

ORIGINAL ARTICLE

Molecular modelling studies of new potential human DNA polymerase α inhibitors

Barbara Zdrazil^{1,*}, Anja Schwanke², Birte Schmitz¹, Monika Schäfer-Korting², and Hans-Dieter Höltje¹

¹Department of Pharmaceutical/Medicinal Chemistry, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany, and ²Institute for Pharmacy (Pharmacology and Toxicology), Freie Universität Berlin, Königin-Luise-Straße 2+4, 14195 Berlin, Germany

Abstract

The human polymerase α (pol α) is a promising target for the therapy of cancer e.g. of the skin. The authors recently built a homology model of the active site of human DNA pol α . This 3D model was now used for molecular modelling studies with eight novel analogues of 2-butanilino-dATP, which is a highly selective nucleoside inhibitor of mammalian pol α . Our results suggest that a higher hydrophobicity of a carbohydrate side chain (pointing into a spacious hydrophobic cavity) may enhance the strength of the interaction with the target protein. Moreover, acyclic acyclovir-like derivatives outperformed those with a sugar moiety, indicating that structural flexibility and higher conformational adaptability has a positive effect on the receptor affinity. Cytotoxicity tests confirmed our theoretical findings. Besides, one of our most promising compounds in the molecular modelling studies revealed high selectivity for the SCC-25 cell line derived from squamous cell carcinoma in man.

Keywords: Actinic keratosis, molecular dynamics simulations, docking, X-Score

Introduction

The superfamily of DNA-dependent DNA polymerases contains essential enzymes for genome replication. They catalyse the synthesis of DNA, built up from deoxyribonucleotides. The mechanism of replication by base incorporation is believed to be identical among all different types of polymerases. According to amino acid sequence analogies DNA polymerases are divided into five families: DNA polymerases I or A family, DNA polymerase α (pol α) or B family, the reverse transcriptase family, the DNA polymerase β family, and the bacterial DNA polymerase III family. Although, it is commonly supposed that all DNA polymerases share a common overall architecture with a 'right hand-shape' and its subdomains determined as 'thumb', 'palm', and 'fingers', only the palm domain is homologous among the different families. Moreover, also within the families the proteins vary a lot in length and amino acid sequences¹.

Representative members of the pol α family are the prokaryotic polymerase II, several eukaryotic and archeal polymerases, as well as the polymerases from viral adenovirus, herpes simplex virus, and phages RB69, T4, and T6. Although crystal structures and nuclear magnetic resonance data of several type B DNA polymerases are available, the 3D structure of the human pol α has yet to be unravelled. Recently, we reported a fractional homology model for the tertiary structure of human pol α based on the crystallographic coordinates of pol α in bacteriophage RB69². Due to a very low overall sequence identity of only ~11%, our model is restricted to the region 836–1102 which is containing the catalytic subunit of the enzyme: the thumb, palm, and fingers region. In these regions, there are at least 20% identical and 28% similar residues.

Inhibitors of pol α are interfering with DNA synthesis. In general, competitive nucleoside inhibitors are 3-deoxy-ribose analogues and have to be transformed by

*Present Address: University of Vienna, Department of Medicinal Chemistry, Althanstraße 14, 1090 Wien, Austria.

Address for Correspondence: Dr. Barbara Zdrazil, University of Vienna, Department of Medicinal Chemistry, Althanstraße 14, Vienna, A-1190 Austria. E-mail: barbara.zdrazil@univie.ac.at

(Received 10 March 2010; revised 12 June 2010; accepted 18 June 2010)

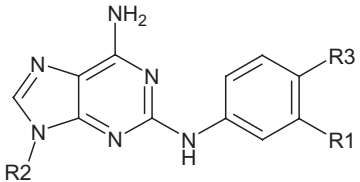
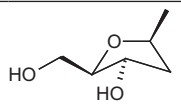
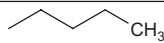
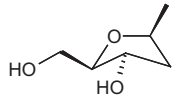
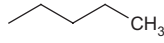
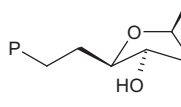
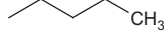
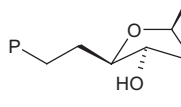
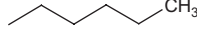
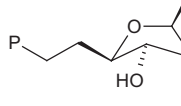
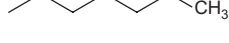
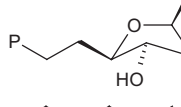
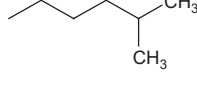
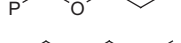
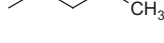
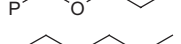
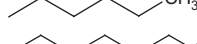
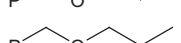
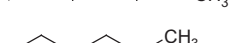

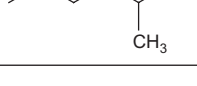
cellular kinases into triphosphates (TPs), which then act as chain terminators after incorporation into the DNA due to the lack of the 3'-OH function. Such chain terminating analogues are used in DNA sequencing³ and clinically for the treatment of viral diseases, like herpes virus and HIV infections⁴. Some nucleoside pol α inhibitors, such as cytarabine and fludarabine, are used systemically for the treatment of leukaemia and gemcitabine in pancreas carcinoma therapy.

We aim at a similar effect in the topical treatment of skin pre-cancerous and cancerous lesions. Actinic keratoses, rough, scaly lesions that commonly occur on sun-exposed areas of the skin, are carcinomas *in situ*, which can progress to squamous cell carcinomas (SCCs)⁵. Organ transplant recipients have particularly high risk rates exceeding those in the immunocompetent population about 100-fold⁶. Currently, actinic keratosis is treated with topically applied 5-fluorouracil (5-FU), photodynamic therapy, a combination of diclofenac and

hyaluronic acid, or with the immune response modifier imiquimod. Typical problems in the therapy with cytotoxic agents are severe local irritations, pain, and secondary infections⁷. Pol α inhibitors may offer an alternative for the treatment of actinic keratoses as well as SCC in particular in high-risk populations. They are supposed to increase cure rates and allow topical application to larger ultraviolet-exposed skin areas.

