POLSKIE TOWARZYSTWO BIOCHEMICZNE

postępy biochemii

PL ISSN
0032-5422

Advances in Biochemistry

TOM 41, NR 5 supplement, 1995

Special issue dedicated to
Professor David Shugar
on the occasion of his 80th birthday

This edition of a supplement to the quarterly *Postępy Biochemii* was made possible through the financial support of the Polish Biochemical Society and the Committee for Scientific Research (KBN)

postępy biochemii

Editor
POLSKIE TOWARZYSTWO BIOCHEMICZNE
Polish Biochemical Society
ul. L. Pasteura 3
02-093 Warszawa
Poland
(2) 658-20-99
fax: (22) 22-45-52

Drukarnia Naukowo-Techniczna
Mińska 65
03-828 Warszawa

Address
REDAKCJA KWARTALNIKA
„POSTĘPY BIOCHEMII"
INSTYTUT BIOLOGII DOŚWIADCZALNEJ
im. M. Nenckiego PAN
ul. L. Pasteura 3
02-093 Warszawa
tel (2) 659-85-71 ext. 332
fax: (22) 22-53-42

CONTENTS

On behalf of the Polish Biochemical Society the Editorial Board of Postępy Biochemii (Advances in Biochemistry) takes pleasure in dedicating this special issue of the journal to Professor David Shugar on the occasion of his 80th birthday. It contains papers written by his coworkers and friends.

On the same occasion a special Symposium was organized in the Institute of Biochemistry and Biophysics, Polish Academy of Sciences. The program of the Symposium is included in this issue of Postępy Biochemii.

Professor David Shugar
delivering his lecture during the Symposium
Warsaw, September the 2nd, 1995

Profile — David Shugar

David Shugar was born in Poland on the 10th of September, 1915. When he was three years old, the family emigrated to Canada. There he grew up, was educated, and obtained his PhD in physics at McGill University (Montreal) in 1940.

His interest in the biological sciences arose during his PhD research when he worked, for financial support, on the construction of an ultracentrifuge for medical purposes at the University hospital. Until 1948 he worked in the field of physics and biophysics in scientific laboratories in Canada. The first occasion to start biological research opened up in postwar Europe, when he received a fellowship at the Pasteur Institute, and subsequently at the Sorbonne. He worked there on kinetic of enzymes, protein denaturation and photochemistry. These studies were continued at the University of Brussels.

In 1952 he published four papers with J. J. Fox on the UV absorption spectra and tautomerism of nucleic acid derivatives. In an elegant way they identified tautomeric forms of pyrimidines and proved N(1) as the position of attachment of a sugar moiety to the base in nucleosides. These findings became fundamental information on the structure of nucleic acids. The simple method of absorption spectra measurement was used for years in studies on the structure of natural and synthetic polynucleotides and their interactions with various agents.

In the early fifties D. Shugar came to Poland, initially for a one-year stay on the invitation of Professor Leopold Infeld and the Ministry of Higher Education. He was a young researcher with broad knowledge, brilliance of mind and the rare ability of finding interesting and important problems that could

be solved under the available conditions. He held a basic belief that in science, nearly in any circumstances, one can do something useful with hard work and a little creativity. He found here the opportunity to use both.

To start with, he held the post of head of a laboratory in the Biochemistry Department at the State Institute of Hygiene in Warsaw. Later, he headed the Department till 1966 when he left the Institute.

In the Institute of Nuclear Research at Świerk he participated in the setting up of an analytical laboratory for measuring the purity of reactor materials.

From 1954 on, D. Shugar was involved in the organization of the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences. At the beginning he organized a Laboratory of Biological Physical-Chemistry, which grew to become a Department, after some time renamed the Biophysics Department. After some years this was divided into two Departments, with Molecular Biology headed by D. Shugar till his retirement in 1985. It is worth mentioning that, due to his organisational activity and influence on the scientific profile, the Institute was given the second part of its name.

D. Shugar organized in 1955, with the help of coworkers and the Faculty of Physics (University of Warsaw), the first course in Poland on isotope techniques for biologists and medical personnel. A collection of the lectures in book form was very helpful for a long time. At the State Institute of Hygiene he designed, Poland's first, cobalt-60 source for research in radiation chemistry and biology, and supervised its construction.

In 1966, at the request of the Ministry of Higher Education and the Faculty of Physics, University of Warsaw, D. Shugar began to organize a Chair of Biophysics, currently the Department of Biophysics, at the Faculty. He was the principal author of the programs for teaching and research, and the head of the Unit for many years. It is the only biophysics department in Poland in a physics faculty; others were organized later usually as parts of biology departments. Over a hundred physicists have already specialized in molecular biophysics in the Department.

D. Shugar served also for some time as a consultant in the Department of Radiobiology in the Institute of Oncology in Warsaw. Due to his cooperation and inspiration there is now a modern laboratory of molecular medicinal biology.

These huge organisational efforts ran simultaneously with experimental studies. D. Shugar started in Poland with a small group of people varying in educational background, shills and ability and age, having available only poor equipment and very limited finances. However, he managed in the first years to perform valuable studies on ribonuclease and lysozyme establishing, for the latter, a unit of enzymatic activity, called the Shugar unit. Research was also conducted on the mechanism of Gram staining in microorganisms and on histochemical localization of enzymes. The latter problem became a flourishing field of investigation in D. Shugar's laboratories, involving synthesis of specific labeled substrates and studies on the properties and function of nucleolytic enzymes.

In 1957 he started research on the photochemical transformation of pyrimidines and polynucleotides in relation to mutagenesis and lethal effects, and after some time in relation to DNA repair as well. The synthesis of polynucleotides was introduced for studies on the structure of nucleic acids, their photochemistry, and also in relation to the genetic code and mechanism of protein biosynthesis. Apart from these main subjects, research was also pursued on the mechanism of bacterial transformation, radiation chemistry of proteins and nucleic acids in relation to mutagenesis and radiation damage.

In the seventies studies on antiviral and anticancer agents became the central problem. These involved new methods for synthesis of analogues of nuclesides and nucleotides modified in the sugar moiety, supported by studies on their conformation to establish correlations between modification, conformation and biological activity. Later extensions of these investigations are studies on protein kinases and their inhibitors. On these problems D. Shugar has been mainly engaged during recent years.

In his curriculum vitae he wrote: "It is obvious that the implementation of such research programs was possible only because of collaboration with young, capable scientific research students, whose names appear on most of the publications. It is also obvious that without such capable coworkers, it would not have been possible for me to carry out such a program". Being deeply involved in scientific research he was able to work with everybody in whom he felt a spark of passion for research. The door to his room was always wide open enabling people to drop in at any time. He attracted people to his laboratories but was equally open for collaboration with other institutions. A list of D. Shugar publications includes over 450 positions (including books and reviews), the majority with one or several coworkers, many from laboratories both inside and outside the country. Friendly and informal contacts with laboratories abroad always served to enrich the studies and possibilities in his laboratories.

The above does not embrace all the activities and achievements of Professor David Shugar. Despite his retirement, he continues to be an active member of the research staffs at the Institute of Biochemistry and Biophysics, and at the Biophysics Department of the University. He attacks research problems with the same eagerness and sparkle in the eye as in preceding years. We hope Professor David Shugar will remain with us for further fruitful years.

Coworkers

Tautomerism of thioguanine. Experimental matrix isolation and theoretical quantum-mechanical studies

KRYSTYNA SZCZEPANIAK[1]
WILLIS B. PERSON[1]
JERZY LESZCZYŃSKI[2]
JÓZEF S. KWIATKOWSKI[3]

Contents:

Abbreviations used: B3LYP — Backe's nonlocal three parameter exchange and correlation functional with Lee-Yang-Parr correlation functional; DFT — density functional theory; DZP — double zeta polarization; MP — Mollett-Plesset; ZPE — zero point energy.

I. Introduction

Stimulated by fruitful and enjoyable discussions with Professor David Shugar, a group of physicists in Warsaw about 25 years ago began studies of tautomerism of isolated pyrimidine and purine bases, first in the vapor phase and then in low temperature inert matrices (noble gases and nitrogen). This latter exciting technique had been brought to Warsaw University from the University of Florida, Gainesville, where one of the authors (KS) spent her postdoctoral study. Matrix isolation infrared spectroscopy appeared to be very useful for studies of vibrational spectra and tautomerism of nucleic acid bases and their analogs, particularly because it does not require the extensive heating of the sample that is necessary in vapor studies. In this technique, the sample is prepared by sublimation of the solid at a temperature much below its melting point into a stream of inert gas followed by rapid condensation to form the inert cold rigid matrix [1]. The procedure avoids thermal decomposition and the isolation in the rigid matrix prevents additional reaction and further tautomeric changes in these sensitive compounds that often present problems for studies of their vapors.

In the late 70's the Warsaw group was the only one in the world doing matrix isolation studies of the vibrational spectra of **isolated** nucleic acid bases and their analogs [2-6]. These studies provided much needed experimental data on tautomerism of isolated pyrimidine and purine bases, which are crucial for verification of the reliability of theoretical calculations. In view of this, it is not surprising that the experimental matrix studies stimulated many new *ab initio* calculations of the stabilities and vibrational spectra of different tautomeric forms. By the 80's other laboratories also began experimental studies of vibrational spectra of these matrix isolated pyrimidine and purine bases. At the present time several groups in different countries perform such studies for these molecules [7-12], with the Warsaw continuing to contribute to these investigations (e.g., see reference [13-17]).

Our laboratory at the Chemistry Department of the University of Florida, joined several years ago by two members of the Warsaw group (KS and Marian Szczesniak), has collected a vast amount of data for many pyrimidine and purine bases and their derivatives. The spectra of these molecules are being analyzed on the basis of better and better *ab initio* molecular orbital calculations [18-31]. The progress in the computational methods is so rapid that before we complete

[1] Department of Chemistry, University of Florida, Gainesville, FL 32611, USA, [2] Department of Chemistry, Jackson State University, Jackson, MS 39217, USA, [3] Institute of Physics, N. Copernicus University, 87-100 Toruń, Poland

the analysis of our spectra using results from one calculation, newer and better calculations are completed, often requiring revision of the interpretation. This fact is partly responsible for the delay in publication of our results. But the main reason for this delay is the extreme difficulty of the interpretation of the complex vibrational spectra of these molecules, particularly when several tautomeric forms are present. The results discussed below illustrate some of the problems involved.

In this paper we shall present briefly results from our recent experimental matrix isolation studies of the vibrational spectra of tautomerism of thioguanine, together with results from some very recent calculations at the Hartree-Fock and density functional theory levels for thioguanine tautomers [32]. This molecule has special biological interest as a known metabolic inhibitor with antitumor and antineoplastic activity, and is used in cancer treatment and research (for example, see reference [33-37], and references therein).

As many other purine and pyrimidine bases do, thioguanine can potentially occur in several tautomeric forms. Figure 1 shows some of the tautomers and rotational isomers that are possible for this molecule. At present, calculations of structures, dipole moments, energies and vibrational spectra have been performed only for the first four forms shown in figure 1.

It is known from X-ray crystallographic studies [38] that thioguanine occurs in the crystal in the amino-thione tautomeric form with hydrogen attached to nitrogen atom N7 of the imidazole ring (structure 2 in Fig. 1), in sharp contrast with guanine, which occurs in the crystal and in solution as the amino-oxo form with the hydrogen on the imidazole ring at N9 [39-41]. The occurrence of the (N7H)amino-thione form in the crystal also is contrary to the prediction of relative tautomeric stabilities by recent *ab initio* calculations at MP2/(full)/DZP [42] and MP2/6-311+G(d,p) [43] levels, which is that the (N9H)amino-thiol form (structure 3 in Fig. 1, or its rotamer with SH *trans* to N1) is the most stable.

Such differences between theoretical predictions of relative stabilities of tautomers and the structure experimentally observed in the crystal may be a result of the effect of the intermolecular interactions, particularly hydrogen bonding, in the crystal which are not considered in the calculations at this relatively high level of theory. This result is somehow puzzling since the N7H-thione tautomer 2 found to occur in the crystal has a much smaller calculated dipole moment (1.78 D [42]) than does the N9H-thione tautomer 1 (7.42 D [42]), thus contradicting the usual belief that the more polar tautomer will be stabilized in the more polar medium of the crystal. In view of these observations, the importance of **experimental verification** of the relative stabilities of tautomers of **isolated** thioguanine

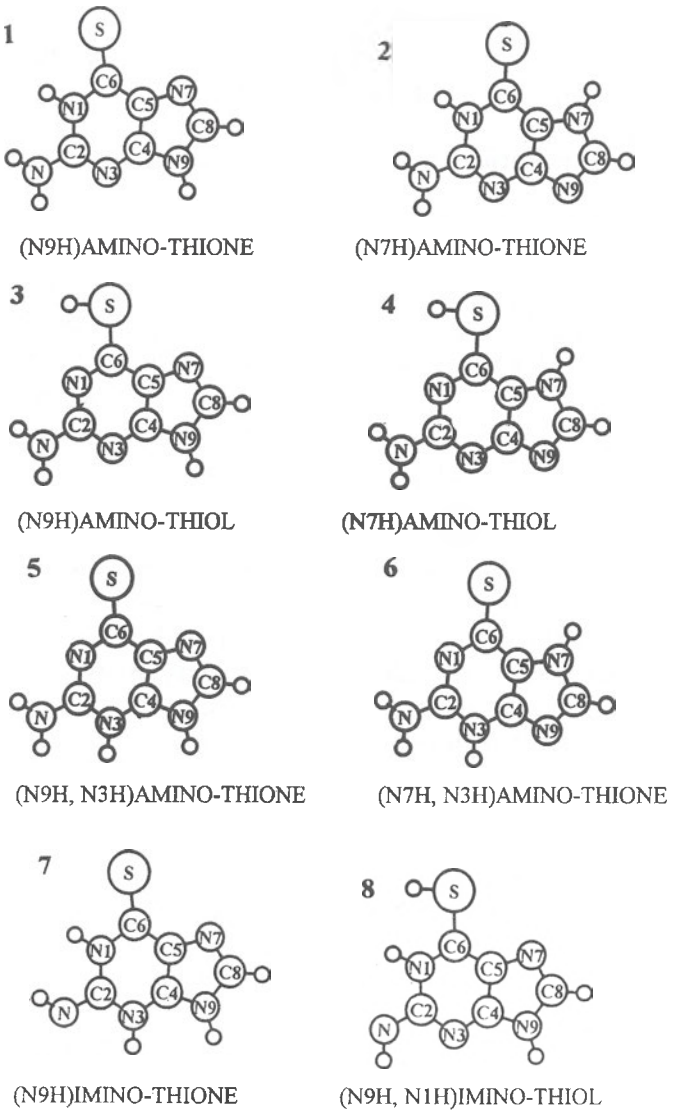


Fig. 1. Some possible tautomers for thioguanine. Structures 1-4 are drawn for the optimized geometry from the DFT(B3LYP)/6-31G(d,p) calculations, which have predicted that the NH2 group is non planar. Structures 5-8 are drawn for arbitrary geometries, starting from the first four structures but with changed positions of the hydrogen atoms.

for comparison with the *ab initio* calculations made for the molecule in vacuum is quite obvious.

II. Methods

II-1. Experimental

The details of the experimental technique of matrix isolation infrared spectroscopy have been described elsewhere [1]. In this study a mixture of the vapor from thioguanine (purchased from Sigma) heated to about 227°C and the matrix gas (argon) was condensed onto a cold (about 12 K) window held in the cryostat (Air Products Model 202E Displex). The infrared spectra were measured at a resolution of 1 cm^{-1} using a Fourier Transform Infrared (FTIR) Nicolet Model 740 spectrometer. The final spectra shown below were drawn by the SpectraCalc program (Galactic) on a personal computer, after electronic transfer of the data from the Nicolet spectrometer to the PC. This

program was also used to subtract the interference fringes caused by the thin matrix film and to correct the background to the horizontal "zero" level.

