Journal of Computer-Aided Molecular Design, 9 (1995) 473–478
ESCOM



J-CAMD 311

# A pseudoreceptor modelling study of the varicella–zoster virus and human thymidine kinase binding sites

Paulette A. Greenidge*, Alfred Merz and Gerd Folkers**

*Department of Pharmacy, ETH Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

Received 15 May 1995
Accepted 14 August 1995

*Keywords:* Yak; Subtype specificity; Structure–activity relationship

**Summary**

A representative range of pyrimidine nucleoside analogues that are known to inhibit herpes simplex virus (HSV) replication have been used to construct receptor binding site models for the varicella–zoster virus (VZV), thymidine kinase (TK) and human TK1. Given a set of interacting ligands, superimposed in such a manner as to define a pharmacophore, the pseudoreceptor modelling technique Yak provides a means of building binding site models of macromolecules for which no three-dimensional experimental structures are available. Once the models have been evaluated by their ability to reproduce experimental binding data [Vedani et al., J. Am. Chem. Soc., 117 (1995) 4987], they can be used for predictive purposes. Calculated and experimental values of relative binding affinity are compared. Our models suggest that the substitution of one residue may be sufficient to determine ligand subtype affinity.

## Introduction

Active-site models of HSV1 TK and human TK have previously been built by means of homology modelling [1]. However, very few specific interactions between the docked ligands and residues of the receptor models were identified. This obviously limits the utility of these models for drug design. In the current study, we use a newly developed pseudoreceptor modelling technique, Yak [2], to construct binding site pockets for human TK1 and for VZV TK. In contrast to the idea inherent in building by homology, here it is not necessary that the pseudoreceptor model structurally resembles the true biological receptor. The main requirement is that it engages its ligands in sufficient, specific noncovalent interactions so as to accommodate them in a manner similar to the biological receptor. Given a pharmacophore (ligand training set), vectors are generated from ligand functional groups based upon the directionality of molecular interactions. Hydrogen-extension vectors originate at H-bond donors and lone-pair vectors originate at H-bond acceptors. Their end points mark the ideal position for an H-bond acceptor and an H-bond donor, relative to the donor and acceptor atoms, respectively. Hydrophobicity vectors originate at apolar hydrogen atoms; their end points mark the approximate position for a hydrophobic moiety relative to the apolar hydrogen atom. Amino acid residues selected from a preference database of the most frequently occurring functional group–amino acid interactions are then automatically docked and oriented. One criterion used for model validation is the ability to reproduce binding data [2]. Test ligands may then be introduced into the receptor model, and estimates of binding affinity can be obtained relative to the training set ligands. For a detailed description of the Yak methodology and protocol, we refer to Ref. 2. Currently, more effective and selective antiherpes drugs are being sought. One possible strategy is the exploitation of differences between virus-specific enzymes and the corresponding host cell enzymes [3]. Thus, the objectives of this study are twofold: (1) to construct receptor models that are capable of reproducing experimental binding data to within an acceptable precision [2]; and (2) to identify the specific interactions of the ligands that modulate subtype affinity. The results are reported in this paper.

*Present address: Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.
**To whom correspondence should be addressed.

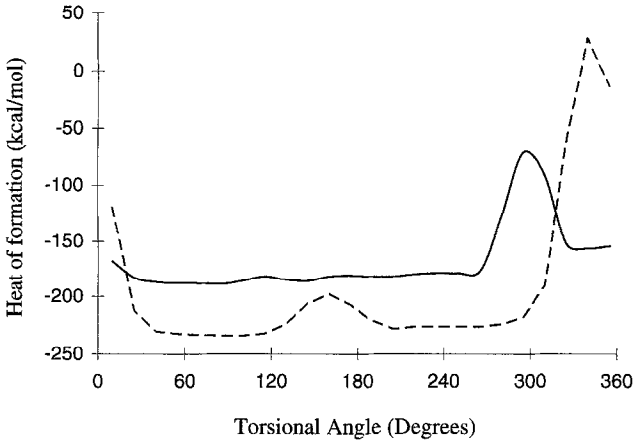



Fig. 1. Dependence of heat of formation upon nucleoside base/sugar C2-N1-C1'-C2' torsional angle. Solid line: compound **1**; dashed line: compound **9** (see Table 1 for structures of the compounds).