Based on our 3D model of the human pol α active site and the knowledge of the binding mechanism of the natural substrate 2'-deoxythymidine-5'-triphosphate (dTTP), we recently studied and described the new pol α inhibitor 2-butylanilino-dATP-OH (BuP-OH)⁸ (for chemical structure see Table 1). It is a derivative of the well-known potent and highly selective nucleoside inhibitor of mammalian pol α BuP (see Table 1). As molecular dynamics (MD) simulations of BuP-TP suggested the possibility of an additional H-bond formation with Tyr865, we introduced a hydroxyl function in ortho position to the butyl-

Table 1. Chemical structures of the novel BuP derivatives.

					
Ligand	M (g/mol)	logP	R ₁	R ₂	R ₃
BuP	398,45		H		
BuP-OH	414,45	2,63	OH		
CpBu	492,46	2,08	OH		
CpPe	506,49	2,57	OH		
CpHex	520,51	3,06	OH		
CpIsohex	520,51	2,99	OH		
OxBu	436,39	1,36	OH		
OxPe	450,42	1,85	OH		
OxHex	464,45	2,34	OH		
OxIsohex	464,45	2,27	OH		

BuP, 2-butylanilino-dATP; M, molecular weight; P, phosphonate group.

moiety of the butylanilino group of BuP. The new ligand BuP-OH-TP indeed showed the formation of a stable H-bond between this OH-function and Tyr865 in the course of MD simulations. Importantly, BuP-OH is proven to be only slightly less active as compared to the standard inhibitor of human pol α , aphidicolin⁸.

As we intend to develop new, topically applicable pol α inhibitors, it is of utmost importance that the new drugs offer sufficient skin penetration which asks for a molecular mass under 500 g/mol and a moderate hydrophilicity (logP 1–3)⁹. As mentioned before, pol α inhibitors are activated *via* stepwise phosphorylation of the applied prodrug. This activation procedure is facilitated, if the applied drug is already containing one phosphate function. Monophosphates, however, are not stable and do not easily penetrate into the skin. Phosphonate derivatives, containing a metabolically stable P-C bond, are less polar and therefore show better penetration rates than the corresponding phosphates.

In order to observe the behaviour of phosphonates as pol α ligands, we recently investigated the binding mechanism of phosphonomethoxyethyl derivatives, such as adefovir (9-[2-(phosphonomethoxy)-ethyl]adenine), in the pol α binding site. These agents belong to a class of acyclic nucleotide analogues with P-C bond. The active adefovir diphosphates (adefovir-DP) fitted well into our model despite lack of the characteristic ribose initiated H-bonds with the protein. In fact, adefovir seems to be a good substrate due to its conformational flexibility and good positional adaptability of the acyclic side chain².

In the context of this paper, we present eight novel derivatives of BuP-OH, forming two congeneric series of compounds with a more or less extended side chain (for chemical structures see Table 1). MD simulations—including the post-processing of relevant average structures by calculation of approximate binding energies—and flexible docking studies have been performed in order to investigate the behaviour of the ligands within

the receptor site of our partial pol α homology model. We studied their binding geometries, and we tried to set up a rank order of the new putative pol α inhibitors according to their relative binding affinities to the target. Finally, the theoretical results from our molecular modelling studies were supported by cytotoxic activities that we obtained by pharmacological testing in normal and transformed keratinocytes.

Computational and experimental methods

Construction of the pol α active site model

The design of the 3D homology model of the active site of the human pol α has been described earlier² and will not be discussed in detail here. Briefly, the model was built on basis of the crystallographic RB69 pol α coordinates in the replicating mode (PDB accession code 1IG9) and alignment of the amino acid sequences of the human, RB69, and T9N7 pol α , using the programs CLUSTALW¹⁰, BLAST¹¹, PSIPRED¹², and SYBYL 7.2 (Tripos, Inc. St. Louis, MO). Validation of the model was performed by PROCHECK¹³. Coordinates for primer/template DNA, the incoming dTTP as well as three Ca²⁺-ions were taken from the RB69 crystal structure. The resulting 3D model of human pol α is restricted to the region 836–1102 that contains the important features for the substrate binding event (fingers, palm, and thumb regions).

Design of BuP-OH derivatives as potential pol α inhibitors

As illustrated in Figure 1, we observed the existence of a spacious hydrophobic cavity next to the butyl chain of BuP-OH-TP in the binding site of the protein. The cavity is surrounded by lipophilic amino acids: Leu960, Leu972, Ala973, Val976, and Ile869. This urged us to design new, putatively more potent analogues that include a pentyl, a hexyl, or an isohexyl function instead of the butyl chain. Variation of the deoxyribose-moiety led to a group of derivatives with

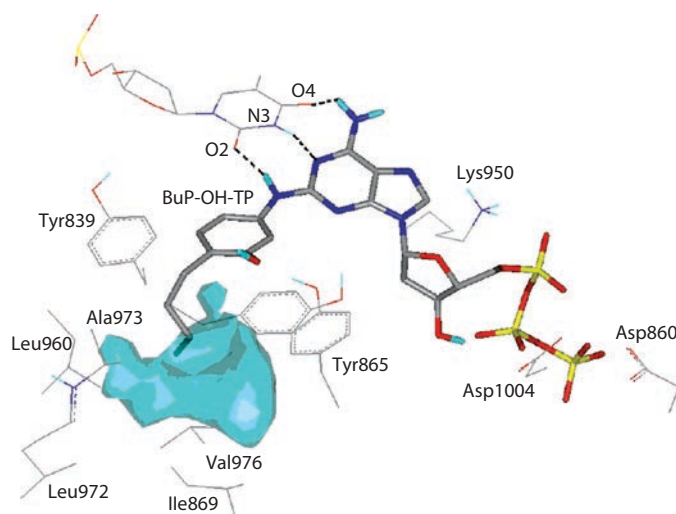


Figure 1. Snapshot of BuP-OH-TP into the active site during MD simulations (hydrophobic cavity marked cyan). Only the strong interactions are shown here. 83x63mm (300 × 300 DPI).

phosphonate function: CpBu, CpPe, CpHex, and CpIsohex (for chemical structures see Table 1). Moreover, in analogy to recent insight into the putative interaction mechanism of phosphonomethoxyethyl derivatives with the binding site² we also studied a series of acyclic BuP-OH derivatives: OxBu, OxPe, OxHex, and OxIsohex (Table 1).

