity of binding of AHIU to binding of the other dT analogs is that with a 1C chair conformation (consistent with the difference electron density) for the anhydrohexitol moiety, the base and the hydroxymethyl group substitute equatorially and axially, respectively. This is in contrast to the crystal structure determined for the small molecule alone.[17]

Subsites of Ligand Binding

In the site of ribose interaction, while there are similarities in the hydrogen bonding patterns that occur, it is probable that the number of hydrogen bonds is a factor in determining binding strength. Doubtless, the presence of two hydroxymethyl groups in penciclovir and ganciclovir contributes to the stronger binding of these drugs compared with aciclovir.[2] In addition, they provide more complete van der Waals interaction with the ribose site than does the unbranched acyclic group of aciclovir. Against this molecular-interaction picture has to be set the inhibitory effectiveness of the triphosphate of each drug against DNA-polymerase.

For the site of interaction of the various bases and their analogs it is useful to consider three groups of ligands: 1) substrate dT, 2) the acyclic guanosine analogs, and 3) the 5-substituted uracils and AHIU. For dT, defining interactions for the base are hydrogen bonding to Arg-176, Tyr-101 (via waters) and Gln-125, and van der Waals contacts with Met-128 and Tyr-172. These are the same for the uracil moiety of the third group of compounds, the 5-substituted uridines. In the case of the second group of compounds, the larger guanine group of the drug molecules inhabits an environment defined by the same residues, and the volume occupied by guanine is approximately "coplanar" with that occupied by