## Methods

*Generation of the ligands*

Compound **1** (Table 1) was constructed from a template within SYBYL [4], following a coarse conformational search about the nucleoside base/sugar bond (torsional angle C2-N1-C1'-C2') with the AM1 Hamiltonian [5] in MOPAC 6.0 [6]; the lowest heat of formation energy was achieved for a torsional angle value of 70°. The structure was then geometry optimised with AM1 and the conformational search was reperformed using smaller increments (Fig. 1); the energy profile is quite flat, with the exception of a peak at a torsion angle value of about 300°. This peak is, however, less pronounced if each conformer is reminimized keeping the torsions constrained (results not shown). The O1'-C1'-N1-C6 torsion angle value of 12.7° compares with the value of 3.5° for the AZT crystal structure (molecule B) and the 32–35° range for 5'-phosphothymidylyl(3'-5')thymidine [7]. The 2' and 5' hydroxyl groups of ligands **9**–**11** were involved in intramolecular H-bonding, such that the 2' hydroxyl was not involved in interactions with residues of the receptor model. The free binding energies of ligands **9**–**11** did not suggest that an extra hydrogen bond would be advantageous compared to the other ligands in Table 1 (cf. Table 2). Compound **9**, unlike compound **1**, is an arabinoside (Table 1), so a further conformational search about the nucleoside base/sugar bond was conducted. However, energy minima were found to be located in a similar region (Fig. 1). Thus, an initial torsional angle of 70° was also used for compound **9**. NMR data suggest that the type of nucleoside, sugar pucker and orientation about the glycosyl (C1'-N1) bond are correlated [8]. The O4'-C1'-N1-C2 torsional angle values of −172.0° and −169.8° for the optimised geometries of ligands **1** and **9**, respectively, are in keeping with the known *anti* preference of pyrimidine nucleosides [8]. The *+synclinal* conformation about the C4'-C5' bond (O5'-C5'-C4'-C3' torsional angle values of 52.8° and 49.1°, respectively, for ligands **1** and **9**) is consistent with the available experimental data [8], as are the C2'-*endo* and C3'-*endo* sugar puckering modes adopted by ligands **1** and **9**, respectively (see Ref. 8 for a