MD simulations

The 3D model of the active site of human pol α served for further investigation by applying MD simulations on the protein-inhibitor complexes. All eight novel inhibitor molecules studied were treated in form of the completely phosphorylated DP of the respective nucleoside phosphonate analogues. The reference compound is a TP of BuP-OH. The inhibitor molecules were inserted manually into the active site by exchange of dTTP. The opposite base of the inhibitor in the DNA strand was exchanged for a thymine instead of the original adenine in order to get a correct Watson-Crick base pairing. The following procedure in order to prepare the interaction complexes for the MD simulations, including the equilibration of the systems and the production simulations, have been carried out with the GROMACS simulation package¹⁴ version 3.3. The interaction complexes were solvated in a cubic box filled with simple point charge (SPC216) water molecules. Subsequently, the whole system was undergoing energy minimization by a steepest descent algorithm. During MD simulations Linear Constraint Solver constraints were used on all protein covalent bonds to maintain constant bond length. A time step of 2 fs was selected. Electrostatic interactions have been calculated with the Particle-Mesh Ewald method and 9 Å distance for the Coulomb cut-off. For calculation of the van der Waals interactions (distance for Lennard-Jones cut-off = 9 Å) twin range cut offs have been used. Temperature (300 K) and pressure (1 bar) were kept constant by applying a Berendsen-thermostat/pressure-coupling. The GROMACS united atom force field and periodic boundary conditions have been applied.

During 1 ps the water molecules have been allowed to equilibrate with positional restraints ($1000 \text{ kJmol}^{-1}\text{\AA}^{-1}$) on all the other atoms. The following simulation was in total 4 ns long. For the whole 4 ns the simulation was performed with positional restraints of $1000 \text{ kJmol}^{-1}\text{\AA}^{-1}$ on the primer- and template-DNA, as well as on the three Ca^{2+} -ions. Ligand and backbone have been restraint during the first 500 ps ($1000 \text{ kJmol}^{-1}\text{\AA}^{-1}$), and subsequently released ($600 \text{ kJmol}^{-1}\text{\AA}^{-1}$ during ps 500–1000; $300 \text{ kJmol}^{-1}\text{\AA}^{-1}$ during ps 1000–1500; no positional restraints from ps 1500–4000).

System coordinates, velocities, and forces were saved every 10 ps for further analysis. Six days of CPU time on a single cluster node with four AMD Opteron Typ 875 processors (2.2 GHz) were necessary to carry out each simulation. Trajectories have been visualized by the GROMACS trajectory viewer ngmx. 3D structures were visually inspected in SYBYL 7.2. We examined the total potential energy of the respective system and the protein backbone root mean square deviations (RMSD) with re-

spect to the solvated, steepest descents minimized structure (after water-equilibration) to see if conformational stability has been achieved during the equilibration. For further examination we calculated short-range (SR) Coulomb and Lennard-Jones (LJ) interaction potentials between the respective ligands and the protein in the course of the simulation. Once equilibrated, we studied the H-bond formations between the respective ligands and the protein/DNA by using the *g_hbond* routine. A hydrogen bond was defined by a maximum donor-acceptor distance of 3 Å and an acceptor-donor-hydrogen angle of 60° or less. In order to give an idea about H-bond occurrence and stability, their lifetime is reported as percentage value of the time period analyzed. Average structures are extracted from the MD trajectories after equilibration, energy minimized and serve for further inspection of the ligand binding modes and affinities. (All analysis tools applied are implemented into the GROMACS software.)

Post-processing of average structures from MD simulations

By means of the *g_cluster* routine (implemented into GROMACS) 10 average structures for each run were extracted from the last 1000 ps (equilibrated region), respectively. These 10 structures represent the centroids of 10 clusters (on basis of protein backbone RMSD values), respectively, resulting in a set of 10 maximum diverse geometries for each protein-ligand complex. Respective pdb files were first pre-processed with OPEN BABEL¹⁵, and then imported into the graphical user interface MAESTRO version 8.5 (SCHRÖDINGER, LLC, New York, NY) of the SCHRÖDINGER software package where the structures again underwent some preparation steps in order to create a readable format. Subsequently, the PRIME version 2.0 (SCHRÖDINGER, LLC, New York, NY) energy minimization tool was used to convert the MD snapshots into more refined structures. Accordingly, these geometries were subject to further investigation by applying the X-Score program¹⁶: the HP-, HM-, HS-Score, and the resulting consensus score X-CCore, as well as an approximation of the predicted binding energies (in kcal/mol) were calculated for each of the snapshots of the receptor-ligand complexes. The mean values of the 10 snapshots for each of the studied ligands should give some first information about their discriminative interaction potentials with the target.

Docking into the active site

In addition to MD simulations flexible docking studies were performed with SURFLEX-Dock¹⁷ in SYBYL 7.2. The original protein homology model, which served as a starting point for MD simulations, did not serve for docking, because ligands were placed partially mirror-inverted into the binding site with the lipophilic alkyl chain to the opposite site compared to the pose of dTTP. We concluded that these artificial poses occur due to insufficient free space in the dTTP protein. Consequently, the following docking studies were performed into the 'wider' binding site of the

protein after MD simulations with CpIsohex-DP, which is the bulkiest of the ligands under consideration. One average structure was extracted from the last 1000 ps of the trajectory (equilibrated region, 3001–4000 ps) of the protein-CpIsohex-DP complex by applying the *g_cluster* routine which is implemented into the GROMACS software package. The selection of the respective MD frame to serve as an average structure is based on protein backbone RMSD values and represents the ‘average’ of all the structures that can be found in the last 1000 ps of the trajectory. In other words, the selected average structure is the centroid of all the structures that appear within this period if all the structures are taken as one unique cluster. This average structure, which appears at ps 3060, was minimized by steepest descent, and then served as a target for docking. Protomol generation was performed ligand-based—the minimized CpIsohex-DP ligand from the average structure at 3060 ps was used to identify the protein active site. Default settings have been used.