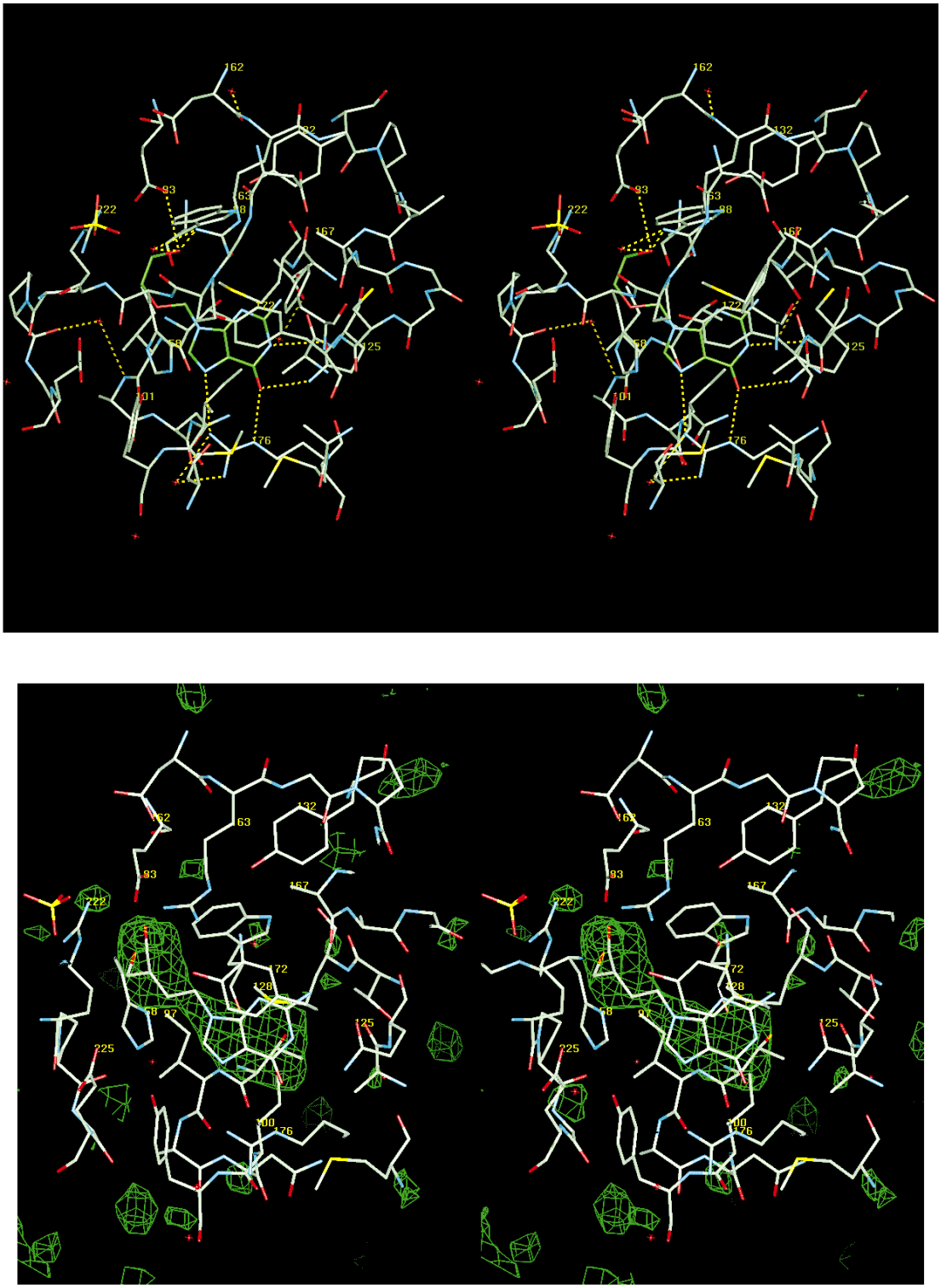


Fig. 4. **Top:** Stereo view of the binding of aciclovir to TK (molecule I), showing active-site residues and intermolecular hydrogen bonding. **Bottom:** Stereo view of the binding of penciclovir to TK (molecule I) superimposed on difference Fourier map contoured at 3σ density.