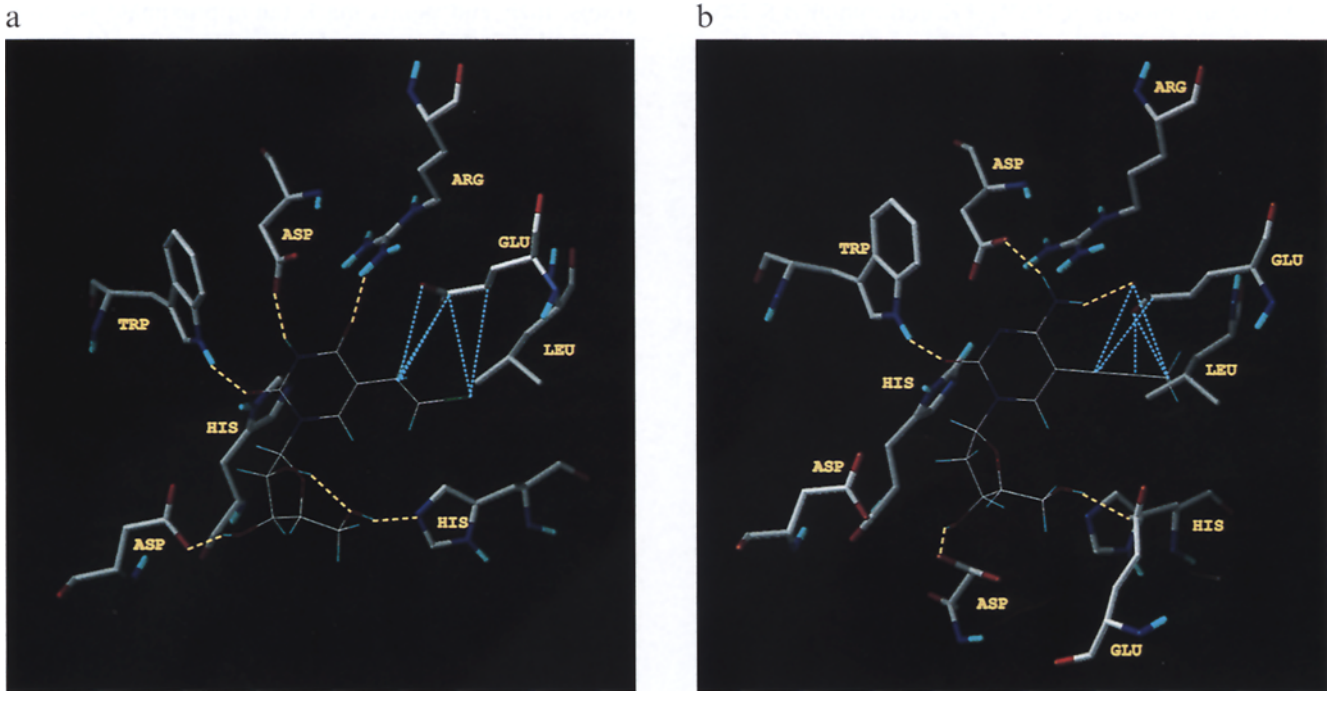



Fig. 2. Views of the VZV TK pseudoreceptor–ligand complex, showing the most important ligand–receptor contacts. (a) Ligand **10** (some residues have been omitted to aid visual clarity); (b) ligand **7** (all residues are shown). Yellow dashed lines represent hydrogen bonding and blue lines indicate significant van der Waals contacts.

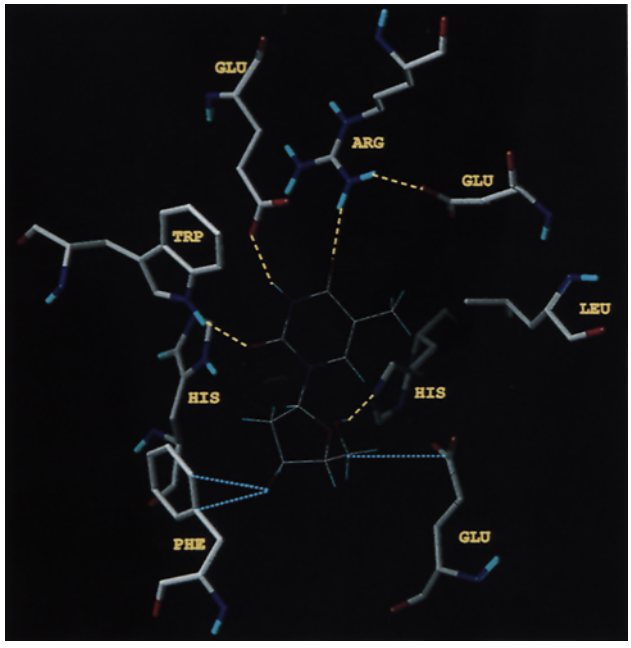



Fig. 3. View of the human TK1 pseudoreceptor–ligand complex with ligand **1**, showing the most important ligand–receptor contacts. The colour coding is as given in the legend to Fig. 2.

detailed discussion on preferred conformations of nucleosides). Ligands **1** and **9** were used as a base from which the other ligands were constructed; the 5-substituent does not greatly affect the overall conformation of a molecule [9]. MNDO [10] electrostatic potential charges were generated for all geometry-optimised ligands.

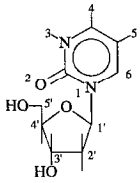
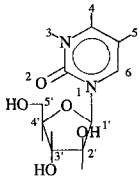
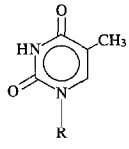
*VZV and human TK1 models*