Chemicals and solutions

Aphidicolin was purchased from Serva Electrophoresis (Heidelberg, Germany), 5-FU and 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazoliumbromid (MTT) from Sigma-Aldrich (Taufkirchen, Germany), BuP-OH⁸ and the analogues were synthesized at Freie Universität Berlin, Department of Chemistry (Berlin, Germany). OxBu and OxHex were synthesized at Chiracon GmbH (Luckenwalde, Germany). Synthesis will be published elsewhere. Stock solutions ($c = 2 \times 10^{-2}$ M) were prepared by dissolving 5-FU in phosphate buffered saline and aphidicolin, BuP-OH, OxBu, OxIsohex, and OxHex in dimethylsulfoxide (DMSO), respectively. In cell culture, maximum concentration of DMSO was 0.5% and cells grown in the presence of the respective solvents served as control.

Cell culture

Normal human keratinocytes (NHK) were isolated from foreskin as described before¹⁸, NHK from three donors were pooled and cultured in keratinocyte growth medium that was prepared from keratinocyte basal medium (Cambrex, Landen, Belgium) by the addition of 0.1 ng/mL epidermal growth factor, 5.0 µg/mL insulin, 0.5 µg/mL hydrocortisone, 30 µg/mL bovine pituitary extract, 50 µg/mL gentamicin sulphate, and 50 ng/mL amphotericin B (Sigma-Aldrich). For all experiments cells of the second to the fourth passage were used. Medium for SCC-25 cells (ATCC # CRL-1628) was prepared from Ham's F12 (Seromed Biochrom, Berlin, Germany) and Dulbecco's modified Eagle's Medium (Sigma-Aldrich; 1:1) by addition of 15% foetal calf serum (Seromed Biochrom, Berlin, Germany), 0.4 µg/mL hydrocortisone, 100 U/mL penicillin and 100 µg/mL streptomycin (Sigma-Aldrich).

Cytotoxicity assay

Cytotoxicity was quantified using the standard MTT dye reduction assay¹⁹. Cells (4×10^4 cells/well) were grown in

growth medium overnight (24 well plates). Medium was replaced by fresh growth medium for keratinocytes or foetal calf serum free medium for SCC-25, respectively. Then the indicated agents were added for 48 h and MTT solution (final concentration of 0.5 mg/mL) for the last 4 h. After removing supernatants and solubilization of formazan crystals in DMSO, optical density was determined at 540 nm (Fluostar Optima, BMG Labtech, Offenburg, Germany) as described²⁰.

Results and discussion

MD simulations

After total release of the protein backbone and the respective ligand (ps 1500–4000) the whole system rapidly reached its equilibrium after ps 2000 (for OxPe-DP) and ps 3000 (for all the other ligands). This was evaluated by the RMSD curves of the protein backbones and the potential energy curves of the whole system, respectively. For the MD simulation with OxPe-DP we consider the period from 2001 to 3000 ps as the equilibrated region, meaning that the backbone RMSD curve as well as the potential energy curve have reached a plateau. For all the other ligands the period from 3001 to 4000 ps during MD simulations offers those equilibrated conditions. Thus, the examination of average structures, mean values of Coulomb and Lennard-Jones interaction energies and the occurrence of hydrogen bonds during this period should give us new insights into the putative binding mode(s).

Ligands with a sugar-moiety: diphosphates of CpBu, CpPe, CpHex, and CpIsohex

Compared to the natural substrate dTTP, BuP-OH-TP (which served as a reference molecule) and its deoxyribose-containing derivatives (CpBu-, CpPe-, CpHex-, and CpIsohex-DP) show quite similar binding modes. In particular, the formation of a stable hydrogen bond between the amide hydrogen of Tyr865 and the oxygen of the 3'-OH-group of the deoxyribose seems to be a common characteristic. Figure 2 exemplifies the formation of this particular H-bond in the case of CpHex-DP (at 3980 ps). We measured the occurrence of this H-bond every 10 ps during the last 1000 ps of the simulation. In the case of BuP-OH-TP it was formed for 85% of the time period analyzed; for the other deoxyribose-containing derivatives

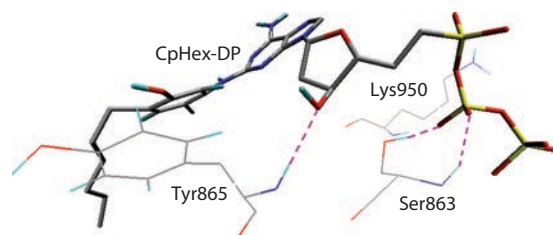


Figure 2. CpHex-Dp within the active site during MD simulations (at 3980 ps): H-bond formation with Tyr865 (plus Ser863). 80x39mm (300×300 DPI).

it appeared even more often (CpBu-DP: 91%, CpPe-DP: 97%, CpHex-DP: 99%, CpIsohex-DP: 88%). Yet, the expected H-bond formation between the OH-function on the phenyl ring of the alkylphenyl-moiety of the ligand and Tyr865 did not appear more often than in 1% of this time period. The adenine ring is stabilized *via* H-bond formation with the opposite base (thymine) in the DNA strand (see Figure 1). On one hand, the amide nitrogen N₃ of the thymine acts as a donor to form H-bonds with the cyclic and acyclic amino counterparts of adenine and the annexed amino group (which itself carries the alkylphenyl-moiety), respectively. On the other hand, the same acyclic amino groups of the ligand are H-bond donors for the opposite thymine N₃ and its vicinal carbonyl oxygens O₂ and O₄. In addition, the adenine base is involved into a weak interaction with Asn954. At the diphosphosphonate terminus, H-bond formation with the conserved residues Ser863, Arg922, Lys950, and sometimes additionally with Leu 864 is of utmost importance. These residues are able to build strong interactions with the respective ligand and thus are capable to additionally lock the ligand into its position within the binding pocket. All of the deoxyribose-containing BuP-OH derivatives show a very stable H-bond formation—with an occurrence of 100%—with Ser863 within the examined period. What differs—regarding the distinct derivatives—is especially the tendency to form H-bonds with Arg922 and Lys950. Most remarkably CpHex-DP (H-bond occurrence with Arg922: 64%, with Lys950: 73%; see Figure 3: CpHex-DP at 3720 ps) and also CpBu-DP (H-bond occurrence with Arg922: 77%, with Lys950: 48%) have a higher frequency of H-bond formation with these two amino acids than the other deoxyribose-containing ligands (Arg922: 6–36%, Lys950: 0–29%).