II-2. Calculation

The calculations reported here were made using density functional theory (DFT) [44-45] with the Becke three parameter exchange functional and the gradient corrected functional by Lee, Yang and Parr (the DFT(B3LYP) method) (see references [46-51], and references therein). The standard 6-31G(d,p) basis set was used. Optimizations were carried out without any constraints on the planarity of the species under study. All quantum mechanical calculations were performed using the GAUSSIAN 92/DFT program [52] at the Mississippi Center for Supercomputing Research. Calculations at this level have recently been the focus of a lot of attention (see reference [53], and references therein) because the calculated vibrational spectra are usually in good agreement with experiment; the agreement for frequencies is often within the uncertainty due to anharmonicity corrections.

On the basis of knowledge of small molecules [54-57] it can be expected that the correction for anharmonicity will be the largest for the X-H (where X is N1, N7, S, C or N for symmetric or asymmetric NH2) stretching frequencies, and relatively smaller corrections for other frequencies. Based on our experience with other molecules, we take these different anharmonicity corrections into account by scaling all N-H and S-H stretching frequencies by 0.955 (force constants by $(0.955)^2$ or 0.912), all C-H stretching frequencies by 0.964 (force constants by $(0.964)^2$ or 0.929), and all other frequencies by 0.973 (force constants by $(0.973)^2$ or 0.947) in the comparisons shown here. The use of these different scaling factors does not change the normal coordinate mixing because it changes only force constants for stretching coordinates which do not mix with other coordinates. These scaling factors are not significantly different from those recommended by Pulay [53] for spectra from DFT(B3LYP)/ /6-31G(d,p) calculations for a large number of molecules. In general, the calculated frequencies after scaling agree with experimental frequencies within an uncertainty of about 10-20 cm^{-1}.

The calculated spectra, structures, and atom displacements in the normal modes displayed here are drawn using the Animol 3.1 program (Innovative Software) installed under Windows in a PC, which reads its initial input directly from the GAUSSIAN output file.

III. Results and discussion

Usually the most straightforward identification of tautomers can be made on the basis of the presence or absence in the infrared spectrum of characteristic absorption bands for the functional groups involved in the tautomeric transitions. In the case of thioguanine these functional groups are **N1H, N3H, N7H, N9H,** asymmetric and symmetric **NH2,** imino (**C=NH**) and **C=S** stretching vibrations. For example the most direct evidence of the presence of the thiol tautomer would be the appearance of the SH stretching absorption band near 2600 cm^{-1} [17, 24]. Unfortunately the infrared intensity of this band is expected to be very low (about 1 km mol^{-1}; see references [17, 32]) and its presence or absence is difficult to decide due to confusion with combinations or overtones which are also expected to absorb in this region.

It is equally difficult to decide about the presence or absence of the C=S stretching band (and thus about the presence of the thione tautomer) because, in contrast with the C=O (carbonyl) stretching mode, the C=S stretching is not characteristic. The C=S stretching coordinate contributes to several normal modes and is mixed strongly in them with other coordinates (see below).

For these reasons the identification of the thiol and thione tautomers can be done with more assurance by examining carefully the infrared absorption in other spectral regions. The analysis relies heavily on the interpretation based on the calculated spectra for the different tautomers.

III-1. N1H, N7H, and NH2 stretching modes

One of the most useful hints about tautomers present in the matrix comes from examining the effect of short (15 minute) irradiation of the matrix isolated sample by UV irradiation from a medium pressure mercury lamp. As a result of this irradiation, some weak bands observed before irradiation disappear; these can be identified as absorption by amino-thione tautomers. Figure 2a shows a comparison of the experimental spectra in the NH stretching region from thioguanine isolated in an Ar matrix before and after the short UV irradiation with the spectra calculated for the amino-thiol tautomers. As can be seen there, the strongest bands observed at **3574, 3564, 3489, 3456** and **3449** cm^{-1} in the spectrum (practically unchanged after irradiation) have frequencies and intensities consistent with those calculated for the asymmetric NH2 stretch, the N9H (and/or the N7H) stretch, and the symmetric NH2 stretch, respectively, for both of the amino-thiol tautomers, but considerably different from the spectrum calculated for the amino-thione forms shown in figure 2b. That figure shows the experimental difference spectrum obtained by subtracting the spectrum after UV irradiation from the original spectrum compared with the spectra calculated for the two thione tautomers shown.

The weak bands observed at **3533, 3526, 3463, 3432, 3426, 3415, 3405,** and **3399** cm^{-1} in the initial spectrum, and also shown in the difference spectrum, have

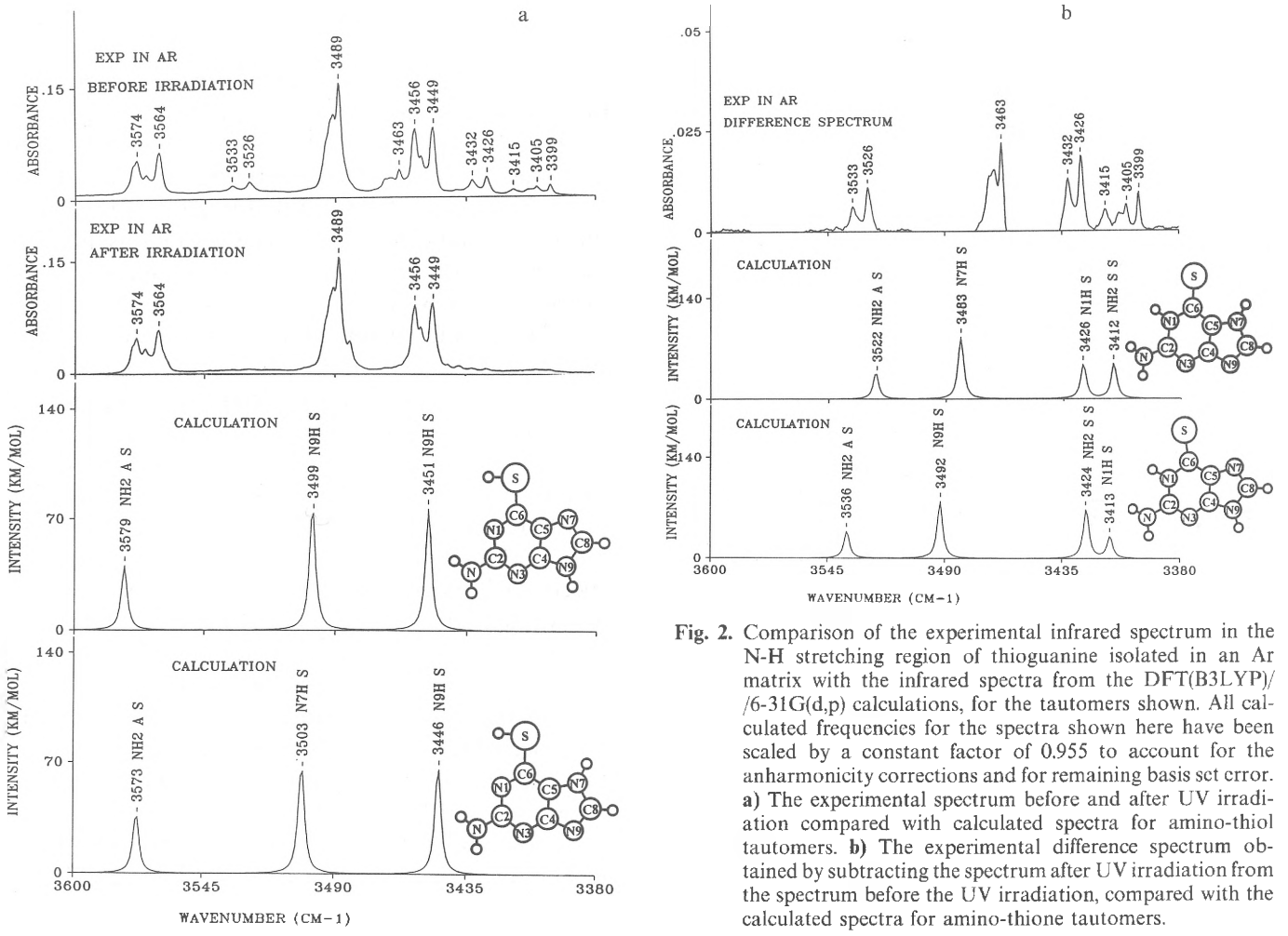


Fig. 2. Comparison of the experimental infrared spectrum in the N-H stretching region of thioguanine isolated in an Ar matrix with the infrared spectra from the DFT(B3LYP)/ /6-31G(d,p) calculations, for the tautomers shown. All calculated frequencies for the spectra shown here have been scaled by a constant factor of 0.955 to account for the anharmonicity corrections and for remaining basis set error. **a)** The experimental spectrum before and after UV irradiation compared with calculated spectra for amino-thiol tautomers. **b)** The experimental difference spectrum obtained by subtracting the spectrum after UV irradiation from the spectrum before the UV irradiation, compared with the calculated spectra for amino-thione tautomers.

frequencies close to those calculated for the asymmetric and symmetric NH2 and the N1H stretching modes in both of the amino-thione tautomers. However, the experimentally observed band at 3463 cm^{-1} is located at lower frequency than calculated for the N7H or N9H stretches (but within the uncertainty of the calculation). It seems most likely that the species responsible for these absorption bands are amino-thione tautomers.

Our first interpretation of the doublet structure for the bands observed for the strong NH2 stretching vibrations (**3574, 3564; 3456** and **3449** cm^{-1}) and for the weak bands **3533, 3526; 3432, 3426;** and **3415, 3405** (and/or **3399** cm^{-1}) was that it was caused by the presence in the matrix of a mixture of N7H and N9H amino-thiol tautomers (strong bands) with N7H and N9H amino-thione tautomers (weak bands). Arguments causing some reservation about this interpretation are: (i) the observed splitting is different than the calculated frequency differences; and (ii) the relative intensities of the components of the experimental doublets differ from one matrix to another. For these reasons we rather prefer to interpret the splitting in the experimental spectrum as resulting from "matrix effects" on the spectrum of one tautomer. However, it is surprising that this doublet pattern is practically insensitive to annealing, in contrast with similar observations for adenine ([16] and references therein). These splittings clearly need further study.

The comparison of experimental and calculated spectra shown in figure 2 suggests that the **predominant** tautomer in the matrix is an **amino-thiol** from, but it is not possible to distinguish whether it is the (N7H) or (N9H) tautomer since the calculated spectra shown in figure 2a for both forms are very similar. The weak bands listed above which disappear after UV irradiation are most probably nearly all due to vibrations of an amino-thione form. Again, the results shown in figure 2 do not allow us to determine with certainty whether it is the (N7H) or the (N9H) amino-thione form.

III-2. S-H stretching mode

Figure 3 shows a comparison of the experimental and calculated spectra in this region (near 2600 cm^{-1}). Six bands are observed in the experimental spectra, each with an absorbance about 30 to 50 times lower that for the NH2 asymmetric stretching mode, consistent with the calculated ratio of intensities for these two modes (note the intensity scale). However, it is very likely that some of the observed bands are combination bands involving the lower frequency fundamentals, which may also be expected to have similar intensities.

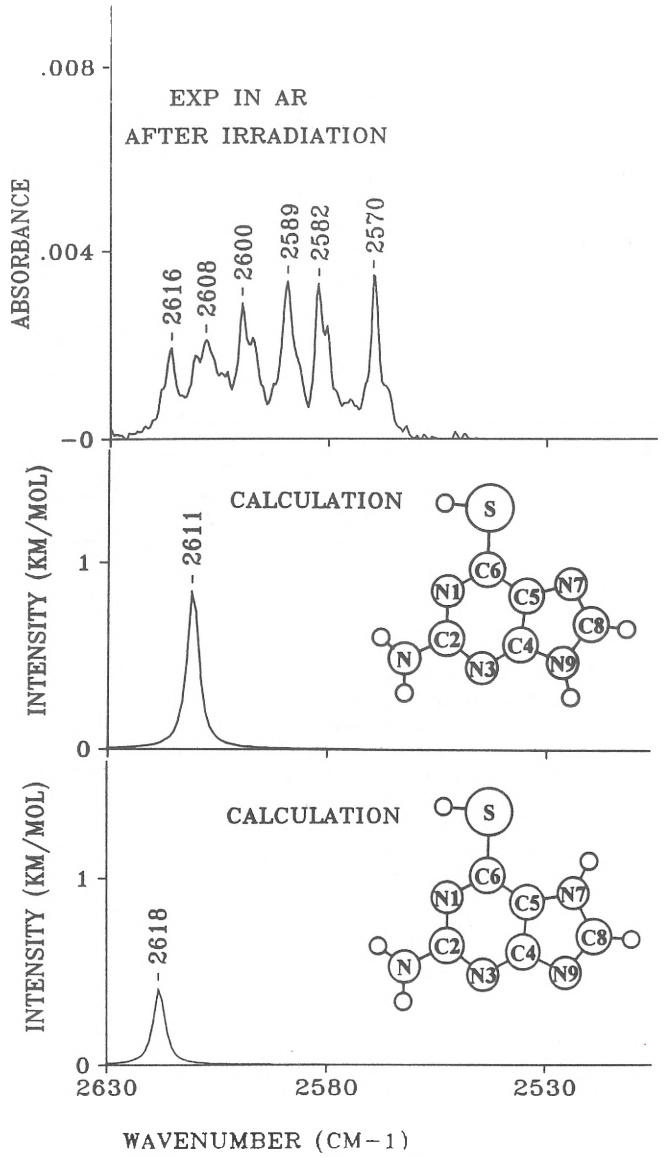


Fig. 3. Comparison of the experimental infrared spectrum in the S-H stretching region of thioguanine isolated in an Ar matrix with the spectra calculated (DFT(B3LYP)/6-31G(d,p)) for the two amino-thiol tautomers. Some of the bands observed in the experimental spectrum may be due to combination bands, such as: (1590+1026 = 2616), (1352+1263 = 2615), (1499+ +1088 = 2587), (1620+961 = 2581), (1263+1314 = 2571), or (1620+956 = 2576) cm^{-1}.

Some possible combinations that could be responsible for some absorption in this region are listed in the caption of figure 3.

Because of the uncertainty about the best value of the frequency scaling factor that should be used the S-H stretch with these DFT calculations, there is uncertainty in the value of the calculated frequency that should be compared in figure 3 with the experimental spectrum. If the errors are the same for both the N7H and the N9H tautomer, the calculation predicts a slightly lower S-H stretching frequency in the latter tautomer.

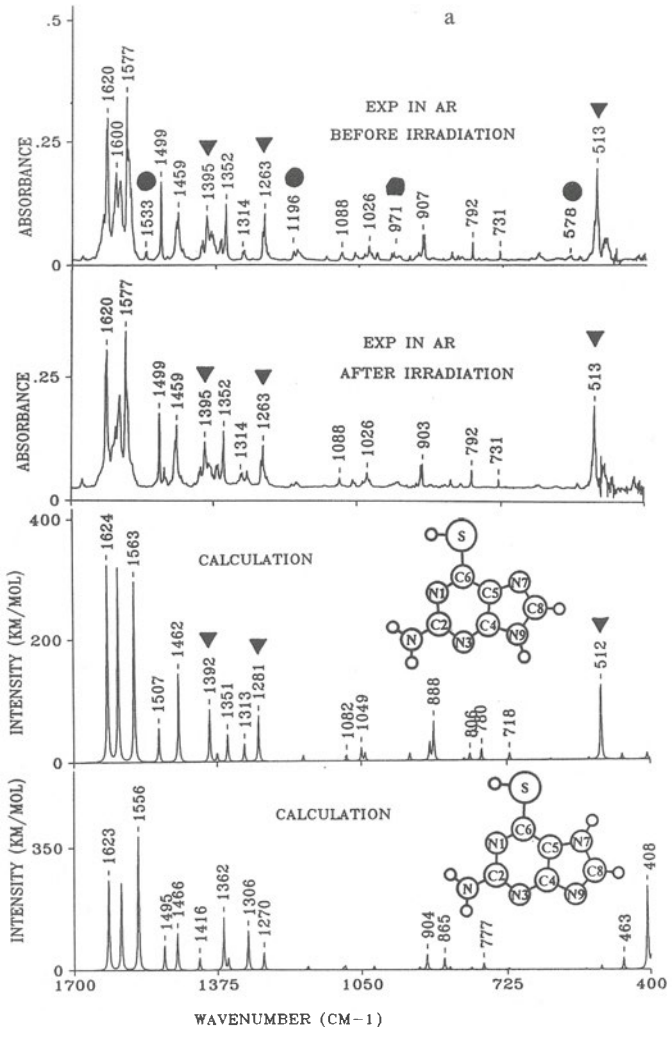
Because of the very low intensity of the S-H stretching mode, the only judgment concerning the presence of thiol tautomers in the matrix from the spectrum in this region can be support for conclusions based on information from other parts of the spectrum.