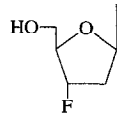
Experimental free energies of ligand binding, $\Delta G^{\circ}_{exp}$, were derived from $K_m$ values [11,12] in the absence of $K_d$ data, and in this instance thus represent only pseudo free energies of ligand binding. Compounds **6**, **7**, **10** and **13** comprised the VZV TK Yak training set, thus all four types of 5-substituent were represented, as were the two types of 4-substituent and a variety of sugar substituents (Table 1). The free binding energies of these ligands effectively span the whole range (Table 2). They were composed into a pharmacophore by using an rms overlay fitting procedure in which molecule **10** served as the template. The test set consisted of the other ligands (with the exception of compound **5**, for which VZV TK binding data were not available). The human TK1 training set was made up of ligands **1**, **2**, **12** and **13**, and ligand **5** was the sole test ligand. Relative free energies of binding were calculated according to the approximation given in Eq. 1 [2]:

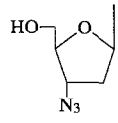
$$\Delta(\Delta G^{\circ}_{calc}) \cong \Delta(\Delta E_{calc}) - \Delta(\Delta G_{solvation,ligand}) \qquad (1)$$

where $\Delta E_{calc}$ is the interaction energy between an individ-

TABLE 1
STRUCTURES OF THE PYRIMIDINE NUCLEOSIDES UNDER INVESTIGATION

| No. | Structure | R | Position 3 | Position 4 | Position 5 | Abbr. | Trivial name |
|---|---|---|---|---|---|---|---|
| 1 |  | | H | O | $CH_3$ | dThd | Thymidine |
| 2 | | | H | O | H | dUrd | Deoxyuridine |
| 3 | | | H | O | CH=CHBr | BVdU | Bromovinyl-deoxyuridine |
| 4 | | | H | O | C≡CCH$_3$ | | 5-(1-Propynyl)-2'-deoxyuridine |
| 5 | | | H | O | F | FdUrd | Fluoro-deoxyuridine |
| 6 | | | – | $NH_2$ | H | dCyd | Deoxycytidine |
| 7 | | | – | $NH_2$ | C≡CCH$_3$ | | 5-(1-Propynyl)-2'-deoxycytidine |
| 8 | | | – | $NH_2$ | $CH_3$ | | 5-Methyl-deoxycytidine |
| 9 |  | | H | O | $CH_3$ | araT | Arabinofuranosyl-T |
| 10 | | | H | O | CH=CHBr | | Arabinofuranosyl-bromovinyl-T |
| 11 | | | H | O | C≡CCH$_3$ | | Arabinofuranosyl-5-(1-propynyl)-T |
| 12 |  | a | | | | FLT | 3'-Fluoro-thymidine |
| 13 | | b | | | | AZT | 3'-Azido-thymidine |
| 14 | | c | | | | d4T | |

a 
2',3'-dideoxy-3'-fluoro-β-D-ribofuranosyl

b 
2',3'-dideoxy-3'-azido-β-D-ribofuranosyl

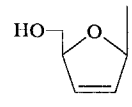
c 
2',3'-dideoxy-2',3'-didehydro-β-D-ribofuranosyl

TABLE 2
COMPARISON OF CALCULATED AND EXPERIMENTAL RELATIVE FREE BINDING ENERGIES OF PYRIMIDINE LIGANDS TOWARDS THE VZV TK RECEPTOR

| Ligand[a] | $\Delta(\Delta G^\circ_{exp})$ (kcal/mol) | $\Delta(\Delta G^\circ_{corr})$ (kcal/mol) | $\Delta(\Delta G^\circ_{corr}) - \Delta(\Delta G^\circ_{exp})$ (kcal/mol) |
|---|---|---|---|
| **Training set**[b] | | | |
| **10**[c] | 0.0 | 0.0 | 0.0 |
| **7** | 2.72 | 2.33 | −0.39 |
| **6** | 3.19 | 3.01 | −0.18 |
| **13** | 3.38 | 3.52 | 0.14 |
| **Test set**[d] | | | |
| **3** | 0.22 | 0.28 | 0.06 |
| **1** | 1.03 | 1.74 | 0.71 |
| **4** | 1.76 | 1.79 | 0.03 |
| **9** | 1.79 | 1.46 | −0.33 |
| **11** | 2.65 | 3.01 | 0.36 |
| **8** | 2.93 | 3.10 | 0.17 |
| **12** | 3.20 | 3.52 | 0.32 |
| **2** | 3.74 | 2.06 | −1.68 |
| **14** | 4.21 | 4.27 | 0.06 |