Acyclic ligands: diphosphates of OxBu, OxPe, OxHex, and OxIsohex

In contrast to the sugar-containing derivatives, the acyclic phosphonates are not suited to build an H-bond with the

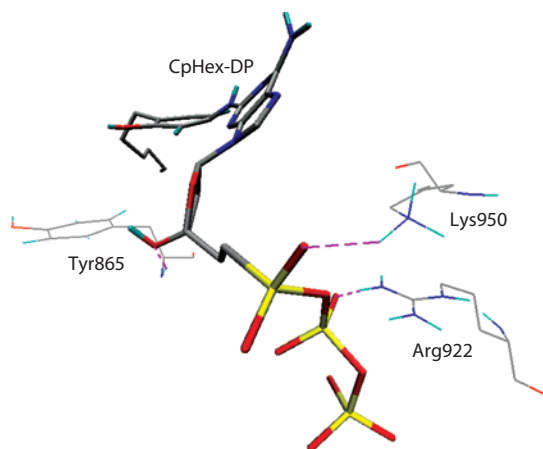


Figure 3. CpHex-DP within the active site during MD simulations (at 3720 ps): H-bond formation with Arg922 and Lys950 (plus Tyr865). 74x58mm (300 × 300 DPI).

amide hydrogen of Tyr865. Instead we could detect the occurrence of an H-bond between the OH-function on the phenyl ring of the alkylphenyl-moiety and Tyr865 for the derivatives OxBu-DP and OxIsohex-DP (see Figure 4: OxIsohex-DP at 3370 ps) in the observed period (3001–4000 ps). The adenine base is engaged in identical interactions as found for the deoxyribose-containing series: H-bond formation with the opposite thymine and additionally with Asn954. Regarding the diphosphosphonate part of the molecule(s), the same amino acids serve for binding interactions as for the cyclic derivatives: Ser863, Leu864, Arg922, and Lys950. Also here, we noticed that there is a very stable Ser863 interaction with all the ligands (occurrence: 98–100%). As before, the appearance of H-bonds with Arg922 and Lys950 seem to vary strongly among the acyclic ligands: OxIsohex-DP (H-bond occurrence with Arg922: 79%, with Lys950: 78%) and OxBu-DP (H-bond occurrence with Arg922: 84%, with Lys950: 76%) show frequent interaction. The other acyclic derivatives show only poor occurrence of H-bonds with Arg922 within the equilibrated period (36% for OxPe-DP and 22% for OxHe-DP respectively). No H-bonds are built with Lys950 in the cases of OxPe-DP and OxHex-DP.

Obviously, BuP-OH derivatives are likely to interact *via* H-bonds with Ser863 on one hand, but also with Arg922 and Lys950 on the other hand, and these interactions seem to be somehow competitive. Only if a molecule is able to optimally adapt its conformation due to ideal flexibility—like in the case of OxIsohex-DP (see Figure 4: OxIsohex-DP at 3370 ps)—it is possible to build H-bonds with all of the mentioned residues at the same time. But this is not a very common scenario—also for OxIsohex-DP. Within the observed period (3001–4000 ps), predominantly H-bond formation with Ser863 competes against those with Arg922 and Lys950. We detected this kind of hydrogen bond competition for all the tested ligands. Especially, the ‘smaller’ molecules OxBu-DP and CpBu-DP tend to oscillate easily between the distinct possible H-bond forms. This is another reason, why the ‘larger’ molecules with e.g. pentyl or hexyl function should be better candidates for pol α inhibitors—besides the fact that greater lipophilic-

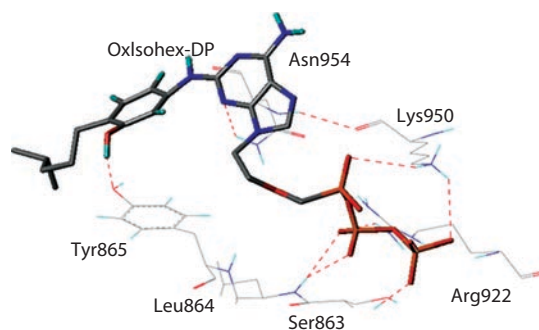


Figure 4. OxIsohex-DP within the active site during MD simulations (at 3370 ps): H-bonds are simultaneously built with Ser863, Arg922 and Lys950 (plus Tyr865 and Asn954). 54x33mm (300 × 300 DPI).

ity of the side chain should enhance binding affinity. These compounds are more likely to adapt rapidly into the binding pocket, with very stable interaction energy levels during the last 1000 ps of the MD simulation. We observed large fluctuation of interaction energies during the mentioned period for the butyl and isohexyl derivatives, but rather small fluctuations for pentyl and hexyl derivatives (data not shown). This effect is even more distinctive when we examine the acyclic derivatives. Thus, the diphosphates of OxPe and particularly OxHex may be good candidates, although they overall lack H-bond formation with Lys950.

The large fluctuations of energy levels regarding the isohexyl compounds are not easy to interpret. Most likely, the bulky side chain acts as an obstacle that represses the ligand from finding its ideal conformation, so that as a consequence the ligand is hopping from one local energy minimum into another.

Sugar-containing versus acyclic series

Comparison of the series of sugar-containing with the series of acyclic ligands regarding their hydrogen bond formation and inspection of average structures within the binding pocket confirmed our suggestions regarding comparison of acyclovir with BuP-OH². Acyclic derivatives are more flexible and thus are able to adopt more easily at the diphosphosphonate terminus. This property seems to counterbalance the fact that they lack H-bond formation of the sugar-3'-OH with Tyr865. Additionally, we rarely observed an intramolecular hydrogen bond that was formed by this 3'-OH with the β -phosphate oxygens. Intramolecular stabilization of the diphosphosphonate terminus of deoxyribose-containing derivatives presumably leads to even greater rigidity of these series of ligands.