III-3. Other vibrations

As has been emphasized above, the comparison of the experimental and calculated spectra of thioguanine tautomers shown for the NH stretching region in figure 2 does not permit the final determination of whether it is the (N7H) or (N9H) form that is actually present in the matrix. Some information bearing on this question can be obtained by examination of the spectra in the lower frequency region (1700-400 cm^{-1}). Figure 4 shows a comparison of the experimental matrix spectra before and after UV irradiation and the difference spectrum with the spectra calculated for the four tautomers shown there. This spectral region includes 37 of the 42 normal modes of thioguanine, and is very difficult to analyze. Full interpretation of the experimental vibrational spectrum based on a detailed analysis of the calculated and experimental spectra will be given in reference [32]. Here we shall focus attention only on some of the most distinctive bands.

As can be seen in figure 4, the calculated spectra for the different tautomers are not significantly different for much of this spectral region. Fortunately, even so, there are a few relatively strong bands in the spectral patterns of the (N7H) and (N9H) thiol tautomers that are calculated to be characteristically different. These include the bands (marked by solid triangles ▼) calculated at **1392, 1281** and, particularly, **512** cm^{-1} in the spectrum of the (N9H)amino-thiol tautomer related to the displacements of atoms for these modes, shown in figure 5, with its contribution from the N9H group. These modes appear to correspond to the bands observed in the experimental spectra at **1395, 1263** and **513** cm^{-1}, respectively. On the basis of this observation, we conclude that the **predominant amino-thiol** tautomer present in the matrix *is* the **(N9H)** form and *not* the (N7H) form. This conclusion is strongly supported by the failure to observe any absorption in the matrix spectrum from the N7H wagging mode calculated to absorb strongly at 408 cm^{-1} for the (N7H)amino-thiol form.

It is much more difficult to distinguish between the (N7H) and (N9H) forms of the amino-thione tautomers, because of the very low concentration of this form in the matrix and consequently the very low absorbances of these bands in the experimental spectra. Some indication that the thione tautomer in the matrix is the (N7H) form comes from the observation in the experimental spectra in figure 4 of the bands at **1533, 1196, 971** and **578** cm^{-1} (marked by filled circles ●) which disappear after UV irradiation and are shown in the difference spectrum in figure 4b. These bands are close to the relatively intense modes at **1523, 1168, 959** and **569** cm^{-1} (also marked by filled circles ● in Fig. 4b) calculated for the (N7H)amino-thione tautomer that are related to the normal mode atom displacements (shown in Fig. 6) with contributions



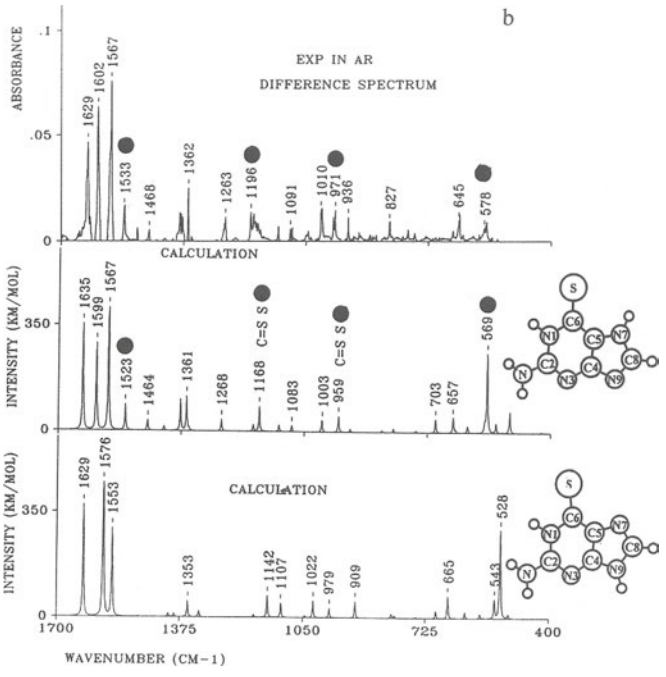


Fig. 4. Comparison of the experimental infrared spectrum in the 1700-400 cm^{-1} region of thioguanine isolated in an Ar matrix with the spectra calculated (DFT(B3LYP)/6--31G(d,p)) for the tautomers shown. All calculated frequencies in this region have been scaled by 0.973 to account for anharmonicity and basis set error. **a)** The experimental spectrum for the sample before and after UV irradiation compared with spectra calculated for two amino-thiol tautomers. **b)** The experimental difference spectrum obtained by subtracting the spectrum after UV irradiation from that before UV irradiation compared with the calculated spectra for two amino-thione tautomers.

from the C=S stretch and the N7H motions, strongly mixed with other displacements.

In summary, our conclusion from the examination of figure 2 and 4 is that the **predominant** tautomer present in the initial deposit of the matrix from the vapor is the **(N9H)amino-thiol** form shown as structure 3 in figure 1 (or the rotamer with SH *trans* to N1). A smaller concentration of the (N7H)amino-thione tautomer 2 in figure 1 is also present.

III-4. Energies

The predominance in the matrix spectrum of absorption by tautomer 3 in figure 1 agrees with the relative stabilities predicted for the isolated molecules by the *ab initio* calculations [42, 43].

If we make the assumption that the analysis given above has correctly identified only two tautomeric forms present in the matrix spectrum of thioguanine (namely the (N9H amino-thiol predominant form and the lesser concentration of the (N7H)amino-thione form), then we may use the relative intensities observed to determine the equilibrium constant for the tautomeric equilibrium at the temperature of the gaseous mixture as the equilibrium was trapped during the formation of the matrix.

The ratio of the observed integrated intensity of the

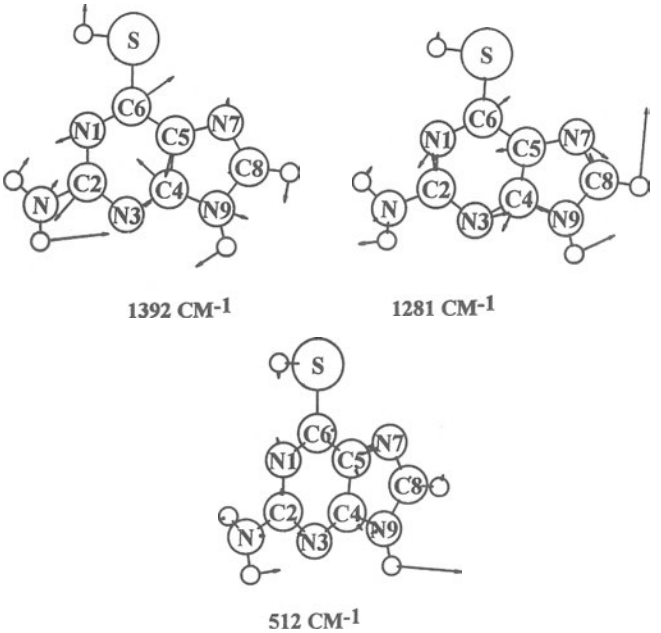


Fig. 5. Calculated displacements of the atoms, shown by the lengths of the arrows, in normal modes for the (N9H)amino-thiol tautomer that are used to distinguish between the (N9H)amino-thiol and (N7H)amino-thiol tautomers in the matrix spectrum. Note that all these normal modes show large contributions from motion both of the N9H group and of the SH group (as well as contributions from motion of several other atoms). The mode at 512 cm^{-1} is an out of plane "wagging" motion.

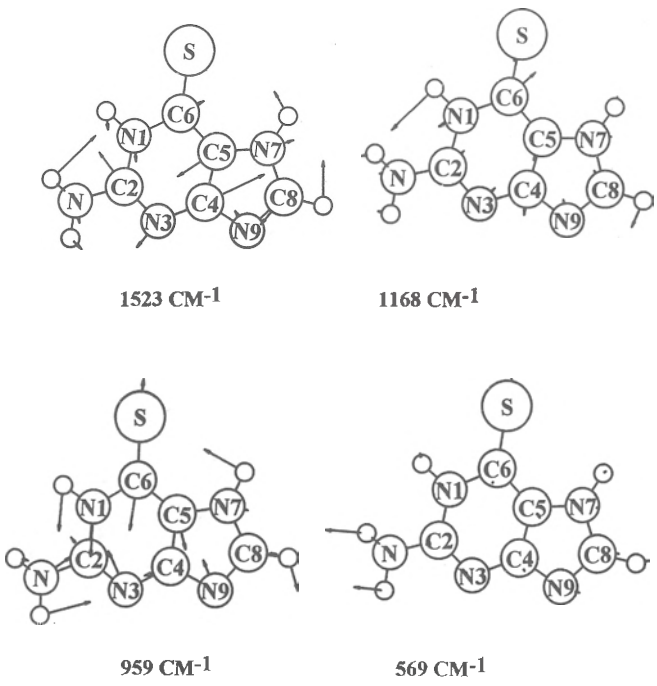


Fig. 6. Calculated displacements of atoms, shown by the lengths of arrows, in normal modes for the (N7H)amino-thione that distinguish between the (N7H)amino-thione and (N9H)amino-thione forms. Note that the N7H bending contributes to the modes at 1523 and 959 cm^{-1}, while C=S stretching and N1H bending motions contribute to the normal modes at 1168 and 959 cm^{-1}. However, all these modes have sizable contributions from motion of other atoms and cannot be considered to be characteristic C=S stretches. The 569 cm^{-1} mode is mostly "out of plane NH2 wagging".

asymmetric NH2 stretch at 3564 and 3574 cm^{-1} in the amino-thiol form to that of the amino thione form at 3526 and 3533 cm^{-1}, together with the corresponding calculated absolute integrated molar absorption coefficients A (41 and 37 km mol^{-1}, respectively) [32], allows us to evaluate the relative concentration of the (N9H)amino-thiol tautomer to that of the (N7H)amino-thione form, trapped in the matrix. The results are that the percentage concentration of the former is 85%, and that of the latter is 15%. For an equilibrium established in the gas phase during sublimation at 500 K (227°C) and subsequently trapped in the matrix, the difference in the standard Gibbs free energy of the more stable amino-thiol form from that for the amino-thione form is thus $\Delta G^0_{500} = -7.1 \pm 1.5$ kJ mol^{-1}.

The values of the change in the internal energy ΔE^0_0 at 0 K have been obtained from the *ab initio* calculations as the sum of the calculated electronic energy differences ΔE_{el} and zero point energy differences (ΔZPE). Both contributions are given in table 1 for tautomeric transitions defined there. According to the calculation, the change in the ZPE is largely responsible for the stability of the thiol tautomer with respect to the thione.

In order to compare the calculated energies with the experimental measurement of the equilibrium constant and the associated value of ΔG^0_{500}, it is appropriate to use the statistical thermodynamic relations [58] to convert the latter to the experimental value of ΔE^0_0, using the calculated rotational constants and frequencies of the normal modes. From this, $\Delta E^0_0 = -5.6$ kJ mol^{-1}.

This value is compared in table 1 with the calculated changes in internal energy for several tautomeric reactions involving tautomers considered here. We see there that the experimental value of ΔE^0_0 (-5.6 kJ mol^{-1}) for the tautomeric reaction from the (N7H)amino-thione form of structure 2 to the (N9H)amino-thiol form of structure 3 is about 3.8 kJ mol^{-1} more negative than calculated value (-1.8 kJ mol^{-1}). Structure 3 of the (N9H)amino-thiol form (in which the S-H bond is *cis* to the N1 atom) has been assumed to be more stable than the rotamer with the

Table 1.

Energies in kJ mol^{-1} for the tautomeric reaction of (N7H)amino-thione thioguanine tautomer to form other tautomers, from calculations. For the reaction:

(N7H)amino-thione → B.

Tautomer (B)	Calculation				Experimental		Calc.[b]
	ΔE_{el}[a]	ΔZPE[a]	ΔE^0_0[a]	ΔE^0_{298}[b]	ΔE^0_0[c]	ΔG^0_{500}[c]	ΔG^0_{298}
(N9H)amino-thiol *cis*	+7.90	−9.7	−1.8	+0.1			−2.5
					−5.6±1	−7.1±1	
trans	—[d]		(−2.7)[e]	−0.8			−3.4
(N7H)amino-thione[f]	0	0	0	0	0	0	0
(N9H)amino-thione	+11.55	−0.55	+11.1	+2.7	> +5	> +5	+2.5
(N7H)amino-thiol *cis*	+27.99	−10.87	+17.0	+4.8	> +5	> +5	

[a] MP2(full)/DZP//HF/DZP calculations given by Leszczynski, reference [42].

[b] Values reported to be ΔE^0_{298} (internal energies including ZPE and thermal energies) or ΔG^0_{298} from MP2/6-311++G(d,p)//HF/6-31 G(d) calculations by Alhambra *et al.*, reference [43].

[c] This work. See text. The values for (N9H)amino-thiol could be for the *cis* or *trans* rotamer. Failure to observe any trace of a tautomer indicates its equilibrium constant is less than 0.05, implying ΔG^0_{500} and ΔE^0_0 values 5-10 kJ mol^{-1} more positive than the energy of the (N7H)amino-thione tautomer.

[d] Not calculated.

[e] Estimated from the calculated value of the energy of the *cis* tautomer, added to the difference calculated between the energies of *cis* and *trans* tautomers in *b*.

[f] Reference molecule. All energies calculated relative to this tautomer.

SH bond *trans* to the N1 atom. However, recent calculations by Alhambra *et al.* [12], of the energies of tautomers of thioguanines have found that the *trans* form is about 0.9 kJ mol^{-1} more stable than the *cis* form (structure 3).

If the observed matrix spectrum is for the *trans* form rather than the *cis* form, then the observed value of ΔE_0^0 (-5.6 kJ mol^{-1}) should be compared with a calculated value for that rotamer. This value can be estimated for the present by correcting the value from Leszczynski [42] in table 1 (-1.8 kJ mol^{-1}) by adding this calculated *cis-trans* energy difference [43], to obtain ΔE_0^0 for the *trans* rotamer to be -2.7 kJ mol^{-1}. Such close quantitative agreement between calculated (-5.6) and experimental (-2.7 kJ mol^{-1}) energies is much better than can be expected, and encourages us to believe that the conclusions about tautomerism in thioguanine from this study may be correct.

IV. Concluding remarks

The main conclusions from this study are summarized below.

(1) Thioguanine occurs in a non-polar (hydrophobic) environment predominantly (about 85% for an equilibrium established at 500 K) as the (N9H)amino-thiol tautomer. Only a small concentration (about 15% at 500 K) of the molecules occur as the (N7H)amino--thione form. It is possible that small concentrations of imino-thione or imino-thiol tautomers may be present in the matrix. Further experimental and theoretical studies are needed to confirm (or contradict) this last suggestion.

(2) This finding for thioguanine is in sharp contrast with the exclusive occurrence of the (N7H)amino--thione form in the polar environment of the crystal. It is also in contrast with the finding for guanine in the matrix environment that three tautomers [(N7H)amino-oxo, (N9H)amino-oxo, and (N9H)amino-hydroxy] occur with *similar* concentrations [4, 59].

(3) The greater stability found in this study for the (N9H)amino-thiol tautomer in the hydrophobic environment confirms the predictions from recent *ab initio* calculations, and the experimental change in energy, ΔE_0^0, for the tautomeric transition from the (N7H)amino-thione form is in good agreement with the value from the calculation. Additional calculations at the MP2(full)DZP level of the energy of the *trans* rotamer of the (N9H)amino-thiol form, which may be the most stable tautomer of thioguanine, are in progress.