[a] See Table 1 for structures of the ligands.
[b] The correlation coefficient for $\Delta(\Delta G^\circ_{exp})$ versus $\Delta(\Delta G^\circ_{calc})$ is 0.989; the rms deviation of $\Delta(\Delta G^\circ_{calc})$ and $\Delta(\Delta G^\circ_{exp})$ is 0.23 kcal/mol.
[c] Ligand **10** is the reference compound.
[d] The rms deviation of $\Delta(\Delta G^\circ_{corr})$ and $\Delta(\Delta G^\circ_{exp})$ is 0.64 kcal/mol.

ual ligand and the pseudoreceptor as calculated by Yak, and $\Delta G_{solvation,ligand}$ is the free energy of ligand solvation, rescaled by means of a linear regression (LR) between $\Delta(\Delta G^\circ_{calc})$ and $\Delta(\Delta G^\circ_{exp})$ according to Eq. 2:

$$\Delta(\Delta G^\circ_{corr}) = \text{slope}^{LR}\,\Delta(\Delta G^\circ_{calc}) \qquad (2)$$

Free energies of ligand solvation were calculated semi-analytically following the approach of Still et al. [13]. For a detailed description of the Yak pseudoreceptor modelling methodology, we refer to Ref. 2. All visualizations and Yak calculations were carried out on a Silicon Graphics XS-24 or INDY workstation and semiempirical calculations were performed using an IBM RS/6000 Model 550.

## Results

*VZV and human TK1 models*

Pharmacophore–pseudoreceptor complexes for the VZV TK and TK1 models (with one representative ligand) are shown in Figs. 2 and 3, respectively, and the corresponding $\Delta(\Delta G^\circ_{corr})$ and $\Delta(\Delta G^\circ_{exp})$ are compared in Tables 2 and 3. The residues composing the VZV TK and TK1 models were Hid-Trp-Hid-Glu-Arg-Asp-Asp-Asp-Glu-Leu and Hid-Trp-Hid-Glu-Glu-Arg-Glu-Phe-Leu, respectively; they were selected based upon suggestions obtained from the Yak preference database. For the training set ligands, the correlation coefficients for $\Delta(\Delta G^\circ_{exp})$ versus $\Delta(\Delta G^\circ_{corr})$ are 0.989 and 0.985, respectively, and the rms deviations of corrected and experimental differences in free energies, $\Delta(\Delta G^\circ)$, are 0.23 and 0.09 kcal/mol, respectively. The rankings of the ligands with respect to their $\Delta(\Delta G^\circ)$ are correctly produced. The test ligands were then added to the pseudoreceptor models and refined in both position and orientation; the pseudoreceptor models were not altered. The rms deviation for the VZV TK predicted free energy differences is 0.64 kcal/mol. The deviation between the predicted and experimental value for **5**, the only TK1 test ligand, is 0.73 kcal/mol.

## Discussion

*Interactions of the ligands with the VZV TK pseudoreceptor*

In general, ligands having large 5-substituents possess better free binding energies than equivalent ligands with smaller substituents (cf. Tables 1 and 2). In our model, this is simulated by engaging the 5-substituent in van der Waals interactions with residues of the pseudoreceptor (Fig. 2). Most contact is made with the glutamate residue that hydrogen bonds with the amino group of compounds **6–8** (Fig. 2b), even though a leucine residue is also present. The finding that the relative free binding energy of ligand **2** (with a small 5-substituent) differs by −1.7 kcal/mol from the experimental value may be an indication that the extent of the van der Waals interactions of ligands having large 5-substituents is underestimated by our model.

Hydrogen bonds are formed by the O2, O4 and N3H atoms of the ligands with tryptophan, arginine and aspartate residues, respectively (Fig. 2). In addition, the O2 of compound **14** is hydrogen bonded to a histidine, whereas the other ligands merely make van der Waals contacts with this residue. The 5' hydroxyl groups of ligands **1–4** and **9–13** hydrogen bond to a common histidine, and the