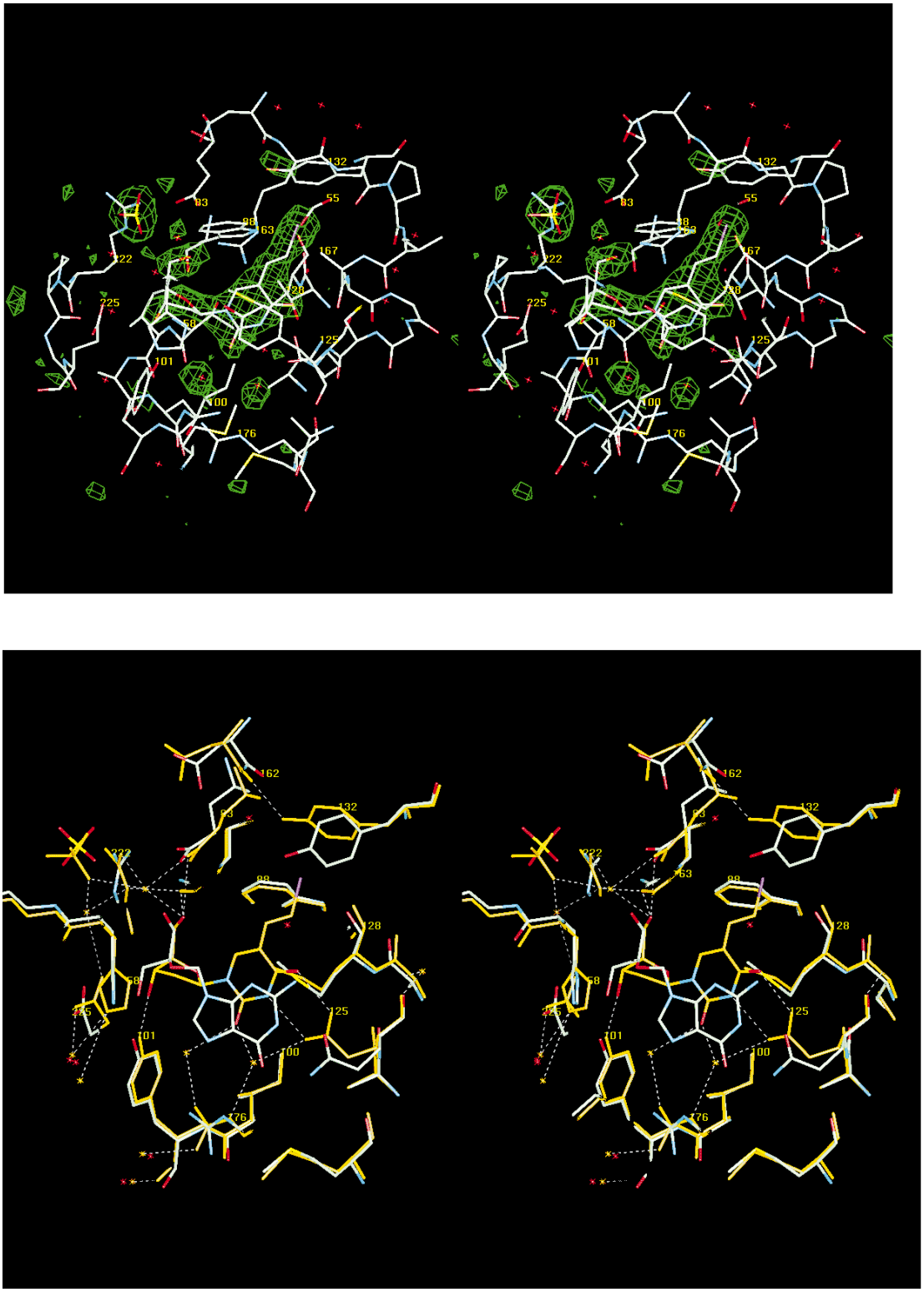


Fig. 5. **Top:** Stereo view of the binding of BVDU binding to TK (molecule I) superimposed on a difference Fourier map contoured at 3σ density. **Bottom:** Stereo view of the active site of the TK/BVDU complex (molecule I) overlaid on the active site of the TK/ganciclovir complex after alignment of enzyme molecules. Intermolecular hydrogen bonding is shown for the TK/BVDU complex.

uracil; however, the guanine is located much closer to Tyr-101 and Arg-176 and direct, not water-mediated, hydrogen bonding is observed in the cases of penciclovir and aciclovir; also, hydrogen bonding with Gln-125 involves a conformational shift of the sidechain, as well as 180° rotation of the amide. A satellite water is present, interacting with the ligand 2-NH_2 group. This water occupies approximately the same space as the 5-methyl group of thymine, highlighting the differing locations of guanine and thymine or uracil ligands. This space is a further significant subsite of the active site, surrounded by predominantly nonpolar groups of Trp-88, Tyr-132, Arg-163, and Ala-168, and it is in this space that the larger substituents of the four nucleoside analogs are located: the bromovinyl and bromothienyl groups of BVDU and BTDU and the iodine of IDU and AHIU. Binding studies have shown BVDU to have the highest affinity for the enzyme.[19] The structural results presented here imply that this could be due to the size of the bromovinyl moiety, by comparison with the iodine of IDU or AHIU or the bromothienyl group of BTDU, being such as to give an optimal fit to this site (allowing for the observed movement of Tyr-132), in terms of van der Waals interactions. The compound 5-thienyluridine, not yet examined by us, shows a strength of binding[18,19] similar to BTDU; it is possible the five-membered ring of this and BTDU is slightly too bulky for the subsite. Figure 5b shows an overlay of the active site regions of the TK/ganciclovir and TK/BVDU complexes, and emphasizes the considerable, albeit flat, volume which is available for ligand binding in the enzyme. The same subsite has been probed by studying the structures of mutated enzyme[4] (M.S.B, J.N.C., F.W., & M.R.S., unpublished).

Attempts have been made to model ligands to the active site of the earlier, incompletely refined structures of TK. In one case,[19] the availability of space for large 5-substituents of uracil, in the context of a dT-like binding mode, was successfully predicted. In another case,[21] a series of N2-phenyl guanines were predicted to bind to TK in a manner distinct from that found for ganciclovir. A preliminary inspection by the present authors indicates that it might, in a ganciclovir binding mode, be possible for an additional N-phenyl group of guanine to be located in the same space as that occupied by the substituent groups of 5-substituted uridines. However, as N-phenyldeoxyguanosine is an inhibitor, rather than a substrate analog,[20] it is reasonable to suppose that its binding mode may be different from those observed in these X-ray studies.

CONCLUSIONS

We report refined X-ray structures of the drug-target thymidine kinase in eight enzyme/ligand complexes: one with one of its natural substrates, dT; three with drug molecules which are acyclic deoxyguanosine analogs, aciclovir, penciclovir, and ganciclovir; and four with 5-substituted uracil derivatives. As well as presenting new complexes, the structure-determinations range from 2.37–2.14 Å resolutions and thus exceed, in quality of data and level of refinement, previously reported HSV-1 TK structures. Detailed information is obtained on the interactions, solvent-mediated or otherwise, that occur between each small molecule and the enzyme. Comparisons between the modes of binding to TK of the various ligands can be summarized as follows: 1) The ribose group and analogs are similarly located in a region of the site rich in polar groups, with generally a constant binding mode for the various hydroxyl groups. The single hydroxyl of aciclovir mimics the 5′-OH interaction of ribose, as do the pro-(S) groups of penciclovir and ganciclovir. 2) Uracil (or thymine) and guanine bases occupy distinct locations on approximately the same geometric plane. With hydrogen bonds being made within the plane and van der Waals interactions perpendicular to it, uracil and guanine show their own specific modes of binding to neighboring residues, reflecting the particular volume the bases occupy. 3) Among the deoxyguanosine analogs, there is only slight variation in the hydrogen bonding pattern for guanine. Among the deoxyuridine analogs, a constant binding mode is observed for the uracil moiety itself; substituents of differing sizes at the 5-position are accommodated in an active-site environment of mostly nonpolar sidechains where a change of conformation of one residue, Tyr-132, accompanies binding of the largest ligands (BVDU and BTDU).