Acknowledgments

First of all, it is a pleasure to acknowledge the inspiration provided by Professor David Shugar, and to dedicate this report to him. Partial financial support (KS and WBP) from NIH Grant No. 32988 and from a Research Development Award from the Division of Sponsored Research of the University of Florida is gratefully acknowledged. We are grateful also to NIH Grant No. 332090, to contract (DAAL 03-89-0038) with the Army Research Office, and to the Mississippi Center for Supercomputing Research for support (JL) of the calculations. JSK wishes to acknowledge support by the Committee for Scientific Research within grant KBN 3T09A 026 08.

References

1. Person WB, Szczepaniak K, Szczesniak M, Del Bene JE (1993) In: Fausto R (ed) Recent Experimental and Computational Advances in Molecular Spectroscopy. NATO ASI Series C: Mathematical and Physical Sciences, vol 406. Kluwer Academic Publishers, The Netherlands, Dordrecht, pp 141-169
2. Nowak MJ, Szczepaniak K, Barski A, Shugar D (1978) *Z. Naturforsch* Teil C **33**: 876-883
3. Nowak MJ, Szczepaniak K, Barski A, Shugar D (1980) *J Mol Struct* **63**: 47-69
4. Shugar D, Szczepaniak K (1981) *Int J Quantum Chem* **20**: 573-585
5. Szczesniak M, Nowak MJ, Szczepaniak K (1984) *J Mol Struct* **115**: 221-224
6. Szczepaniak K, Szczesniak M (1986) *J Mol Struct* **156**: 29-42
7. Radchenko ED, Plochotnichenko AM, Sheina GG, Blagoi YuP (1983) *Biophysics* **28**: 559-563 (English edn)
8. Barnes AJ, Stuckey MA, LeGall L (1984) *Spectrochim Acta* **40A**: 419-431
9. Radchenko ED, Sheina GG, Smorygo NA, Blagoi YuP (1984) *J Mol Struct* **116**: 387-396
10. Stepanian SG, Sheina GG, Radchenko AD, Balgoi YuP (1985) *J Mol Struct* **131**: 333-346
11. Stepanian SG, Radchenko ED, Sheina GG, Blagoi YuP (1990) *J Mol Struct* **216**: 77-90
12. Graindourze M, Smets Y, Zeegers-Huyskens Th, Maes G (1990) *J Mol Struct* **222**: 345-364
13. Nowak MJ (1989) *J Mol Struct* **193**: 35-49
14. Nowak MJ, Lapinski L, Fulara J (1989) *Spectrochim Acta* **45A**: 229-242
15. Fulara J, Nowak MJ, Lapinski L, Les A, Adamowicz L (1991) *Spectrochim Acta* **47A**: 595-613
16. Nowak MJ, Rostkowska H, Lapinski L, Kwiatkowski JS, Leszczynski J (1994) *J Phys Chem* **98**: 2813-2816
17. Nowak MJ, Rostkowska H, Lapinski L, Leszczynski J, Kwiatkowski JS (1991) *Spectrochim Acta* **47A**: 339-353
18. Szczesniak M, Nowak MJ, Rostkowska H, Szczepaniak K, Person WB, Shugar D (1983) *J Am Chem Soc* **106**: 5969-5976
19. Chin S, Scott I, Szczepaniak K, Person WB (1984) *J Am Chem Soc* **106**: 3415-3422
20. Szczesniak M, Nowak MJ, Szczepaniak K, Chin S, Scott I, Person WB (1985) *Spectrochim Acta* **41A**: 223-235
21. Kwiatkowski JS, Person WB, Szczepaniak K, Szczesniak M (1987) *Acta Biochim Polon* **34**: 165-181
22. Szczesniak M, Szczepaniak K, Kwiatkowski JS, KuBulat K, Person WB (1988) *J Am Chem Soc* **110**: 8319-8330
23. Szczepaniak K, Szczesniak M, Person WB (1988) *Chem Phys Lett* **153**: 39-44
24. Rostkowska H, Barski A, Szczepaniak K, Szczesniak M, Person WB (1988) *J Mol Struct* **176**: 137-147
25. Person WB, Szczepaniak K, Szczesniak M, Kwiatkowski JS, Hernandez L, Czerminski R (1989) *J Mol Struct* **194**: 239-258
26. Wojcik MJ, Rostkowska H, Szczepaniak K, Person WB (1989) *Spectrochim Acta* **45A**: 499-502
27. Jaworski A, Szczesniak M, Szczepaniak K, KuBulat K, Person WB (1990) *J Mol Struct* **223**: 63-92
28. Rostkowska H, Szczepaniak K, Nowak MJ, Leszczynski J, KuBulat K, Person WB (1990) *J Am Chem Soc* **112**: 2147-2160

29. Szczepaniak K, Szczesniak M, Szajda W, Person WB (1991) *Can J Chem* **69**: 1705-1720
30. Szczesniak M, Leszczynski J, Person WB (1992) *J Am Chem Soc* **114**: 2731-2377
31. Person WB, Szczepaniak K (1993) In: Durig JR (ed) Vibrational Spectra and Structure, vol 20. Elsevier, The Netherlands, Amsterdam, pp 239-325
32. Szczepaniak K, Kasende O, Person WB, Leszczynski J, Kwiatkowski JS (Full interpretation of the vibrational spectrum of thioguanine; manuscript in preparation for J Am Chem Soc)
33. Hunter CJ, Deen DF, Marton LJ (1989) *Int J Cancer* **44**: 658-660
34. Pieters R, Huismans DR, Loonen AH, Hahlen K, Veerman AJP (1991) *Jpn J Cancer Res* **82**: 1051-1055
35. Hunter CJ, Deen DF, Marton LJ (1989) *Int J Cancer* **44**: 658-660
36. Phillips MD, Nascimbeni B, Tice RR, Shelby MD (1991) *Environ Mol Mutagen* **18**: 168-183
37. Pieters R, Huismans DR, Loonen AH, Peters GJ, Hahlen K, Van der Does-van den Berg A, Van Wering ER, Veerman AJP (1992) *Int J Cancer* **51**: 213-217
38. Bugg CE, Thewald U (1970) *J Am Chem Soc* **92**: 7441-7745
39. Thewald U, Bugg CE, Marsh RE (1971) *Acta Cryst* **B27**: 2358-2363
40. Majoube M (1984) *J Chim Phys* **81**: 303-315
41. Shugar D, Psoda A (1990) In: Nucleic Acids. Landolt-Bernstein New Series of Biophysics, Part I. Springer Verlag, Berlin
42. Leszczynski J (1994) *J Mol Struct (Theochem)* **311**: 37-44
43. Alhambra C, Luque FJ, Estelrich J, Orozco M (1995) *J Org Chem* **60**: 969-976
44. Parr RG, Yang W (1989) In: Density Functional Theory of Atoms and Molecules. Oxford University Press, New York
45. Labanowski JK, Anzelm JW (eds) (1991) In: Density Functional Methods in Chemistry. Springer Verlag, New York
46. Becke AD (1992) *J Chem Phys* **96**: 2155-2160
47. Becke AD (1992) *J Chem Phys* **97**: 9173-9177
48. Becke AD (1993) *J Chem Phys* **98**: 5648-5652
49. Lee C, Yang W, Parr RG (1988) *Phys Rev* **B41**: 785-789
50. Vosko SH, Wilk L, Nusair M (1980) *Can J Phys* **58**: 1200-1211
51. Ziegler T (1991) *Chem Rev* **91**: 651-667
52. Frish MJ, Trucs GW, Schlegel HB, Gill PMW, Johnson BG, Wong MW, Foresman JB, Robb MA, Head-Gordon M, Replogle ES, Gomberts R, Andres JL, Raghavachari K, Binkley JS, Gonzales C, Martin RL, Fox DJ, Defrees DJ, Baker J, Stewart JJP, Pople JA (1993) In: Gaussian 92/DFT Revision F.2, Gaussian, Inc. Pittsburg, PA
53. Rauhut G, Pulay P (1995) *J Phys Chem* **99**: 3093-3100
54. Herzberg G (1950) In: Molecular Spectra and Molecular Structure, vol 1. Van Nostrand Reinhold Co, New York, pp 502-580
55. Mills IM (1975) In: Specialist Periodical Reports of the Chemical Society, Theoretical Chemistry, vol 1 pp 110-159
56. Green WH Jr, Willetts A, Jayatilaka D, Handy NC (1990) *Chem Phys Lett* **169**: 127-137
57. Duncan JL (1991) *Spectrochim Acta* **47A**: 1-27
58. Atkins PW (1994) Physical Chemistry, 5th edn. Freeman, New York, chap. 20, pp 691-716; particularly Equation 33
59. Szczepaniak K, Szczesniak M, Szajda W, Person WB, Leszczynski J, Kwiatkowski JS (Tautomerism of guanine; manuscript in preparation for J Am Chem Soc)

This work is dedicated to Professor David Shugar my teacher and best friend

Mechanisms of action of methyl methanesulfonate on *Escherichia coli*: Mutagenesis, DNA damage and repair

CELINA JANION*

Contents:

Abbrevations used: XP — xeroderma pigmentosum, FA — Fanconi's anemia; BS — Bloom's syndrome; TFHII — protein subunits of transcription factor in human, required for initiation of transcription of genes and nucleotide excision repair; HNPCC — hereditary nonpolyposis colorectal cancer; MMS — methyl methanesulfonate; MNNG — *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine; MNU — *N*-methyl-*N*-nitrosourea; MFD — mutation frequency decline; TRCF — transcription repair coupling factor.

* Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5A, 02-106 Warsaw, Poland

I. DNA repair processes and cancer induction

Recent studies have provided astonishing data on the role of DNA repair in avoidance of cancer. It has long been known that a deficiency, or disturbance, in

DNA repair may lead to genetic instability, and a predisposition to cancer, in patients with xeroderma pigmentosum (XP), Fanconi's anemia (FA), Bloom's syndrome (BS); and to such hereditary neurological and neuromuscular diseases as Hungtinton's or Cockayne's syndrome. The origin of these diseases is a deficiency in some of the systems of DNA repair: nucleotide excision repair, DNA-ligase, preferential nucleotide excision repair (human gene *ERCC6* corresponds to *mfd* of *E. coli*), in TFHII subunit required for initiation of gene transcription and nucleotide excision repair [1-3]. Quite recently it was discovered that hereditary nonpolyposis colorectal cancer (HNPCC) in humans develops because of mutation in *hMSH2*, *hMLH1*, *hPMSH1*, or *hPMSH2* genes and deficiency in DNA mismatch repair [4-8].

All these DNA repair systems were first discovered and well recognized in bacteria, then in yeast and/or in human cells.

In the last decade, much attention was focused on the *p53*-human gene encoding the p53 protein. A mutated form of this protein was found in about 50% of all human cancer tissues [9, 10]. In the normal tissues, p53 protein is undetectable.

Protein p53 is a transcription factor, anti-oncogene product, suppressor of cancer, and a genomic guardian of the human cells. The synthesis of p53 is much increased after damage to DNA (e.g. by ionizing radiation) and this stimulates synthesis of the other proteins involved in arrest of cell cycle, DNA synthesis and repair [11-13].

The cascade of reactions which occurs after DNA damage in human cells resembles that operating in *E. coli* where it is known as the SOS system. The hypothesis that malignancy develops because the cells cannot scope with DNA repair, in 80-90% of cancer cases, seems to be true [14].

Organisms are continuously exposed to DNA damage from insults of endogenous and exogenous (environmental) mutagens. Methylating agents are widespread in the environment. In this paper mechanisms of damage and DNA repair, the systems defending *E. coli* cells against the killing and mutagenic effects of methyl methanesulfonate (MMS), will be briefly reviewed.

II. MMS-induced lesions and defence systems against alkylating agents

MMS is a monofunctional agent which methylates bases predominantly at the nitrogens (S_N2 type of reaction, in contrast to S_N1-type acting reagents alkylating to greater extent the oxygens) in DNA. As a result, about 82% of 7meG (non-mutagenic), 10.8% of 3meA and 0.6% of 3meG (mutagenic and highly toxic) and 0.3% of O^6meG (highly mutagenic) is formed [15]. The presence of 7meG and O^6meG in DNA does not arrest DNA synthesis (although O^6meG, with methyl group in *syn* conformations, may slow it down [16]); however, O^6meG (but not 7meG) may occasionally mispair with T, and lead to mutations of GC→AT transitions [17, 18]. 3meA and apurinic sites, which may be formed by excision of 3meA, 3meG or 7meG, stop synthesis of DNA and may, in a *umuDC*-dependent manner, induce AT→TA or GC→TA transversions [19-22].

So far, two constitutively expressed proteins, defending the cells against the chemical methylation of DNA, are known: (i) *ogt*-encoded O^6methylguanine-DNA methyltransferase, able to remove methyl, or (with a much lower efficiency) ethyl group from the O^6alkylG in DNA [23, 24], and (ii) *tag*-encoded 3-methyladenine DNA glycosylase I, excising 3-methyladenine (3meA) and leaving the apurinic sites in DNA [25]. O^6meG and apurinic sites may be excised by UvrABC-excinuclease known to remove many of the base adducts from DNA [26, 27]. The level of the UvrA and UvrB proteins much increases after SOS induction.

Alkylating agents may activate two inducible DNA repair systems: (i) the SOS system [28], including about 20 genes involved in DNA mutagenesis (*umuDC*) and repair, and (ii) the adaptive response, including four genes at least, triggered by low, weakly mutagenic, doses of methylating agents [29]. The cellular signal to induce the SOS system is damage to DNA and arrest of DNA synthesis [30, 31]. For the alkylating agents, the primary signal is the presence of 3meA (or 3etA) in DNA. Ability to induce expression from the *sfi::lacZ* fusion (*sfi* belongs to the SOS-regulated system) occurs at 5-50-fold lower doses of the alkylanes in bacteria devoid of both constitutive, *tagA*, and inducible *alkA*-encoded, 3meA DNA glycosylases, than in the wild type cells [32].

The primary signal for inducing the adaptive response is methylation of the Ada-protein at Cys-69. Ethylating agents are poor inducers of the adaptive response. Ada protein contains two (of the twelve) cysteines able to accept methyl groups: Cys-69, and Cys-321. Ada-me-Cys-69 (or Ada-me-Cys-69, me-Cys--321) protein is a potent positive inducer of the *ada* regulon, including its own and three other genes, at least: *alkA*-encoding 3meA DNA glycosylase II, and *alkB*- and *aidC*- encoding proteins with as yet unknown functions [33-36]. The 39 kDa Ada protein itself, and 19 kDa carboxyl-terminal fragment resulting from Ada proteolysis (with Ompt protein [37]), has the activity of a O^6alkyl-DNA alkyl transferase able to accept alkyl group from O^6-alkylguanine (and O^4-methylthymine) in DNA on their Cys-321 residue. Ability of *ada* to demethylate O^6meG is 7-fold higher than that to deethylate of O^6etG. Both, Ada protein, as well as its 19-kDa fragment, which structurally and functionally resembles of *ogt*-encoded protein, are not in *sensu stricto* enzymes, and after alkyl acceptation are not able to recover their activity again; i.e. they are

"suicide proteins".

Constitutively expressed 19-kDa Ogt and inducible, derived of Ada, 19 kDa O^6meG-DNA methyltransferase, differ in theirs affinities. Both, *ogt*- and *ada*-encoded alkyl transferases remove alkyl groups (with decreasing efficiency) from O^6meG, O^6etG and O^4meT in DNA; but the ability of Ogt protein to remove ethyl group from O^6etG is 145-fold greater, and to remove methyl from O^4meT, is 7-fold greater, than for Ada alkyltransferase [38].

Mammalian tissues contain non-inducible O^6meG--DNA methyltransferase, with its activity resembling of that of the bacterial Ogt protein [39]. Expression of Ogt protein in mammalian cells following transformation is frequently suppressed. Human cell lines devoid of O^6meG-DNA methyltransferase activity (Mer^- phenotype), are highly sensitive to killing effect by MNNG and MNU [40].