TABLE 3
COMPARISON OF CALCULATED AND EXPERIMENTAL RELATIVE FREE BINDING ENERGIES OF PYRIMIDINE LIGANDS TOWARDS THE HUMAN TK1 RECEPTOR

| Ligand[a] | $\Delta(\Delta G^\circ_{exp})$ (kcal/mol) | $\Delta(\Delta G^\circ_{corr})$ (kcal/mol) | $\Delta(\Delta G^\circ_{corr}) - \Delta(\Delta G^\circ_{exp})$ (kcal/mol) |
|---|---|---|---|
| **Training set**[b] | | | |
| **1**[c] | 0.0 | 0.0 | 0.0 |
| **13** | 0.11 | 0.18 | 0.07 |
| **12** | 0.89 | 0.64 | −0.25 |
| **2** | 1.78 | 1.81 | 0.03 |
| **Test set**[d] | | | |
| **5** | 0.92 | 1.65 | 0.73 |

[a] See Table 1 for structures of the ligands.
[b] The correlation coefficient for $\Delta(\Delta G^\circ_{exp})$ versus $\Delta(\Delta G^\circ_{calc})$ is 0.985; the rms deviation of $\Delta(\Delta G^\circ_{calc})$ and $\Delta(\Delta G^\circ_{exp})$ is 0.09 kcal/mol.
[c] Ligand **1** is the reference compound.
[d] The rms deviation of $\Delta(\Delta G^\circ_{corr})$ and $\Delta(\Delta G^\circ_{exp})$ is 0.73 kcal/mol.

3' hydroxyl groups of ligands **1–4** and **9–11** hydrogen bond to a common aspartate residue (Fig. 2a). Ligands **6–8**, however, are forced to adopt a slightly different position in the binding site because of the unfavourable steric separation between their amino moiety and the arginine residue (Fig. 2b). Consequently, their 3' hydroxyl groups form a hydrogen bond with an aspartate residue, but with a different one compared to the uracil compounds; the 5' hydroxyl groups form hydrogen bonds with a glutamate residue. In the biological receptor, phosphorylation occurs at the 5' hydroxyl. However, in our receptor model no attempt was made to specifically simulate this.

*Interactions of the ligands with the human TK1 pseudoreceptor*

The pseudoreceptor model (Fig. 3) is based on the one previously described, as they share four common ligands. The only major difference, apart from the reminimization of the ligand–pseudoreceptor complex, is the replacement of the aspartate residues that interact with the 3'-substituent by a phenyl residue. This modification, achieved using the tools for residue mutation and manipulation contained within Yak, was necessary, because otherwise the ligands lacking the hydroxyl substituent are calculated to have much poorer relative binding energies than is experimentally observed. It is documented that aromatic rings are capable of acting as hydrogen-bond acceptors [14]. In general, following minimization of the whole pseudoreceptor complex, the orientation in space of the residues remained similar. However, the aspartate residue moves away from the 5-substituent, as ligands **6–8** are no longer present in the binding site to hold it in position. Instead, it forms a hydrogen bond with the side chain of the arginine residue. This leaves more space for the leucine residue; in this model it is located directly along the bond of the 5-substituent.

*Subtype specificity of the ligands*

Even though it is not possible to directly compare binding data from different studies, the data still give some indication about the relative binding affinity of the ligands in the two kinases. Tables 2 and 3 clearly indicate that in human TK1, the relative difference in experimental free binding energy between compounds **1** and **13** is much reduced compared to the situation in VZV TK. Our two models were designed to cope with this difference. In the VZV TK model the 3'-substituent of the ligands interacts with an aspartate residue. In the case of ligand **1**, this involves formation of a strong hydrogen/electrostatic bond, but for ligand **13** it involves only van der Waals–electrostatic interactions. However, in the human TK1 model the hydrogen bond formed by ligand **1** with the centre of the phenyl ring is weaker. Since both ligands are identically substituted at the 5-position (Table 1), it then seems reasonable that both ligands should have more similar free binding energies.