The structural knowledge presented here is a valuable starting point for understanding how the enzyme acts as an initial target for antiviral drugs. It provides a basis for the exploitation of the enzyme structure by design of novel compounds not confined to nucleoside analogs. Such compounds may well be realized as inhibitors, rather than substrate analogs, as is the case with current antiherpetics, and would thus give rise to a new type of therapy in which the virus, while able to persist, may be rendered avirulent. These results are also useful as the initial stage of a comparison of herpes TK with other enzymes of related function and possibly not-dissimilar structure, such as human TK, as well as with the UL97 gene product from cytomegalovirus, which is known to phosphorylate ganciclovir.[33] It is intended to extend the studies of herpes simplex TK to other TK ligands, and also to observe the structural changes that accompany known mutations in the enzyme. Study of the binding of known or new drug molecules to mutant enzyme would be of considerable value in gene therapy systems in which TK is expressed and used as a drug target.

ACKNOWLEDGMENTS

We thank Dr. M. Sohi and Mr. T. Rutherford for expert advice and technical assistance, and Dr. S.M. Fabiane and Ms. C.R. Kirwan for assistance with the illustrations. J.N.C is supported by a grant from the Wellcome Trust.

REFERENCES

1. Roizmann, B., Sears, A.E. Herpes simplex viruses and their replication. In: "Fundamental Virology." Fields, B.N., Knipe, D.M., Howley, P.M. (eds.). New York: Lippincott-Raven, 1996:1043–1107.
2. Griffiths, P.D. Progress in the clinical management of herpesvirus infections. Antiviral Chem. Chemother. 6:191–209, 1995.
3. Darby, G.K. In search of the perfect antiviral. Antiviral Chem. Chemother. 6:54–63, 1995.
4. Black, M.E., Newcomb, T.G., Wilson, H.-M.P., Loeb, L.A. Creation of drug-specific from herpes simplex virus type 1 thymidine kinase mutants for gene therapy. Proc. Natl. Acad. Sci. USA 93:3525–3529, 1996.
5. Borelli, E., Heyman, R., Hsi, M., Evans, R.M. Targeting of an inducible toxic phenotype in animal cells. Proc. Natl. Acad. Sci. USA 85:7572–7576, 1988.
6. Klazmann, D., Philippon, J., Valery, C.A., Bensimon, G. Clinical protocol: Gene therapy for glioblastoma in adult patients: Safety and efficacy evaluation of an *in situ* injection of recombinant retroviruses producing cells carrying the thymidine kinase gene of the herpes simplex type 1 virus, to be followed with the administration of ganciclovir. Hum. Gene Ther. 7:109–126, 1996.
7. Timbury, M.C. "Notes on Medical Virology." London, Harcourt Brace, 1994.
8. Salemme, F.R., Spurlino, J., Bone, R. Serendipity meets precision: The integration of structure-based drug design and combinatorial chemistry for efficient drug discovery. Structure 5:319–324, 1997.
9. Bugg, C.E., Carson, W.M., Montgomery, J.A. Drugs by design. Sci. Am. 269:60–66, 1993.
10. Kempf, D.J., Marsh, K.C., Denissen, J.F., et al. ABT-538 is a potent inhibitor of human immunodeficiency virus protease, and has high oral bioavailability in humans. Proc. Natl. Acad. Sci. USA 92:2484–2488, 1995.
11. Reich, S.H., Melnich, M., Davies, J.F., et al. Protein structure-based design of potent orally bioavailable, nonpeptide inhibitors of human immunodeficiency virus protease. Proc. Natl. Acad. Sci. USA 92:3298–3302, 1995.
12. Wild, K., Bohner, T., Aubry, A., Folkers, G., Schulz, G.E. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. FEBS Lett. 368:289–292, 1995.
13. Brown, D.G., Visse, R., Sandhu, G., et al. Crystal structures of the thymidine kinase from herpes simplex virus type-1 in complex with deoxythymidine and ganciclovir. Nat. Struct. Biol. 2:876–881, 1995.