Regulated by methylated, or not methylated, Ada-protein [41] is *alkA*-encoded 31-kDa 3meA DNA glycosylase II. This glycosylase, is much less specific than 21-kDa 3meA glycosylase I, and apart from 3meA, excises a wide spectrum of modified nitrogen bases: 3meG, N^7meG [42], O^2meT, O^2meC [43], N^2, 3ethenoG [44], and possibly, 3,N^4ethenoC [45]. Recently, it has been found that 3-alkyl DNA glycosylase II may excise the oxidative-damaged pyrimidines: 5-formylU with the same rate as 3meA, and 5-hydroxymethylU with a rate of 1-3 order of magnitude slower [46]. Therefore the *alkA* gene may function in the repair of oxidative-damaged DNA lesions, occurring as a consequence of aerobic metabolism of cells, which in humans may be a cause of aging and cancer [47].

There is an interesting relationship between the *tag*- and *alk*-encoded 3meA DNA glycosylases. After induction of the Alk enzyme, the level of Tag enzyme decreases in the cells. Overproduction of *alk* DNA glycosylase in the wild type cells rendered them (most probably because formation of too many apurinic sites) highly sensitive to MMS. However, overproduction of *tag* DNA glycosylase, almost completely suppresses sensitivity to methylating agents of alk^- cells [48].

Recently, it has been found that *tag*-encoded enzyme may, albeit with 70-fold lower efficiency than *alk*--encoded enzyme, remove 3meG from DNA. This explains why an *alkA* mutation renders the cells much more sensitive to MMS then a *tag* mutation, and implies that 3meG is a highly cytotoxic lesion [49].

III. Mechanisms of the induction of the adaptive response

The mechanism of MMS-induced mutagenesis is closely related to the ability of the cells to repair premutagenic lesions. Induction of the adaptive response in *E. coli* appears to be a highly specific process. Expression of the Ada-protein, and proteins involved in the *ada* regulon, begins when Cys-69 become methylated. The methylating substrate is only one of the two stereoisomers (Sp) of methylphosphotriesters in DNA [36]. However, Cys-69 in Ada may be directly methylated by the methylating agents, and the adaptive response may be triggered without a contribution of methylated DNA [50]. Therefore one can guess that the ways of triggering the adaptive response directly, or *via* methyl phosphotriesters in DNA, should depend on the ability of the methylating agents to methylate phosphodiesters in DNA and/or Cys-69 in the Ada protein.

It has been also shown that O^6meG-DNA methyltransferases, both from mammalian and *E. coli* cells, may be inactivated by a direct action of the methylating agents [50-52]. O^6meG-DNA methyltransferases may be directly inactivated by a much lower concentration of MMS, that by MMS-methylated DNA. Among the methylating agents are those which inactivate O^6meG-DNA methyltransferase directly at a lower (S_N2 type reagent, including MMS) at a similar (MNNG), or at a much higher (MNU) concentration than by MMS-(MNNG or MNU) methylated DNA [50, 51].

Thus MMS is a direct inducer of the adaptive response, and an agent inactivating (or a substrate for?) O^6meG-DNA methyltransferase. One can conclude, therefore, that: (i) Cys-69 in Ada protein is a better acceptor for methyl group from MMS, than phosphotriester (and perhaps the nitrogen bases) in DNA and, (ii) the adaptive response in MMS-treated cells is induced earlier, before DNA can be damaged. That can explain the low mutagenic potency of MMS as compared to other alkylating agents. The enzymes involved in repair of premutagenic lesions, *ada*-encoded O^6meG-DNA methyltransferase and *alkA*-encoded 3meA-DNA glycosylase II, are induced ***before*** DNA becomes methylated, and the cells can remove MMS-induced lesions in DNA very efficiently. Accordingly to above, we did not observe an essential decrease in MMS-induced mutations, when cells were pretreated, before MMS mutagenic treatment, with the low doses of MNNG to induce the adaptive response (unpublished results and [53]).

The different reactivity of a methylating agent with DNA and with Cys-69 of Ada protein must be reflected in its mutagenic specificity and killing potency. Therefore, after MMS-treatment of *E. coli*, GC→AT transitions in DNA, due to efficient O^6meG repair are happened rarely, and induction of Alk enzyme leads to a great increase in the apurinic sites which, when UmuD′ UmuC proteins are available, may result in AT→TA [21, 22, 54] or GC→TA [19] transversions.

IV. Specificity of MMS-induced mutations

MMS is regarded as a low efficient, and largely

umuDC-dependent, mutagen, creating a low number of O^6meG in DNA. In *umuC*-defective cells, the frequency of MMS-induced mutations is 20-30% of that induced in *umuDC*-proficient cells [55-58]. However, when bacteria are deficient in *ogt*, or to an even greater extent, both in *ada* and *ogt* genes, the cells become hypermutable by MMS [59]. Furthermore, it has been calculated that in spite of S_N2 reactivity of MMS, the quantity of O^6meG in MMS-treated cells is astonishly high, and equals that obtained by isopropyl methanesulfonate treatment, a S_N1 acting reagent, with a high ability to alkylate bases at the oxygens [60].

Mutability of MMS is highly increased when bacteria are devoid of the mismatch repair system, and then the majority of the mutations are not *umuDC*-dependent [21, 61]. This may suggest that premutagenic lesions are formed in the nucleotide pools, then modified deoxynucleotides are incorporated into DNA and removed by action of the *dam*-instructed mismatch repair system [62, 63].

By analyzing of the $argE3_{(OC)} \rightarrow Arg^+$ revertants in *E. coli* AB1157 strains [64], it has been found that a fraction of MMS-induced, *umuDC*-dependent, Arg^+ revertants arise by supL(N) suppressor formation as a result of AT→TA transversions [21, 22]. This class of revertants is very low in $umuDC^-$ (or in *mutS*-devoid of the mismatch repair system, regardless of whether $umuDC^+$, or $umuDC^-$) strains. This was then confirmed by using bacteria with a deleted *umuDC* operon and transformed (or not) with plasmids bearing *umuD'C*, *umuDC* or *umuD* genes, respectively. The highest level of AT→TA transversions was observed in transformants bearing plasmid *umuD'C* genes; whereas no effect (as expected) was observed in bacteria bearing plasmid *umuD*. The level of MMS-induced mutations relies in practice on the accessibility of the UmuD' and UmuC proteins [22].

Obviously different lesions induced by MMS are leading to a different specificity of mutations. It is generally assumed that *umuDC*-dependent AT→TA transversions arise by apurinic sites formation (due to activity of 3meA DNA-glycosylase), and translesion (after dAMP insertion) replication in reaction with DNA polymerase III, UmuD', UmuC and RecA [20, 65, 66].

The high mutagenic potency of MMS, in an *ogt*, *ada* strain, testifies that O^6meG residues are formed in DNA, which in the wild type cells are quickly repaired. The great increase in mutation by MMS in the *mutS* strains may indicate that mutagenic modification occurs in high yields at bases in the nucleotide pools. The *mutSLH*, *dam*-dependent mismatch repair system corrects mismatches by excising a fragment containing a wrongly paired base from the newly synthesized strand of DNA [62, 63], hence in $mutS^+$ cells these mutations do not arise.

Increased sensitivity to MMS is shown by bacteria with mutations in *dam*, *recA* or *polA* genes, encoding proteins involved in repair of modified DNA. Inactivation of *polA* greatly decreases ability of cells, preadapted with a low level of MMS, to reactivate viability of MMS-treated bacteriophages [67].

V. Mutation frequency decline (MFD)

Earlier studies on UV-induced mutagenesis in *E. coli* B/r bacteria have shown, that certain mutations (at suppressor genes) did not arise when synthesis of proteins, immediately after irradiation, was transiently arrested. The effect known as mutation frequency decline (MFD), [68] was not observed in *mfd* mutant [69]. Recently, it was discovered that UV-induced lesions in DNA are preferentially repaired at the transcribed strand, and the repair requires transcription repair coupling factor (TRCF). It was subsequently shown, that TRCF and product of the *mfd* gene (*ERCC6* in humans) are the same [70, 71]. Mutated *mfd* gene from *E. coli* B/r *mfd* was transfered to AB1157 *E. coli* K-12 strain and it was shown that MFD may occur in a K-12 strain. Analysis of UV-induced mutations in *lacI* on F'*prolac* in *E. coli* AB1157 has revealed that mutations in the wild type strain arose predominantly from the lesions on the nontranscribed strain (A→G transitions) indicating that transcribing strand was preferentially repaired [72].

VI. Does MFD operate after MMS-treatment?

It has been found that a fraction of *umuDC*-dependent, MMS-induced mutations which occurs by AT→TA transversion at anticodon $tRNA^{lys}_{\dot{A}AA}$ is susceptible to the MFD effect, i.e. the level of those mutants falls rapidly, when the cells, immediately after MMS treatment, were incubated in minimal medium devoid of the amino acids required for bacterial growth [22]. Formation of supL(N) suppressor fulfills the condition for MFD. A modified adenine residue (most probable 3meA) is situated on the template (transcribed) strand of DNA. As it was also shown, by analysis of plasmid DNAs isolated from MMS-treated cells, apurinic sites were formed in DNA, and under condition of MFD, when synthesis of proteins was inhibited, were progressively repaired. Whether it is coincidental with MFD, or reflects MFD, remains to be elucidated.

VII. Conclusions

— There are many pathways in organisms allowing for mutation-avoidance and maintaining DNA stability. The great diversity of cellular responses to alkylating agents reflects not only their differences in chemistry of reaction with nucleic acid, but also their different potencies to induce (and to be repaired by) DNA repair system.

— The low mutagenic potency of MMS on *E. coli* is

result of its direct, and instant, ability to induce the adaptive response ***before*** DNA has been insulted. This allows for efficient repair of O^6meG, and rapid excision of methylated purines with formation of abasic sites in DNA. The majority of mutations then occur, most probably, as a result of translesion DNA replication, and dAMP insertion in a reaction with DNA polymerase III and UmuD′, UmuC, and RecA proteins. — Predominantly formed supL(N,G) suppressor mutations, monitored in *argE3*→Arg$^+$, MMS-induced revertants, undergo the MFD effect: a decrease in the suppressor mutations due to the preferential repair of the transcribed strand of DNA. This was supported by identification of the abasic sites in plasmid DNA isolated from MMS-treated bacteria and their subsequent repair, under conditions of MFD.

Acknowledgments

This work was partially supported by the State Committee for Scientific Research, Grant No. G-31.

References

1. Hoeijmakers JHJ, Bootsma D (1992) *Nature Genetics* **1**: 313-315
2. Hoeijmakers JHJ (1993) *Trends Genet* **9**: 173-177
3. Hoeijmakers JHJ (1993) *Trends Genet* **9**: 211-217
4. Jiricny J (1994) *Trends Genet* **10**: 164-168
5. Cleaver JE (1994) *Cell* **76**: 1-4
6. Hanawalt PC (1994) *Science* **266**: 1957-1958
7. Lindahl T (1994) *Curr Biol* **4**: 249-251
8. Koi M, Umar A, Chauhan DP, Cherian SP, Carethers JM, Kunkel TA, Boland CR (1994) *Cancer Res* **54**: 4308-4312
9. Hollstein M, Sindransky D, Vogelstein B, Harris CC (1991) *Science* **253**: 49-53
10. Soussi T, Legros Y, Lubin R, Ory K, Schlichtholz B (1994) *Int J Cancer* **57**: 1-9
11. Smith ML, Chen I-T, Zhan Q, Bae I, Chen C-Y, Gilmer TM, Kastan MB, O'Coonor PM, Fornace AJ Jr (1994) *Science* **266**: 1376-1380
12. Waga S, Hannon GJ, Beach D, Stillman B (1994) *Nature* (Lond) **369**: 574-578
13. Gottlieb TM, Jacson SP (1994) *Trends Biochem Sci* **19**: 500-503
14. Sancar A (1994) *Science* **266**: 1959-1960
15. Beranek P (1990) *Mutation Res* **231**: 11-30
16. Dosanjh MK, Loechler EJ, Singer B (1993) *Proc Natl Acad Sci USA* **90**: 3983-3987
17. Snow E, Foote RS, Mitra S (1984) *J Biol Chem* **259**: 8095-8100
18. Dosanjh MK, Singer B, Essigmann JM (1991) *Biochemistry* **30**: 7027-7033
19. Foster PL, Davis EF (1987) *Proc Natl Acad Sci USA* **84**: 2891-2895
20. Strauss B (1991) *BioEssays* **13**: 79-84
21. Sledziewska-Gójska E, Janion C (1989) *Mol Gen Genet* **216**: 126-131
22. Grzesiuk E, Janion C (1994) *Mol Gen Genet* **245**: 486-492
23. Potter PM, Wilkinson MC, Fitton J, Carr FJ, Brennand J, Cooper DP, Margison GP (1987) *Nucleic Acids Res* **17**: 9177-9193
24. Rebeck GW, Samson L (1991) *J Bacteriol* **173**: 2058-2076
25. Karran P, Lindahl T, Ofsteng I, Evensen GB, Seeberg E (1980) *J Mol Biol* **140**: 101-127
26. Voigt JM, Van Houten B, Sancar A, Topal MD (1989) *J Biol Chem* **264**: 5172-5176
27. Lin J-J, Sancar A (1989) *Biochemistry* **28**: 7979-7984
28. Radman M (1974) In: Prakash L, Sherman F, Miller M, Lawrence C, Tabor HW (eds) Molecular and Environmental Aspects of Mutagenesis, vol 3. Thomas Charles C Publisher, Springfield Ill pp 128-142
29. Samson I, Cairns J (1997) *Nature* (Lond) **267**: 281-283
30. Little JW, Mount DW (1982) *Cell* **29**: 11-22
31. Walker GC (1984) *Microbiol Rev* **48**: 60-83
32. Costa de Oliveira R, Laval J, Boiteux S (1986) *Mutation Res* **183**: 11-20
33. Teo I, Sedgwick B, Kilpatrick MW, McCarthy TV, Lindahl T (1986) *Cell* **45**: 315-324
34. Lindahl T, Sedgwick B (1988) *Annu Rev Biochem* **57**: 133-157
35. Volkert MR (1988) *Environm Molec Mutagen* **11**: 241-255
36. Shevell DE, Friedman BM, Walker GC (1990) *Mutation Res* **233**: 53-72
37. Sedgwick B (1989) *J Bacteriol* **171**: 2249-2251
38. Wilkinson MC, Potter PM, Cawkwell L, Georgiadis P, Patel D, Swann PF, Margison GO (1989) *Nucleic Acids Res* **17**: 8476-8484
39. Nakatsu Y, Hattori K, Hayakawa H, Shimizu K, Sekiguchi M (1993) *Mutation Res* **293**: 119-132
40. Karran P, Bignami M (1994) *BioEssays* **16**: 833-836
41. Nakabeppu Y, Sekiguchi M (1986) *Proc Natl Acad Sci USA* **83**: 6297-6301
42. Karran P, Hjelmgren T, Lindahl T (1982) *Nature* (Lond) **296**: 770-773
43. Mc Carthy TV, Karran P, Lindahl T (1984) *EMBO J* **3**: 545-550
44. Habraken Y, Carter CA, Sekiguchi M, Ludlum DB (1991) *Carcinogenesis* **12**: 1971-1973
45. Mroczkowska M, Kolasa K, Kuśmierek JT (1993) *Mutagenesis* **8**: 341-348
46. Bjelland S, Birkeland N-K, Bonneche T, Volden G, Seeberg E (1994) *J Biol Chem* **269**: 30489-30495
47. Ames BN, Gold LS (1991) *Mutation Res* **250**: 3-16
48. Kaasen I, Evensen G, Seeberg E (1986) *J Bacteriol* **168**: 642-647
49. Bjelland S, Bjoras M, Seeberg E (1993) *Nucleic Acids Res* **21**: 2045-2049
50. Takahashi K, Kwazoe Y, Sakumi K, Nakabeppu Y, Sekiguchi M (1988) *J Biol Chem* **263**: 13490-13492
51. Brent TP (1986) *Cancer Res* **46**: 2320-2323
52. Sledziewska-Gójska E (1995) *Mutation Res* **336**: 61-67
53. Sledziewska-Gójska E (1993) *Mutation Res* **294**: 1-8
54. Elespuru RK, Stupar LL, Gordon JA (1991) *Carcinogenesis* **7**: 1161-1167
55. Todd PA, Monti-Bragadin C, Glickman BW (1979) *Mutation Res* **62**: 227-233
56. Schendel PF, Defais M (1980) *Molec Gen Genet* **177**: 661-665
57. Schanabruch WG, Rein RP, Behlau I, Walker GC (1983) *J Bacteriol* **153**: 33-44
58. Langer P, Perry KL, Walker GC (1985) *Mutation Res* **150**: 147-158
59. Abril N, Roldan-Arjona T, Prieto-Almo M-J, van Zeeland AA, Pueyo C (1992) *Environm Molec Mutagen* **19**: 288-296
60. Eder E, Wiedenmann M, Deininger C, Kutt W (1990) *Toxic in Vitro* **3**: 167-174
61. Janion C, Plewako S, Bębenek K, Sledziewska-Gójska E (1989) *Mutation Res* **210**: 15-22
62. Radman M, Wagner R, Glickman BW, Meselson M (1980) In: Alcevic M (ed) Progress in Environmental Mutagenesis. Elsevier, North Holland Biomed Press, Amsterdam, pp 121-130
63. Modrich P (1991) *Annu Rev Genet* **25**: 229-253
64. Sledziewska-Gójska E, Grzesiuk E, Płachta A, Janion C (1992) *Mutagenesis* **7**: 41-46
65. Echols H, Goodman MF (1990) *Mutation Res* **236**: 301-311
66. Rajagopalan M, Lu C, Woodgate R, O'Donell M, Goodman MF, Echols H (1992) *Proc Natl Acad Sci USA* **89**: 10777-10781
67. Janion C (1982) *Mol Gen Genet* **186**: 419-426
68. Doudnay CO, Haas FL (1958) *Proc Natl Acad Sci USA* **44**: 390-401