No comments can be made about compound **5**, since VZV TK binding data are not available for this ligand. The difference in relative experimental free binding energy between molecules **1** and **2** is reduced in the human TK1 compared to the VZV TK receptor. According to our models, since the 3' hydroxyl groups of both ligands experience a reduction in the strength of the hydrogen bond formed with the receptor in going from the VZV TK to the human TK1 model, we would have expected the relative free binding energy difference to remain effectively constant. However, another factor to consider is the reduced importance of van der Waals interactions between the 5-substituent and the pseudoreceptor in the latter model. Thus, ligands with small 5-substituents should be less discriminated against in the TK1 model compared to the VZV TK model. The experimental free binding energy of ligand **12** relative to **1** is reduced in human TK1 compared to VZV TK (Tables 2 and 3). Our explanation for this reduction, provided above for compound **13**, should also hold for compound **12**. Ligand **14** is not a substrate for TK1 [10]. The predicted $\Delta(\Delta G^\circ)$ difference between ligands **14** and **1** in VZV TK is ~3 kcal/mol. In TK1, this value increases to ~9 kcal/mol, although a larger test set would have been preferable in order to validate this model.

## Conclusions

Pseudoreceptor models based upon a pharmacophore of interacting ligands (training set) were built for the varicella–zoster virus and human cytosolic thymidine kinases. The computed relative binding affinities agree in a semi-quantitative manner with the available experimental data. The VZV TK model was also shown to have good predictive capabilities for the test set ligands (ligands different from the training set). The functional groups of the ligands engage in specific interactions with residues of the receptor model, so it may be possible to use the models for de novo ligand design purposes. Previously constructed active-site models [1] are very much less well defined, to such an extent that to make a comparison with our present model is not useful.

The 3' and 5 positions display the greatest substituent variability among the ligands. In VZV TK, ligands having a hydroxyl group as their 3'-substituent display much better binding affinities compared to those that have not (Tables 1 and 2). In human TK1, this superiority is eroded (Table 3). This is simulated in our receptor models by exchanging an aspartate (strong hydrogen-bonding partner) by a phenylalanine (weaker hydrogen-bonding partner). However, a structural alteration of the receptor, which only allowed for less than optimal hydrogen bonding with the 3' hydroxyl, is also a feasible alternative.

While it is not assumed that our models structurally resemble the true binding site or that an exact correspondence exists between the amino acid residues of the models and those composing the pockets of the enzymes, there should be some analogy in the types of amino acids, e.g., hydrophobic or hydrogen bonding. An assessment of the utility of the models will be provided by their ability to support the design of effective and selective ligands.

## Acknowledgements

We thank Drs. Petra Kern, Martina Michael, Angelo Vedani and Peter Zbinden for their helpful discussions and suggestions.

## References

1 Folkers, G., Trumpp-Kallmeyer, S., Gutbrod, O., Krickl, S., Fetzer, J. and Keil, G.M., J. Comput.-Aided Mol. Design, 5 (1991) 385.
2 Vedani, A., Zbinden, P., Snyder, J.P. and Greenidge, P.A., J. Am. Chem. Soc., 117 (1995) 4987, and references cited therein.
3 Goodchild, J., Porter, R.A., Raper, R.H., Sim, I.S., Upton, R.M., Viney, J. and Wadsworth, H.J., J. Med. Chem., 26 (1983) 1252.
4 TRIPOS Associates, Inc., St. Louis, MO.
5 Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P., J. Am. Chem. Soc., 107 (1985) 3902.
6 Stewart, J.J.P., J. Comput.-Aided Mol. Design, 4 (1990) 1.
7 Dewar, M.J.S. and Thiel, W., J. Am. Chem. Soc., 99 (1977) 4899.
8 Saenger, W., Principles of Nucleic Acid Structure, Springer, New York, NY, 1988, pp. 51–104.
9 Kulikowski, T., Pharm. World Sci., 16 (1994) 127.
10 Camerman, A., Mastropaolo, D. and Camerman, N., Proc. Natl. Acad. Sci. USA, 84 (1987) 8239.
11 Roberts, G.B., Fyfe, J.A., McKee, S.A., Rahim, S.G., Daluge, S.M., Almond, M.R., Rideout, J.L., Koszalka, G.W. and Krenitsky, T.A., Biochem. Pharmacol., 46 (1993) 2209.
12 Munch-Petersen, B., Cloos, L., Tyrsted, G. and Eriksson, S., J. Biol. Chem., 266 (1991) 9032.
13 Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., J. Am. Chem. Soc., 112 (1990) 6127.
14 Levitt, M. and Perutz, M.F., J. Mol. Biol., 201 (1988) 751.