14. Wild, K., Bohner, T., Folkers, G., Schulz, G.E. The structures of thymidine kinase from herpes simplex virus type-1 in complex with substrates and a substrate analogue. Protein Sci. 6:2097–2106, 1997.
15. Lavie, A., Vetter, I.R., Konrad, M., Goody, R.S., Reinstein, J., Schlichting, I. Structure of thymidylate kinase reveals the cause behind the limiting step in AZT activation. Nat. Struct. Biol. 4:601–604, 1997.
16. Verheggen, I. Van Aerschot, A., Toppet, S., et al. Synthesis and antiherpes virus activity of 1,5-anhydrohexitol nucleosides. J. Med. Chem. 36:2033–2040, 1993.
17. Verheggen, I., Van Aerschot, A., Van Meervelt, L., et al. Synthesis, biological evaluation, and structure analysis of a series of new 1,5-anhydrohexitol nucleosides. J. Med. Chem. 38:826–835, 1995.
18. Luyten, I., De Winter, H., Busson, R., et al. Synthesis of 5-(isothiazol-5-yl)-2′-deoxyuridine and its interaction with the HSV-1 thymidine kinase. Helv. Chim. Acta 79:1462–1474, 1996.
19. DeWinter, H., Herdewijn, P. Understanding the binding of 5-substituted 2′-deoxyuridine substrates to thymidine kinase of herpes simplex virus type-1. J. Med. Chem. 39:4727–4737, 1996.
20. Xu, H., Maga, G., Focher, F., et al. Synthesis, properties and pharmacokinetic studies of N^2-phenylguanine derivatives as inhibitors of herpes simplex virus thymidine kinases. J. Med. Chem. 38:49–57, 1995.
21. Gaudio, A.C., Takahata, Y., Richards, W.G. Prediction of the binding mode of N^2-phenylguanine derivative inhibitors of herpes simplex virus type-1 thymidine kinase. J. Comput. Aided Mol. Des. 12:15–25, 1998.
22. Tung, P.P., Respass, J., Summers, W.C. 3′-Amino thymidine affinity matrix for the purification of herpes simplex virus thymidine kinase. Yale J. Biol. Med. 69:495–503, 1996.
23. Sanderson, M.R., Freemont, P.S., Murthy, H.M.K., Krane, J.F., Summers, W.C., Steitz, T.A. Purification and crystallisation of thymidine kinase of herpes simplex virus type-1. J. Mol. Biol. 202:917–919, 1988.
24. Otwinowski, Z. Oscillation data reduction program. In: "Data Collection and Processing." Sawyer L., Isaacs N.W., Bailey S., (eds.). Warrington, UK: Daresbury Laboratory, 1993:55–62.
25. Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. Acta Crystallogr. D50:760–763, 1994.
26. Brünger, A.T., Kuriyan, J., Karplus, M. Crystallographic *R* factor refinement by molecular dynamics. Science 235:458–460, 1987.
27. Brünger, A.T. The free R-value: A novel statistical quantity for assessing the accuracy of crystal structures. Nature 355:472–474, 1992.
28. Jones, T.A., Zou, J.-Y., Cowan, S.W., Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A47:110–119, 1991.
29. Fry, D.C., Kuby, S.A., Mildvan, A.S. ATP-binding site of adenylate kinase: Mechanistic implications of its homology with *ras*-encoded p21, F_1-ATPase, and other nucleotide binding proteins. Proc. Natl. Acad. Sci. USA 83:907–911, 1986.
30. Karkas, J.D., Germershausen, J., Tolman, R.L., et al. Stereochemical considerations in the enzymatic phosphorylation and antiviral activity of acyclonucleosides: Phosphorylation of 2′-nor-2′-deoxyguanosine. Biochim. Biophys. Acta 911:127–135, 1987.
31. Vere Hodge, R.A., Darlison, S.J., Earnshaw, D.L., Readshaw, S.A. Use of isotopically chiral [4′-^{13}C] penciclovir and ^{13}C-NMR to determine the specificity and absolute configuration of penciclovir phosphate esters formed in HSV-1 and HSV-2 infected cells and by HSV-1 encoded thymidine kinase. Chirality 5:583–588, 1993.
32. Wigerinck, P., Pannecouque, C., Snoeck, R., Claes, P., DeClercq, E., Herdewijn, P. 5-(5-Bromothien-2-yl)-2′-deoxyuridine and 5-(5-Chlorothien-2-yl)-2′-deoxyuridine are equipotent to (E)-5-(2-bromovinyl)-2′-deoxyuridine in the inhibition of herpes simplex virus type I replication. J. Med. Chem. 34:2383–2389, 1991.
33. Darby, G.K. A history of antiherpes research. Antiviral Chem. Chemother. 5:3–9, 1994.