69. George DL, Witkin EM (1974) *Mol Gen Genet* **133**: 283-291
70. Selby C, Sancar A (1994) *Microbiol Rev* **58**: 317-329
71. Witkin EM (1944) *BioEssays* **16**: 437-444
72. Oller AR, Fijałkowska I, Dunn RL, Schaaper RM (1992) *Proc Natl Acad Sci USA* **89**: 11036-11040

Secretion of proteins by *Escherichia coli* and its application in production of recombinant proteins

MAGDALENA M. FIKUS[1], BOŻENNA REMPOŁA[2]

Contents:

I. Introduction

In microorganisms, expression coupled to secretion of proteins seems an attractive possibility for production of recombinant proteins, rendering them easy to isolate and purify. Due to the recent progress in understanding of the mechanisms of natural secretion in *Escherichia coli*, its secretion systems have been exploited for the above purposes; various vectors and host strains have been constructed and examined.

The overall process of protein secretion includes their synthesis, membrane insertion, translocation, signal peptide recognition, its cleavage by signal peptidase and final localization of the mature protein. Most of the native secretory proteins synthesized by *E. coli* are transported to the periplasmic space [1]. Similarly, secretion of foreign proteins by *E. coli* occurs mostly to the periplasm; however, small recombinant polypeptides are exported outside the cell.

Export of proteins depends on the properties of the exported protein and on the export machinery of the host cell.

[1] To whom correspondence should be addressed. [1,2] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5a, 02-106 Warsaw, Poland.

II. Secretory proteins, what do they look like?

Outer membrane proteins and periplasmic proteins of *E. coli* are synthesized in the precursor form, with prepeptide N-extension of 20-25 amino acids called the signal peptide [2]. The signal peptide is essential for protein translocation across the inner membrane. It is believed that it also slows down preprotein folding.

Signal peptides in microorganisms do not show any sequence homology, they only display similar physico-chemical properties:

the N-terminal portion, 5 to 6 amino-acid residues, is of a basic character, i.e. it usually contains two positively charged amino acids,

these amino acids are followed by a long stretch of hydrophobic uncharged residues (10 to 15 amino acids with glycine in the center),

the C-terminal end contains 5 to 6 polar amino acids with alanine, glycine, threonine and/or serine in close vicinity of the cleavage site.

Preproteins with negatively charged (or uncharged) N-terminal part of a signal peptide accumulate in the cytoplasm, probably because of the lack of interaction with negatively charged C-terminus of the signal peptide or with negatively charged membrane phospholipids. The hydrophobic core helps in insertion of

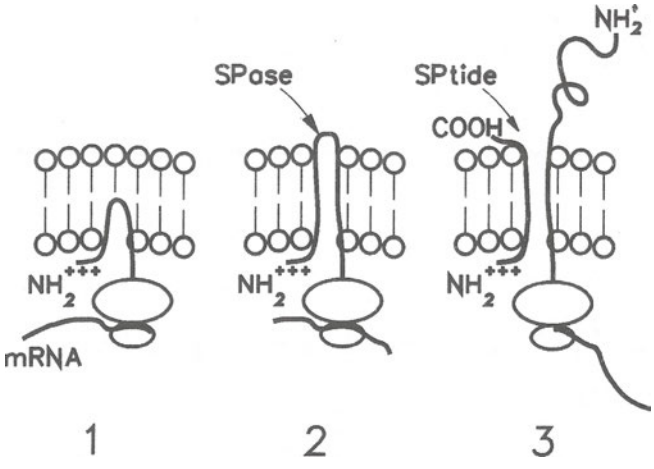


Fig. 1. "Loop model" showing an initial stage of secretion, where the preprotein sequence is pushed into the membrane with positively charged N-terminus oriented on the cytoplasmic side (1). Signal peptidase (SPase) cleaves the signal peptide (SPtide) on the outer side of the membrane (2) and the mature protein sequence is gradually exported across the membrane (3).

the prepeptide into the membrane; then the translocation process may be initiated. The hydrophobic amino acids may act as a recognition signal for membrane proteins (see below) which form the translocation channel [3].

It is assumed that preproteins are secreted across the membrane by forming a loop and a signal peptide is cleaved by a signal peptidase located in the membrane giving the mature secretory protein [4]. According to the "positive-inside rule" the hydrophobic part of the signal peptide should be aligned in the membrane in such a manner that its positively charged terminus is oriented towards the cytoplasm [5, 6] (Fig. 1).

It is generally believed that in prokaryotic secretion systems the net charge around the cleavage sequence should equal zero. A common feature for the composition of the region of about 20 amino acids at the N-terminus of the mature protein is that the presence of positively charged residues would prevent formation of the translocation loop [7-9]. On the other hand, it may be promoted by negatively charged amino-acid residues at position 2 of the N-terminal sequence of the mature protein [10]. However, examples are known when the replacement of positively charged amino-acid residues from the N-terminus of the mature protein did not promote secretion. It also seems that there is no promoting secretion consensus sequence in the mature part of the exported protein [11]. Before the problem of the role of the mature part of secreted proteins in translocation will be solved in more detail the lack of understanding is usually masked with vaguely described "translocation competent configuration" of the recombinant preprotein, needed for proper translocation [12]. At the moment it seems that the "Universal Signal Peptide" does not exist for *E. coli.*

III. Secretion, why is it so advantageous

The periplasm is enclosed between the inner and outer membranes of Gram-negative bacteria. Its volume in *E. coli* is estimated as about 30% of the total cell volume. The periplasm is filled with a peptidoglycan matrix with high concentration of bound water. The periplasmic proteins constitute about 5% of total cellular proteins.

The periplasmic space has a different proteolytic environment to that of the cytoplasm, and it is believed that here protein degradation is less pronounced than in the cytoplasm. On the secretory pathway the newly synthesized protein undergoes specific cleavage of the N-terminal signal sequence and in the oxidizing environment of the periplasm it folds, due to formation of disulphide bridges, to the native conformation. As the periplasmic space abounds in hydrolases the presence of ATP-dependent chaperones which could assist in proper folding of periplasmic proteins, native and recombinant, is highly improbable. However, the possible role of cellular chaperones in recombinant protein production is implicated by the fact that mutations in the σ^{32} subunit of RNA polymerase (*rpo*H gene), which takes part in transcription of stress proteins, increase the yield of secreted proteins [13].

The mature protein can be easily recovered by osmotic shock and readily purified from culture medium. Even toxic proteins can be produced in these systems.

When the use of the secretory pathway in production of recombinant proteins is sought, the nucleotide sequences coding for a signal peptide with its C--terminal specific amino-acid sequence, recognized and cleaved by a signal peptidase, is introduced into the expression cassette. Such a construction makes possible synthesis of recombinant proteins with any desired N-terminus.

IV. Host cell assistance in translocation across the inner membrane

The biosynthesis of proteins occurs in the cytoplasm, and when secretory preproteins are transported across it to the membrane, they remain in an unfolded state, chaperoned by proteins GroEL and SecB. Transport through the inner membrane is promoted by the signal sequence and the general export Sec system. It is believed that the main genes of *E. coli* (*secA*, *secB*, *secD*, *secE*, *secF*, *secG* and *secY*) involved in this pathway have already been discovered [14].

Protein SecB forms a complex with newly synthesized preprotein in the cytoplasm, preventing misfolding. This complex binds to the inner membrane through affinity binding to SecA, the peripheral subunit of *E. coli* preprotein translocase. SecA shows affinity for the SecYEG complex and for acidic phospholipids. It is suggested that the SecA protein also acts as an inhibitor of reverse translocation.

Binding of preprotein activates ATPase activity of SecA, which in turn drives the translocation of a short

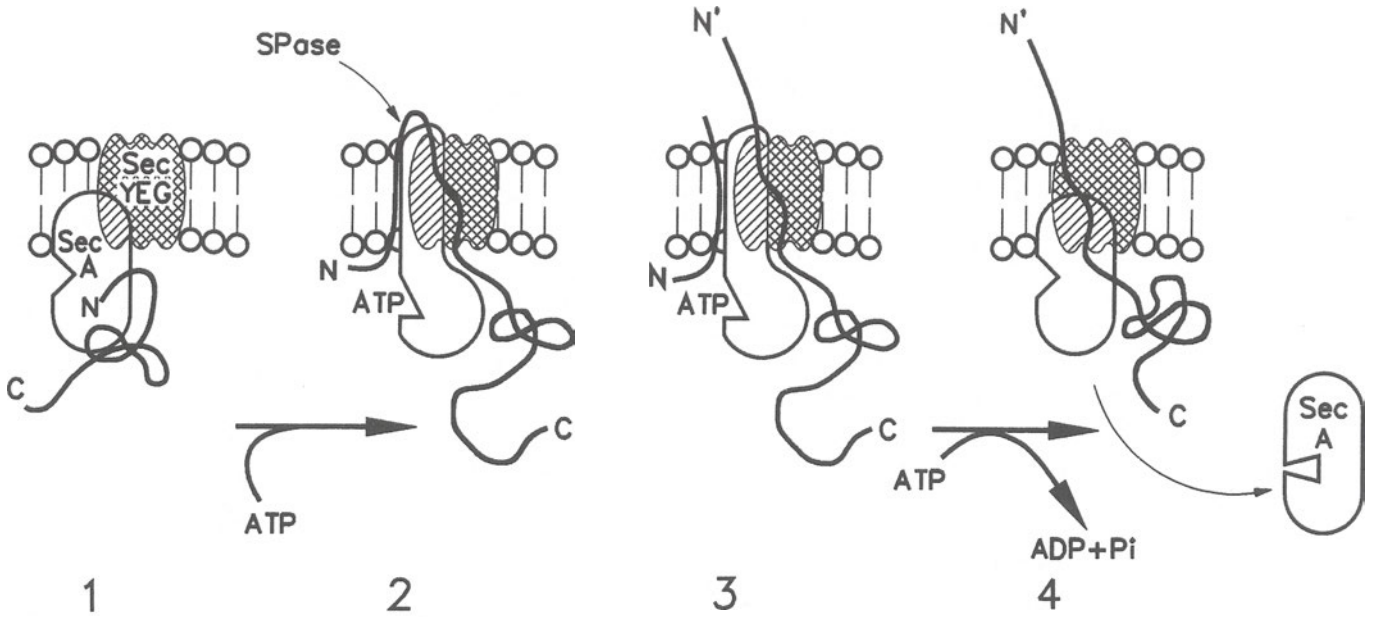


Fig. 2. Bacterial general export system based on interaction of Sec proteins. SecA protein has an affinity to the preprotein sequences and to the translocation membrane complex SesYEG (1). Preprotein activates SecA for the binding to SecYEG with accompanying ATP hydrolysis. When SecA binds ATP its conformation changes and it is driven together with the prepeptide across the membrane, forming a translocation channel with the SecYEG complex (2). The signal peptide is cleaved by signal peptidase (3), SecA uses a second ATP molecule to deinsert from the membrane (4).

segment of the preprotein into the membrane translocation channel formed by SecE, SecG and SecY proteins. Once a protein is translocated SecA detaches from the membrane [14].

Products of *secD* and *secF* genes regulate the channel assembly-disassembly reaction *via* the SecA insertion-deinsertion step. The final translocation step across the inner membrane of the preprotein free of SecA depends on the proton motive force [12, 15, 16] (Fig. 2).

SecD, an inner membrane protein, acts as a chaperone and prevents side-to-folding-reactions; it also assists in the release of the mature protein from the inner membrane.

From the point of coupling ATP energy to the preprotein movement across the membrane, mechanisms of translocation were solved in distinct ways in eukaryotes and prokaryotes [16], whereas numerous proteins involved in this process, from archaebacteria to mammals, are highly homologous [17].

V. Proteolysis: maturation and degradation of newly synthesized secretory preproteins

Low concentration of proteolytic enzymes in the periplasm is one of the strongest arguments in favour of production of recombinant proteins in secretion systems. However, secretion reduces but does not abolish proteolysis of recombinant proteins, and proteolysis remains an important factor lowering the yield of recombinant secretory proteins.

Known *E. coli* proteases assisting the secretory pathway are compartmentalized: signal peptidases I and II, protease DegP, IV and V were localized within the inner membrane, protease III was found in the periplasm, and proteases V and VI (ompT) in the outer membrane [18].

The level of expression of the signal peptidase can influence the export and maturation of secretory proteins [19].

It was recently shown that mutations in all genetic loci known to affect stability of proteins in the cytoplasm also affect the stability (yield) of heterologous, periplasmic proteins [13]. Examination of single, double and triple mutants in genes coding for proteolytic enzymes proved that none of the proteases in question was essential for strain growth. This result points to the fact that there are some undissolved envelope proteases that are essential for growth.

VI. Protein folding

Due to the reduced mobility of proteins in the periplasm folding of exported proteins is under specific control.

Experimental data prove that folding is the rate limiting step in accumulation of native *E. coli* proteins in the periplasm. The major determinant of the extent of correct folding is the rate of folding as opposed to the rate fo aggregation. *In vivo* folding is assisted by a large set of proteins playing catalytic and structural roles in the process. Since the existence of ATP in the periplasm is highly improbable, there are doubts about the existence of ATP-dependent periplasmic chaperones.

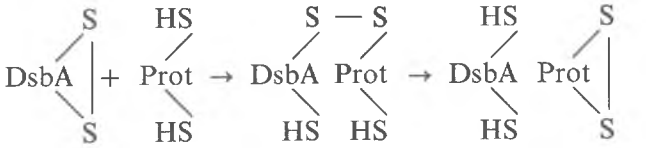
In the case of recombinant proteins prepared for secretion, even if some periplasmic proteins perform the task of chaperones for them, they do not accelerate folding, but bind to recombinant proteins inhibiting

premature folding, misfolding and aggregation.

In vitro formation and rearrangement of disulphide bonds, and the *cis-trans* isomerization of prolyl peptide bonds (the main processes directing folding) are slow chemical reactions, taking hours and even days. Indubitably *in vivo* both should be assisted by enzymes.