Interaction energies

Inspection of the mean SR Coulomb and Lennard-Jones interaction energies (calculated by the *g_energy* routine of GROMACS) should not lead to a definite rank order for classification of the compounds into putatively stronger and weaker pol α inhibitors at this point of investigation. What we learned from the ranking of the ligands according to their mean Coulomb and LJ interaction potentials during the last 1000 ps of the simulation, respectively (see Table 2), is, that almost all of the novel ligands seem to have at least the same potential to inhibit the target as the reference compound BuP-OH-TP. Regarding the

SR-Lennard-Jones interaction energies we can clearly detect the influence of the size of the lipophilic side chain. The hexyl- (CpHex-DP: -272 kJ/mol; OxHex-DP: -276 kJ/mol) and isohexyl (CpIsohex-DP: -285 kJ/mol; OxIsohex-DP: -264 kJ/mol) derivatives of the sugar-containing as well as of the acyclic series, show stronger interaction energies than those of the respective pentyl- (CpPe-DP: -267 kJ/mol; OxPe-DP: -237 kJ/mol) and butyl- (CpBu-DP: -250 kJ/mol; OxBu-DP: -202 kJ/mol) derivatives (see Table 2). Most remarkably, all of the ligands have stronger Coulomb-interaction energies than BuP-OH-TP (-19 kJ/mol, Table 2).

Post-processing by X-Score

Applying the empirical scoring functions implemented into the X-Score program on representative snapshots from the MD production runs should provide rough binding strength estimations. Three independent scoring functions (the HP-, HM-, and HS-Score)—which differ in the way how the hydrophobic effect is modelled—are combined into a consensus scoring function (X-CCScore) and finally these scores are translated into estimated binding energies. It should be stated clearly, that the average accuracy of this consensus scoring algorithm in calculating the absolute binding free energies is reported to be approximately 2.2 kcal/mol (tested on an independent set of 30 protein-ligand complexes)¹⁶. Besides the term for the hydrophobic effect, the three scoring functions also include algorithms which account for the van der Waals interactions, hydrogen bonding, and deformation penalties. Additionally, the regression constant reflects the effects due to the translational and rotational entropy loss in the binding process¹⁶.

Table 3 lists the mean scoring and binding free energy values for each ligand-protein complex. Each of the applied scoring functions is able to reflect the influence of the size of the hydrophobic side chain. However, differences of the predicted binding energies are insignificant considering the reported probability of error of the program (~2.2 kcal/mol).

Deformation effects of the ligand and the protein during the binding process lead to entropic changes. This effect is even more distinctive, the more flexible the ligand is. In our special case the scoring function might lead to an overestimation of the entropy loss in the case of the acyclic series of ligands. This leads us to the conclusion that the results proposed by X-Score have to be interpreted with caution, and that finally the two series

Table 2. Mean SR-Coulomb (SR-Coul) and SR-Lennard-Jones (SR-LJ) interaction potentials (kJ/mol) and the respective standard deviations (SD) between the respective ligands and the protein during the last 1000 ps of the simulation (equilibrated region).

	SR-Coul (SD)	SR-LJ (SD)	Sum (SD)		SR-Coul (SD)	SR-LJ (SD)	Sum (SD)
OxIsohex-DP	-77 (26)	-264 (18)	-341 (22)	CpIsohex-DP	-55 (26)	-285 (15)	-340 (28)
OxHex-DP	-27 (18)	-276 (12)	-303 (17)	CpHex-DP	-80 (27)	-272 (14)	-352 (26)
OxPe-DP	-27 (19)	-237 (14)	-264 (19)	CpPe-DP	-36 (19)	-267 (15)	-304 (20)
OxBu-DP	-84 (33)	-202 (14)	-286 (33)	CpBu-DP	-94 (27)	-250 (16)	-343 (30)
				BUP-OH-TP	-19 (31)	-270 (12)	-290 (35)

BuP, 2-butylanilino-dATP; DP, diphosphates; SR, short-range; TP, triphosphate.

can be taken as equivalent concerning their approximate binding energies.

Moreover, it is very likely that the enhancement of the drug in the binding site as well as the building of the protein-ligand complex occurs more difficult/slowly for the cyclic than for the acyclic derivatives. This cannot be considered by a scoring function, but is a determining factor for the bioavailability of a certain drug. Thus, adefovir-like derivatives are believed to offer certain benefits that probably turn them into slightly better candidates for pol α inhibition compared to the sugar-containing series.

Flexible docking studies

Docking with the flexible algorithm of SURFLEX-Dock, should provide us with additional information regarding the ideal binding poses of the ligands. Complementary to the results obtained from MD simulations and the scoring of representative structures with X-Score, independent docking studies should allow us to set up a final rank order of the ligands according to their binding affinities represented by the total score. SURFLEX-Dock uses the so called 'protomol'¹⁷—a computational representation of the intended active site (an absolute docking envelop)—to which then the putative ligands are aligned on basis of maximum similarity. In the course of protomol generation, molecular fragments ('probes') are put into the protein binding site in multiple positions. In our case, ligand-based protomol generation means that the steepest descent minimized structure of CpIsohex-DP at ps 3060 (average structure) was used for the identification of the binding site.

Diverse types of probes are used for protomol generation: a steric, hydrophobic probe (CH_4), a hydrogen bond donor probe (N-H) and a hydrogen bond acceptor probe (C=O). Like X-Score, SURFLEX-Dock also uses an empirically derived scoring function, but it is based on atom/atom pairwise interactions. There are hydrophobic, polar, repulsive, entropic, and solvation terms included in its calculation²¹. By combining the results from flexible docking studies with those from MD simulations

and post-processing by X-Score, we are able to detect the influence of the hydrophobic side chain more clearly and we could finally differentiate between the two congeneric series of ligands in terms of binding affinity to the target protein.

As already described, docking was performed into the protein after MD simulation with the bulkiest ligand CpIsohex-DP (for details see Methods section). We also repeated to whole docking procedure with the protein from the simulation with OxIsohex-DP, in order to avoid a preferential treatment of the deoxyribose-containing series of ligands, but we did not succeed with this template. Some of the ligands did not adapt correctly with their lipophilic chain into the hydrophobic pocket, but were also placed kind of mirror-inverted into the active site—like in the case of the use of the original protein homology model as a template. Moreover, the influence of the selected ligand as a basis for protomol creation should be marginal, as docked ligands are scored in the context of the receptor, and not in the context of the protomol.

From the 10 best docking poses for each molecule we chose the one with the highest total score and most reliable conformation due to our knowledge. This was the very first (and also highest ranked) pose for all ligands, except for BuP-OH-TP and CpBu-DP. For BuP-OH-TP the first three highest ranked poses as well as the highest ranked pose for CpBu-DP placed the molecules correctly into the binding pocket regarding the ring system, but the TP and diphosphosphonate termini, respectively, protrude from the active site of the homology model. Thus, we selected the fourth pose of BuP-OH-TP and the second pose of CpBu-DP in order to be able to discriminate the ligands in terms of their scores based on comparable binding geometries. Figure 5 displays the superimposed docking poses of all our eight novel putative pol α inhibitors plus our reference compound BuP-OH-TP. As demonstrated, the ring systems of the ligands were placed onto each other during the docking process with quite low deviations. In contrast, the flexible

Table 3. Mean values for the predicted binding energy, the HP-, HM-, and the HS-Score, as well as the X-CCScore for each protein-ligand complex.

	Predicted binding energy	X-Score	HP-Score	HM-Score	HS-Score
OxIsohex	-11.31	8.29	8.35	8.42	8.10
OxHex	-11.10	8.14	8.27	8.23	7.93
OxPe	-10.66	7.82	7.92	7.92	7.62
OxBu	-10.56	7.74	7.76	7.88	7.59
CpIsohex	-12.00	8.80	8.82	8.94	8.73
CpHex	-11.88	8.71	8.70	8.86	8.58
CpPe	-11.83	8.67	8.73	8.81	8.48
CpBu	-11.23	8.23	8.16	8.36	8.17
BUP-OH	-11.15	8.18	8.16	8.30	8.07

Predicted binding energies are displayed in kcal/mol, all scores in units of $-\log(K_d)$.

BuP, 2-butylanilino-dATP.

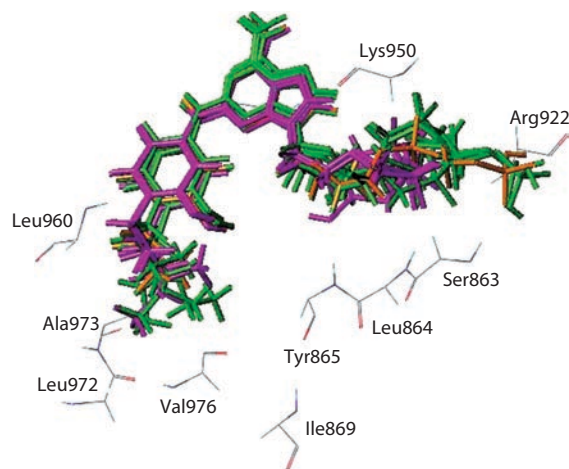


Figure 5. Docking poses of all ligands superimposed. 60x47mm (300×300 DPI).

lipophilic side chains as well as the TP/diphosphophosphonate termini show more expressed variations in their positions.

Finally, we examined the absolute ranking order of all our ligands—plus BuP-OH-TP as a reference—due to the docking scores, calculated by the docking algorithm (see Table 4). According to the total scores, the acyclic derivatives are clearly favoured over the sugar-containing compounds. In addition, the more lipophilic hexyl and isohexyl derivatives were ranked higher than those with a pentyl-moiety, followed by the butyl compounds, which took the lowest ranking positions, respectively. Moreover, all of our investigated compounds—except for CpBu-DP—got higher scores than the reference BuP-OH-TP. Thus, by performing flexible docking studies, we were able to support the results obtained by MD simulations and could set up a final rank order of the novel putative inhibitors due to their binding affinities with the pol α active site.

Pharmacological activity

Since the hydrophilic diphosphates of the eight novel compounds with a phosphonate function would not be taken up into the cells, cytotoxicity was investigated with the nucleoside phosphonate analogues themselves. This is in contrast to molecular modelling which is based on the completely phosphorylated derivatives (diphosphates). The phosphorylation of the substances is done intracellularly by cellular kinases which are expressed both by NHK and SCC-25 cells. In fact, first approaches to derive potential pol α inhibitors by molecular modelling allowed us to identify non-physiological purine (BuP-OH) and pyrimidine bases which proved cytotoxic though at rather high concentrations⁸. To investigate the potency of nucleoside phosphonate analogues to inhibit pol α , we tested OxBu, OxIsohex, and OxHex on NHK and SCC-25 cells. Activity was compared to the cytotoxicity of BuP-OH and 5-FU, which is the current standard for topical therapy of actinic keratosis and to the efficacy of the

highly cytotoxic aphidicolin. Activity of these substances is proven experimentally, the MTT reduction assay appeared discriminative⁸. Due to the fact that only viable and metabolically active cells will reduce MTT to the blue formazan, the amount of dye is directly related to the cell number.

The maximum inhibitory effect of aphidicolin on the viability of NHK and SCC-25 cells was even superseded by BuP-OH. Importantly, BuP-OH affected NHK and SCC-25 cells in the same concentration range whereas aphidicolin was most toxic to normal keratinocytes. Yet, since the active concentrations of BuP-OH were almost four orders of magnitude higher than with aphidicolin (Table 5) the agent appears not suitable for clinical development. In contrast, 5-FU was slightly selective for SCC-25 cells. The aim of superseding 5-FU effects, however, could be reached with BuP-OH only in part. The maximum cytotoxicity of BuP-OH clearly surmounted maximum 5-FU effects whereas the latter agent was toxic in lower concentrations.

Next we decided to synthesize and determine the cytotoxic activity of the phosphonates most efficiently interacting with the target protein as derived from the molecular modelling studies. These are three congeners of the acyclic series: OxBu, OxHex, and OxIsohex. Their synthesis will be described separately. Most importantly, the phosphonates of acyclic nucleotide analogues clearly differentiated between NHK and SCC-25 cells. With OxHex the toxic concentrations in NHK exceeded those in SCC-25 cells more than 10-fold, and with OxBu as well as with OxIsohex no damage of NHK at all was detected. Referring to the IC_{50} value, OxBu and OxHex are 1000-fold more toxic for SCC-25 cells than 5-FU and aphidicolin. OxIsohex reduced viability of SCC-25 cells by about 20% at maximum only, whereas OxBu reached with 39% maximum inhibition almost the effect of 5-FU. Although OxHex affects also NHK, the maximum inhibition is clearly stronger in SCC-25 cells, which adds to the much lower toxic concentrations SCC-25 cells (Table 5).