Two *E. coli* prolyl *cis-trans* isomerases (PPI A and B) were discovered. The product of the *rot* gene, PPI A, was localized in the periplasm [20]. Coexpression of the genes coding for fragments of antibodies, for PPI A, and for DsbA (see below) did not influence the rate of folding of the recombinant antibody nor its final yield [21, 22].

Formation of disulphide bonds in the periplasm of *E. coli* is assisted by proteins, products of *dsbA*, *dsbB* and *dsbC* genes [23]. The key catalyst is DsbA protein, acting according to the general scheme:



Careful examination of catalytic properties of DsbA suggests its role *in vivo* as an oxidizing agent. A recently discovered product of the *dsbC* gene can substitute for DsbA protein. The protein shows disulphide oxidoreductase activity *in vitro*, and is more active than DsbA in disulphide-rearrangement reactions, acting as disulphide isomerase. DsbB, the integral inner membrane protein, probably takes part in the reoxidation of DsbA (Fig. 3). Both DsbA and DsbB proteins are essential for proper folding of recombinant proteins in *E. coli* [24].

It has still not been unequivocally established if the regulation of the rate of disulphide bond formation is important for the efficiency of secretion [25]. Cocloning of a gene coding for secretory protein with *dsbA* in the presence of GSH/GSSG increased the yield of the investigated protein [26].

Increase of the expression — secretion yield was observed when two genes, both with their own promoters, namely *rpo*H (coding for σ^{32} factor), and *dsbA* were included and overexpressed in one vector [27].

VII. Secretion vectors for *E. coli*

The secretory pathway in *E. coli* could be exploited for production of recombinant proteins only when appropriate expression — secretion vectors were constructed. Presently, various secretion systems exist, created for particular research purposes, in which for instance the signal sequence is coded by the pre-region of the *phoA* gene (coding for alkaline phosphatase) [28, 29], or the nucleotide sequence coding for the signal peptide of the OmpT protein is included [30].

The most widely used secretion vectors are those in which the nucleotide sequence coding for the signal peptide of native *E. coli* OmpA protein was inserted.

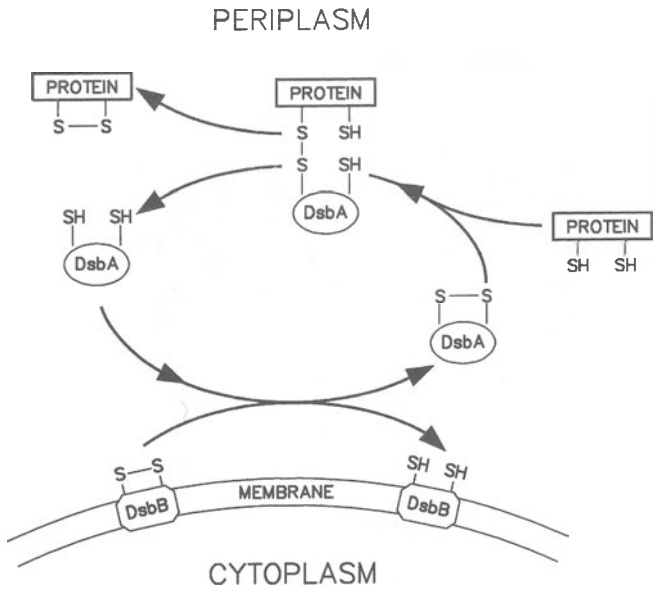


Fig. 3. The general scheme for formation of S-S bridges in proteins to be folded in the periplasmic space of *E. coli*, with participation of DsbA and DsbB enzymes.

The sequence of 21 amino acids of this signal peptide displays all the features required by the loop model. A group of pIN-OmpA plasmids has been constructed, applied as vectors and generoulsy offered to the scientific community by Dr. M. Inouye [3].

In the pIN-OmpA vectors expression is regulated by a constitutive lpp^p promoter and *lac*UV5 inducible promoter operator fragment (lac^{po}). The cloned gene may be inserted into the downstream polylinker sites. Translation termination codons, and a rho-independent transcription termination signal are included. These vectors have been successfully applied for gene cloning, expression and secretion of many different proteins of bacterial and eukaryotic origin.

VIII. Factors affecting a yield of properly folded secretory recombinant proteins

Usually, recombinant polypeptides are produced in higher concentrations in the cytoplasm than when secreted. However, as explained earlier, the secretory system for many sound reasons may become an attractive alternative to the cytoplasmic ones. Those of the recombinant proteins that are secreted by their native host cells (hormones, serum proteins) are easily accepted by the secretory machinery of *E. coli*. Construction of a secretory cassette successfully coding for typical cytoplasmic proteins still remains as non routine task.

Expression of recombined genes and secretion of recombinant proteins could be increased by co-expression of genes involved in protein folding and functions of the secretory machinery of the cell. In addition to the previous examples, the yield of mature recombinant hIL-6 increased 4.3-fold when additional copies of *secY* and *secE* were supplied to the cell; the ratio of mature to preprotein forms increased from 1.2 to 10.8 [31].

The amount of signal peptidase does not seem to limit a secretion yield [32]. However, when the expression yield is high, some secreted polypeptides still bear an uncleaved signal peptide at the N-terminus [13, 30, 31].

General conditions may [33] or may not [34] affect the yield, depending on the host strain used. The various external factors such as nutritional elements, reducing agents, temperature of growth, a culture phase, concentration of the inducer, may play a role in the yield and rate of secretion. Once again the "protein conformation competent for secretion" may be invoked [34].

The size of the secretory peptide is obviously important: it has to be long enough for the translocation loop formation. The conclusion from many experiments is that at least, 60 amino acids' long preprotein may enter the export pathway and be served by the general Sec-dependent system [35].

Shorter peptides, if exported, probably bypass the central export machinery [17]. The shortest peptides produced and secreted in *E. coli* are the scorpion insectotoxin, mature form has 35 amino acids, four disulphide bridges [36], β endorphine — 31 amino acids [37], potato carboxypeptidase inhibitor — 39 amino acids, three disulphide bridges [38], and plant serine protease inhibitor — 29 amino acids, three disulphide bridges [39], all found exclusively in a culture medium.

The ultimate localization of the peptide will depend on its size, sequence and structure.

Various prokaryotic and eukaryotic proteins have been produced and secreted in *E. coli*: enzymes and their inhibitors, cytokines, hormones, recombinant antibodies [3, 9, 12, 21, 22, 25, 27, 31, 34, 39, 40-45]. Unfortunately, various authors apply different criteria for estimating of the expression yield, and these renders difficult direct comparison of their results. In general, very high expression levels (several percent of the total cell protein) resulted in periplasmic aggregation of the produced proteins [22, 43]. Under proper conditions in the most successful experiments secreted proteins were obtained in a soluble state at levels close to 30% of the total periplasmic protein, which represents several dozen miligrams of protein per 1 litre of culture [22, 44].

IX. Conclusions

Expression and secretion of the recombinant proteins in *E. coli* depend on numerous factors such as signal peptide and mature proteins sequence and structure, type of elements controlling expression and secretion, vector and *E. coli* strain employed, physiological conditions of the experiment. Secretion is the system of choice when the microbial host for production of a recombinant protein in its native folded state, including formation of disulphide bridges, is preferred.

References

1. Holland IB, Mackman N, Nicuad J-M (1986) *Bio/Technology* **4**: 427-431
2. Izard JW, Kendall DA (1994) *Mol Microbiol* **13**: 765-773
3. Duffaud GD, March PE, Inouye M (1987) In: Wu R, Grossman L (eds) Methods in Enzymology, vol 153. Academic Press, San Diego, New York, Berkeley, Boston, London, Sydney, Tokyo, Toronto, pp 492-507
4. DiRienzo JM, Nakamura K, Inouye M (1978) *Annu Rev Biochem* **47**: 481-532
5. von Heijne G (1992) *J Mol Biol* **225**: 487-494
6. Boyd D, Beckwith J (1990) *Cell* **62**: 1031-1033
7. Geller B, Zhu H-Y, Cheng S, Kuhn A, Dalbey RE (1993) *J Biol Chem* **268**: 9442-9447
8. Li P, Beckwith J, Inouye H (1988) *Proc Natl Acad Sci USA* **85**: 7685-7689
9. Groch N, Hahn U, Heinemann U (1993) *Gene* **124**: 99-103
10. von Heijne G (1986) *J Mol Biol* **192**: 287-290
11. MacIntyre S, Henning U (1990) *Biochimie* **72**: 157-167
12. Cheah K-C, Harrison S, King R, Crocker L, Wells JRE, Robins A (1994) *Gene* **138**: 9-15
13. Meerman HJ, Georgiou G (1994) *Bio/Technology* **12**: 1107-1110
14. Kim YJ, Rajapandi T, Oliver D (1994) *Cell* **78**: 845-853
15. Economou A, Wickner W (1994) *Cell* **78**: 835-843
16. Wickner WT (1994) *Science* **266**: 1197-1198
17. MacIntyre S, Henning U (1990) *Biochemie* **72**: 157-167
18. Enfors S-O (1992) *Trends Biotechnol* **10**: 310-314
19. van Dijl JM, Jong A, Smith H, Bron S, Benema G (1991) *Mol Gen Genet* **227**: 40-48
20. Liu J, Walsh C (1990) *Proc Natl Acad Sci USA* **87**: 4028-4032
21. Knappik A, Krebber C, Pluckthun A (1993) *Bio/Technology* **11**: 77-83
22. Beck E, Burtscher H (1994) *Protein Expression and Purification* **5**: 192-197
23. Bardwell JCA (1994) *Mol Microbiol* **14**: 199-205
24. Bardwell JCA, Lee J-O, Jander G, Martin N, Belin D, Beckwith J (1993) *Proc Natl Acad Sci USA* **90**: 1038-1042
25. Wulfing C, Pluckthun A (1994) *Mol Microbiol* **12**: 685-692
26. Wunderlich M, Glockshuber R (1993) *J Biol Chem* **268**: 24547-24550
27. Wulfing C, Pluckthun A (1994) *J Mol Biol* **242**: 655-669
28. Inouye H, Michaelis S, Wright A, Beckwith J (1981) *J Bacteriol* **146**: 668-675
29. Weiss E, Orfanoudakis G (1994) *J Biotechnol* **33**: 43-53
30. Teixeira AV, Dowdle EBD, Botes DP (1993) *Biochim Biophys Acta* **1217**: 16-22
31. Perez-Perez J, Marquez G, Barbero J-L, Gutierrez J (1994) *Bio/Technology* **12**: 178-180
32. Denefle P, Kovarik S, Ciore T, Gosselet N, Benichou J-C, Latt M, Guinet F, Ryter A, Mayaux J-F (1989) *Gene* **85**: 499-510
33. Hussain K, Bowler C, Roberts LM, Lord JM (1989) *FEBS Lett* **244**: 383-387
34. Halfman G, Brailly H, Bernadac A, Montero-Julian FA, Lazdunski C, Baty D (1993) *J Gen Microbiol* **139**: 2465-2473
35. von Heijne G (1994) *FEBS Lett* **346**: 69-72
36. Pang S-Z (1992) *Gene* **116**: 165-172
37. Nagahari K, Kanaya S, Munakata K, Aoyagi Y, Mizushima S (1986) *EMBO J* **41**: 3589-3592
38. Molina MA, Aviles FX, Querol E (1992) *Gene* **116**: 129-138
39. Rempoła B, Wilusz T, Markiewicz W, Fikus M (1995) *Acta Biochim Polon* **42**: 109-114
40. Auerswald EA, Rossler D, Mentele R, Assfalg-Machleidt I (1993) *FEBS Lett* **321**: 93-97
41. Quaas R, McKeown Y, Stanssens P, Frank R, Blocker H, Hahn U (1988) *Eur J Biochem* **173**: 617-622

42. Parker K C, Wiley D C (1989) *Gene* **83**: 117-124
43. Ghrayeb J, Kimura H, Takahara M, Hsiung H, Masui Y, Inouye M (1984) *EMBO J* **3**: 2437-2442
44. Taguchi S, Yoshida Y, Matsumoto K, Momose H (1993) *Appl Microbiol Biotechnol* **39**: 732-737
45. Becker G W, Hsiung H M (1986) *FEBS Lett* **204**: 145-150

Synthetic virus-like gene transfer system

STANISŁAW SZALA*

Abbreviation used: NLS — nuclear localization signal.

In a broad sense gene therapy relies on the insertion of a functioning gene into cells to correct an error of the metabolism or to provide the cells with a new function.

For several cancers, genetic intervention based on substitutive gene therapy of many recessive and, to some extent also dominant genetic disorders, does not appear feasible. The current strategies are based on the elimination of tumor cells rather than on the reversion of their malignant phenotype (for the latest review see [1]).

Direct elimination of cancer cells capitalizes on the insertion of suicide genes that may selectively kill transduced tumor cells by activating autotoxic mechanisms.

Indirect elimination of a tumor exploits the insertion of an immunostimulatory gene into cancer cells in order to stimulate the host antitumor immune response.

Several methods of DNA transfer have been developed [2]. Some of them, based upon virally mediated gene transfer, are currently used in *ex vivo* and *in vitro* gene transfer. Since virally mediated gene transfer does not yet allow cell targeting, it is not suitable for *in vivo* gene transfer approaches. For this reason, the search for other vectors allowing cell specific targeting has gained considerable significance.

Specific cancer cell targeting can be achieved by two different approaches: transductional and transcriptional targeting [3]. Transductional approach utilizes differences in the expression of cell surface proteins (e.g. different cell surface receptors) to target cancer cells. A specific ligand or monoclonal antibody is bound to the complex of DNA molecules/polycation polymer to be delivered. Targeting is thus achieved by specific binding of the vectors to cognate cell surface molecules.

The transcriptional targeting approach relies on the observation that tumors are capable of differentially and specifically expressing certain genes (e.g. melanoma expresses tyrosinase, gastrointestinal tumors — carcinoembryonic antigen, hepatoma — α-fetoprotein, etc.), and thus exploits this distinctive pattern of gene regulation. In this regard, therapeutic genes under the control of promoters specific for these tumor marker genes offer a potential mode for specific targeting by restricting expression to these tumor cell targets. If the inserted gene encodes a nonmammalian enzyme, which can convert a nontoxic prodrug to a toxic metabolite, then all transfected tumor cells are selectively killed ("suicide genes' strategy") whereas normal transfected cells remain alive (see for review: [4]).

The aim of this short review is to emphasize the potential of transductional targeting which employs cellular mechanisms of receptor-mediated endocytosis of macromolecules and viruses for the transfer of DNA molecules into cells.

The functional design of molecular conjugate vectors resembles natural viruses: they imitate the entry of viruses into the cell but are devoid of viral gene elements as well as potential safety hazards resulting from the presence of such viral gene sequences (see for review: [5]).

The first step in the cellular pathways of macromolecular transport includes binding of macromolecules or viruses (ligands) to specific cell-surface receptors and their internalization through coated vesicles into endosomal compartments. In general, a vector developed to deliver DNA *via* the receptor-mediated endocytosis pathway (called a molecular conjugate vector) possesses, at least, two distinct functional domains: a polycationic domain for binding DNA and a ligand domain, which binds to a particular cell surface receptor. The polycationic domain is chemically linked to the ligand [6, 7]. DNA charge neutralization by polycations results in condensation of DNA in the form of a compact structure. The complexes (DNA/polycations/ligand) interact with cell surface receptors *via* its ligand domain (a cognate domain). This interaction allows the linked DNA to be transported into the cell vesicle system.

In practice, the main problem in this type of gene targeting is that gene transfer *via* the corresponding internalization pathway does not always result in effective expression of transferred heterologous DNA. The inefficient gene transfer is not related to initial binding and internalization steps but rather reflects some events occurring after complex localization wi-

* Department of Tumor Biology, Institute of Oncology, 44-100 Gliwice, Poland

thin the endosome. It seems that the conjugate DNA complex may be entrapped within cellular endosomes and thus transported DNA cannot reach the nucleus.