According to the modelling results, OxIsohex should be a very active congener of the nucleoside phosphonates, too. This was not seen in the *in vitro* experiments and should be due to hindered access of the non-linear side chain of OxIsohex to the active site of the enzyme. Currently the molecular mechanism of cell specific toxicity remains open since major differences in the expression of enzymes of nucleoside metabolism and nucleoside transporters have not been detected⁸. In addition, binding experiments to the isolated DNA

Table 4. Docking scores of the novel potential pol α inhibitors.

Ligand	Total score	Ligand	Total score
OxIsohex-DP	15.02	CpIsohex-DP	12.80
OxHex-DP	15.80	CpHex-DP	11.64
OxPe-DP	14.50	CpPe-DP	10.97
OxBu-DP	14.30	CpBu-DP	9.98
		BuP-OH-TP	10.70

BuP, 2-butylanilino-dATP; DP, diphosphates; pol α , polymerase α ; TP, triphosphate.

Table 5. Cytotoxicity of polymerase α inhibitors in normal human keratinocytes (NHK) and SCC-25 cells.

MTT test 48 h	Aphidicolin	5-FU	BuP-OH	OxBu	OxHex
NHK	8,26 \pm 0,41 (44%)	5,08 \pm 0,48 (41%)	4,58 \pm 0,24 (66%)	n.d.	n.d.
SCC-25	6,29 \pm 0,29 (62%)	6,01 \pm 0,28 (42%)	5,05 \pm 0,3 (65%)	9,43 \pm 0,22 (39%)	4,93 \pm 0,26 (20%)

–log IC_{50} values (M) \pm standard error and maximum inhibition of the dye reduction were derived from viability (MTT-reduction) assay.

Three independent experiments were performed in triplicate. n.d.: log IC_{50} is not determinable with maximum inhibition < 20%.

BuP, 2-butylanilino-dATP; 5-FU, 5-fluorouracil; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazoliumbromid; SCC, squamous cell carcinoma.

polymerase isoforms will be helpful to unravel, if the differences in activity result from improved receptor binding or improved cellular uptake—although the latter appears clearly less plausible with the limited structural differences of OxBu and OxHex (Table 1). Binding experiments on the various isoforms of DNA polymerases will also unravel, if the selectivity of OxBu and OxHex for the cancer cell is due to a preferential inhibition of an isoform predominantly expressed in SCC-25 cells. Yet, we have to postpone this investigation because of the need for another highly challenging synthesis step to form the active diphosphate until the final decision for the agent of further development.

Conclusions

MD and docking studies of two congeneric series of BuP-OH-TP derivatives revealed important features which are determining potent nucleoside inhibitors of pol α . These studies allowed us to select candidates for pharmacological studies and apparently the range order of activity in a squamous carcinoma cell line seems to meet our predictions. Congeners of the acyclic series may offer a new approach for skin cancer, if active by the topical application route which is used to control unwanted side effects. The equipotent agents of the sugar series, however, appear less suitable due to slightly minor binding affinities to the target and a molecular weight close to the critical size with respect to skin penetration.

Declaration of interest

Two authors (Hans-Dieter Hölte, Monika Schäfer-Korting) are inventors of the respective patent applied for by RIEMSER Arzneimittel AG (Greifswald – Insel Riems), European Appl. No.: 07090098.0. Moreover, Monika Schäfer-Korting acts as a consultant for RIEMSER Arzneimittel AG (Greifswald – Insel Riems) with respect to the Development of dermatics. The authors alone are responsible for the content and writing of the paper.

References

- Steitz TA. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem* 1999; 274:17395–17398.
- Richartz A, Hölte M, Brandt B, Schäfer-Korting M, Hölte HD. Targeting human DNA polymerase α for the inhibition of keratinocyte proliferation. Part 1. Homology model, active site architecture and ligand binding. *J Enzyme Inhib Med Chem* 2008; 23:94–100.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74:5463–5467.
- Mitsuya H, Yarchoan R, Broder S. Molecular targets for AIDS therapy. *Science* 1990; 249:1533–1544.
- McIntyre WJ, Downs MR, Bedwell SA. Treatment options for actinic keratoses. *Am Fam Physician* 2007; 76:667–671.
- Lindelöf B, Sigurgeirsson B, Gäbel H, Stern RS. Incidence of skin cancer in 5356 patients following organ transplantation. *Br J Dermatol* 2000; 143:513–519.
- Chakrabarty A, Geisse JK. Medical therapies for non-melanoma skin cancer. *Clin Dermatol* 2004; 22:183–188.
- Hölte M, Richartz A, Zdravil B, Schwanke A, Dugovic B, Murruzzu C, Reissig HU, Korting HC, Kleuser B, Hölte HD, Schäfer-Korting M. Human polymerase α inhibitors for skin tumors. Part 2. Modeling, synthesis and influence on normal and transformed keratinocytes of new thymidine and purine derivatives. *J Enzyme Inhib Med Chem* 2010; 25:250–265.
- Korting HC, Schäfer-Korting M. Novel drug delivery systems. *Handbook of Pharmacology*. Heidelberg: Springer, 2009.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; 22:4673–4680.
- Altschul SF, Madden TL, Schäfer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; 25:3389–3402.
- Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 1999; 292:195–202.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993; 26:283–91.
- Lindahl E, Hess B, Van der Spoel D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model* 2001; 7:306–17.
- Guha R, Howard MT, Hutchison GR, Murray-Rust P, Rzepa H, Steinbeck C, Wegner J, Willighagen EL. The Blue Obelisk interoperability in chemical informatics. *J Chem Inf Model* 2006; 46:991–8.
- Wang R, Lai L, Wang S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comput Aided Mol Des* 2002; 16:11–26.
- Jain AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem* 2003; 46:499–511.
- Vogler R, Sauer B, Kim DS, Schäfer-Korting M, Kleuser B. Sphingosine-1-phosphate and its potentially paradoxical effects on critical parameters of cutaneous wound healing. *J Invest Dermatol* 2003; 120:693–700.
- Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* 1983; 65:55–63.
- Gysler A, Lange K, Korting HC, Schäfer-Korting M. Prednicarbate biotransformation in human foreskin keratinocytes and fibroblasts. *Pharm Res* 1997; 14:793–797.
- Jain AN. Scoring non-covalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J Comput Aided Mol Des* 1996; 10: 427–440.