The internalized DNA could escape entering cellular pathways if the conjugate complex included an endosome escape mechanism such as inactivated replication-incompetent adenovirus or endosome-disruptive protein [8, 9]. The escape from the cell vesicle system depends on viral capsid proteins. The penton base and/or peripentonal capsid components appear to be crucial capsid proteins involved in endosome disruption [5]. The N-terminal of the influenza virus hemagglutinin subunit HA-2 contains a membrane-active peptide sequence that, at neutral pH, is prevented from adopting an α-helical conformation. Upon lowering the pH to 5-6 (i.e. endosomal pH) the charges are neutralized by protonation, allowing a transition to an α-helical amphipathic structure, which is able to interact with and destabilize lipid membranes [9].

The active transport of molecular conjugate vectors to the nucleus is achieved by the presence of nuclear localization signals (NLS) [10]. The vector capitalizes on the active transport of nuclear proteins through the nuclear pore.

For the nuclear transport to occur several cytosolic factors are required (see also Fig. 1) [11]. The specific receptor (importin) recognizes NLS-bearing proteins and directs them to the binding site at the pore. The activity of the receptor is stimulated by its interaction with another protein (p97). Once the receptor--NLS-bearing protein complex is bound at the pore, a complex of other proteins (TC4 and B-2) stimulates translocation. After translocation to the nucleus the complex is dissociated and released DNA is transcribed by the transcription machinery.

In fact, the synthetic multifunctional vector utilizes two distinct basic mechanisms of transporting macromolecules into and within the cell: the receptor-mediated endocytosis pathway and the active transport to the nucleus. The functional design of conjugate vectors that exploits these efficient cellular and nuclear entry mechanisms should include at least four distinct domains:

1) domain for binding and packaging of DNA,
2) domain for receptor binding,
3) domain for disruption of endosome,
4) domain for nuclear localization signals (NLS) (for details and examples see: Tab. 1).

In theory, the system which is based on polycations/DNA core particles is flexible enough: it should allow construction of a vector capable of specific transduction targeting. Using monoclonal antibodies [12], glycoconjugates [13] or lectins [14] bound to polycations such a system can be targeted to the cell of interest.

Self-aggregation and/or intermixing of DNA with polycations or amphiphiles having covalently bound

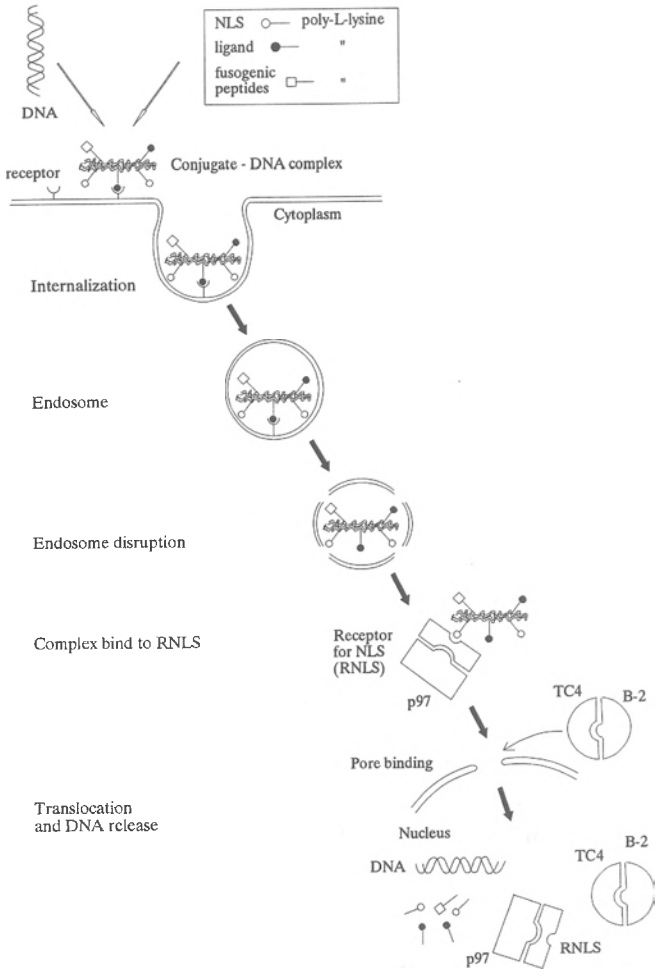


Fig. 1. Targeted gene transfer based on two distinct basic mechanisms of transporting macromolecules into a cell: receptor-mediated endocytosis and active transport to the nucleus. A multifunctional molecular conjugate is emoloyed to bind DNA and transport it *via* a macromolecular transport mechanism. A vector consists of a DNA-binding domain (poly-L-lysine) which is covalently bound to a ligand for its cell surface receptor. All other domains (for disruption of endosomes and for nuclear localization signals, NLS) are linked to other molecules of poly-L-lysine. When the ligand domain is bound by its corresponding cell surface receptor, the conjugate is internalized. The endosome is disrupted by the presence of the domain that comprises fusogenic peptides. The active transport of the conjugate to the nucleus is achieved by binding of the complex to cytosolic receptors that recognize the nuclear localization signals (NLS) included in conjugate vectors.

peptides (fusogenic, NLS) as well as different ligands should produce synthetic vectors with marked plasticity, mimicking natural viral vectors [9, 15].

It is also structurally possible to construct a vector that is, in fact, a fusion protein composed of 3-4 domains [10, 16]. In this case, the cognate domain might be derived from a single-chain monoclonal antibody to a defined cell surface antigen. All other domains may be derived from amino-acid sequences of transcription factors (a domain for DNA binding), exotoxin A from *Pseudomonas* (endosome disrupting domain), sequences from histone H1 or from a large T antigen of SV40 virus for transport to the nucleus (NLS sequences) (see: Table 1).

In such synthetic vectors, the size and shape of

Table 1.
Basic domains of virus-like particles

Domain	Function	Examples
DNA binding	DNA condensation	Cationic polymers (e.g. poly-L-lysine, histone H1, lipopolyamine, etc.); DNA binding sequence of transcription factors (e.g. yeast GAL4).
Receptor binding	cell targeting	Ligands: growth factors (e.g. insulin, etc.); transporting protein (e.g. transferrin, etc.); carbohydrate (e.g. galactose, glucose, mannose, mannose-6-phosphate, *N*-acetylgalactosamine, etc.); lectins (e.g. wheat germ agglutinin, peanut agglutinin, concanavalin A, etc.); monoclonal antibodies (e.g. monoclonal against antigen erb-B2, receptor growth factors, etc.).
Endosomolytic components	cytoplasm entry	Inactivated viruses (e.g. adenovirus, etc.); fusogenic peptides (e.g. peptides from influenza virus, etc); bacterial toxin (e.g. Pseudomonas exotoxin A, listeriolysin O, etc.).
Nuclear localization signals (NLS)	nucleus entry	Sequence present in several nuclear polypeptides (e.g. histone H1, SV40 large antigen, etc.).

condensed molecules of DNA are most important factors related somehow to the efficiency of gene transfer. Perales *et al.* [17] have clearly shown that the controlled condensation of DNA molecules with, for instance, poly-L-lysine in certain concentration of NaCl, results in complexes of defined size and shape (10-12 nm in diameter). This compact structure of complexes is probably an important factor in many steps of gene transfer *in vivo* such as DNA stability while in the circulation, receptor-ligand interaction, endocytosis, etc., as well as for long and persistent expression of the transgene (several months) in targeted cells.

Gene targeted delivery *via* receptor-mediated endocytosis pathways can be effective for transferring suicide genes into cell deposit of primary tumors or to the cells that metastased several different organs. The injectable vector with the tumor specific promoter (transcriptional targeting) can be, in fact, additionally targeted and expressed only in the tumor cells (this twice targetable vector is important to avoid unnecessary killing of normal cells).

The low efficiency of nonviral gene transfer (in the 10-20% range of all tumor cells) as compared to the gene transfer efficiency of natural viral vectors (in the 20-90% range) should be theoretically compensated by 1-10% of the so-called "bystander effect" of suicide genes (1-10% bystander effect means that expression of a suicide gene in 1-10% of tumor cells leads to the death of virtually all bystander cells after treatment with adequate prodrug) [18].

Targeted systems can also be used for *in vivo* transfer of antisense oligonucleotides as well as DNA fragments up to 48 kb [19, 20].

There are two major limitations of the above system at present: so far its lower efficiency of gene transfer than in other systems (natural viral vectors), and transient expression of the transferred gene. Nevertheless, the system is more targetable and safer than the natural viral vectors.

Finally, one would like to stress that the real progress in tumor gene therapy will depend mainly on the developments in the area of targeted gene delivery (vehicles) rather than on the "DNA-drug" itself.

The new vector design (virus-like vectors) could arise from a novel formulation of existing molecules or be based on a different principle [21, 22]. Nevertheless, different designs should be consistent with the concept of the targetable, injectable gene transfer vector. In other words: the new vector deliverable *in vivo* by systemic route should be "stealthed" from humoral factors and should be capable of delivering genes efficiently into a cell type of interest.

But, frankly, it would be cautious to remember a famous statement of H. L. Mencken: "To every human problem there is a solution that is simple, neat, and wrong!"

Acknowledgment

I thank my colleagues W. Widłak and A. Sochanik for help in preparation of the manuscript. This work was supported by a grant from the State Committee for Scientific Research 4 PO5A.028.08.

References

1. Herrmann F (1995) *J Mol Med* **73**: 157-163
2. Mulligan RC (1993) *Science* (Washington) **260**: 926-932
3. Schmidt MJ, Rousculp MD, Goldsmith KT, Curiel DT, Garver RI Jr (1994) *Human Gene Therapy* **5**: 29-35
4. Blaese RM, Ishii-Morita H, Mullen C, Ramsey J, Ram Z, Oldfield E, Culver K (1994) *Eur J Cancer* **8**: 1190-1193
5. Michael SF, Curiel DT (1994) *Gene Therapy* **1**: 223-232
6. Wu GT, Wu CH (1987) *J Biol Chem* **262**: 4429-4432
7. Wagner E, Zenke M, Cotten M, Beug H, Birnstiel ML (1990) *Proc Natl Acad Sci USA* **87**: 3410-3414
8. Curiel DT, Agarwal S, Wagner E, Cotten M (1991) *Proc Natl Acad Sci USA* **88**: 8850-8854

9. Wagner E, Plank C, Zatloukal K, Cotten M, Birnstiel M L (1992) *Proc Natl Acad Sci USA* **89**: 7934-7938
10. Chen J C, Stickles R J, Daichendt K A (1994) *Human Gene Therapy* **5**: 429-435
11. Powers M A, Forbes D J (1994) *Cell* **79**: 931-934
12. Waldmann (1991) *Science* (Washington) **252**: 1657-1662
13. Monsigny M, Roche A-C, Midoux P, Mayer R (1994) *Adv Drug Delivery Rev* **14**: 1-24
14. Batra R K, Wang-Johanning F, Wagner E, Garver R I Jr, Curiel D T (1994) *Gene Therapy* **1**: 255-260
15. Behr J-P (1994) *Bioconjugate Chem* **5**: 382-389
16. Fominaya J, Wels W (1995) *3rd Symposium on Gene Therapy. Berlin-Buch April* 20-22, Abstract P22
17. Perales J C, Ferkol T, Beegen H, Rattnoff O D, Hanson R W (1994) *Proc Natl Acad Sci USA* **91**: 4086-4090
18. Sorscher E J, Peng S, Bebok Z, Allan P W, Bennett L L Jr, Parker W B (1994) *Gene Therapy* **1**: 233-238
19. Freese J Jr, Wu C H, Wu G Y (1994) *Adv Drug Deliv Rev* **14**: 137-152
20. Cotten M, Wagner E, Zatloukal K, Phillips S, Curiel D T, Birnstiel M L (1992) *Proc Natl Acad Sci USA* **89**: 6094-6098
21. Hodgson C P (1995) *Bio/Technology* **13**: 222-225
22. Hodgson C P (1995) *Exp Opin Ther Patents* **5**: 1-10

Scope and limitations of boron neutron capture therapy as the selective tool against cancer

HANNA WÓJTOWICZ[1],
HALINA THIEL-PAWLICKA[2],
KRZYSZTOF GOLANKIEWICZ[3]

Contents:

Abbreviations used: BBB — blood-brain-barrier; BNCT — boron neutron capture therapy; BOPP — di-potassium tetrakis-carborane-carboxylate esters of 2,4-bis-(a,b-dihydroxyethyl)-deuteroporphyrin IX; BPA — *p*-boronophenylalanine; BSH — sodium borocaptate; BTPP — boronated tetraphenylporphyrin; CBA — carboranylalanine; CBU — 2′-2′-*O*-(1,2-carboran-1-ylmethyl)uridine; DBDU — 5-dihydroxyboryl-2′-deoxyuridine; DCP-AES — direct current plasma-atomic emission spectroscopy; ICP-AES — inductive coupled-atomic emission spectroscopy, LET — linear energy transfer; PCR — polymerase chain reaction; PLD — potentially lethal damage; SLD — sublethal damage; SPECT — single photon emission computer tomography; TU — 2-thiouracil; VCDP — boronated hematoporphyrin.

I. Introduction

Boron compounds have been widely used in medicine for decades [1]. Once in a while in public media there appear the rumours about their harmful side effects. It is difficult to evaluate in the present commercialised world how many of these rumours are based on the acquired additional true knowledge and how many are the results of ruthless competition between pharmaceutical companies.

In this paper we would like to outline the possibility of selective destruction of cancer cells. This possibility involves the boron containing chemical compounds and thermal neutrons. It is called Boron Neutron Capture Therapy (BNCT) and was suggested [2] long before the technical development enabled the first experiments [3]. Nevertheless, technical complexity of the application of this unique method for destroying the malignant cells was responsible for the slow progress in this field. During the last decade we observe the increasing interest in this method thanks to the rapid development of the nuclear engineering. Already several excellent reviews covering the field appeared in the world literature [4-6].

II. Nuclear properties of boron

Boron consists of two nonradioactive isotopes, boron-10 and boron-11 in the ratio 81.6% to 18.5%.

[1,2,3] Uniwersytet im. A. Mickiewicza, Wydział Chemii, Zakład Syntezy i Struktury Związków Organicznych, 60-780 Poznań, ul. Grunwaldzka 6, Poland

Boron-10 displays the very useful ability to capture thermal neutrons and to produce a localised nuclear reaction [7] according to the simplified equation:

$$^{10}B + {}^{1}n \rightarrow [^{11}B] \rightarrow {}^{4}He + {}^{7}Li + \gamma + 2.79 \text{ MeV}$$

The products of this nuclear reaction except the γ-photons are heavy nonradioactive particles and their kinetic energy transfer takes place on a very short path length. The rate of linear energy transfer (LET) of these particles is very high and the immense energy of this reaction is dissipated in a very small volume. In average, the range of the alpha particles released in a tissue is about 10 μm, what is about equal to the diameter of a single cell. Therefore the lethal effect is confined only to 1-2 cells containing boron-10.

III. The requirements for the successful use of BNCT

BNCT can be regarded as a binary system in which each constituent is non lethal to the malignant cell but their combination is lethal. Moreover, the sublethal damage (SLD) and potentially lethal damage (PLD) caused by BNCT are not repairable [8]. It is worthy to note that SLD and PLD caused by gamma photons and X-rays may be repaired [9, 10]. There are three factors which limited the broad application of BNCT in clinical practice:

1). Neutron source.
2). Boron-10 carriers to the tumor site.
3). The precise determination of boron atoms in malignant cells, blood and normal tissues.

ad 1). Currently BNCT is dependent on nuclear reactors as a source of neutrons. The type of neutron beam extracted from the reactor depends on moderators or filters. With no moderators the reactors produce the fast neutron beams with energies greater than 10,000 eV. These neutrons cause the biological damage through ionisation and the results can be compared with cobalt-60 therapy. So, the fast neutrons are useless in BNCT. Thermal neutrons with energies 0.025 eV can be obtained by filtering fast neutrons through scattering media such as heavy or light water or graphite.

The first BNCT experiments with the application of thermal neutrons gave dubious results [5]. W.H. Sweet at Massachusetts General Hospital in Boston in the late 1950s did not observe the enhancement of survival times of his patients with brain tumor after employment of BNCT. Sodium borate ($Na_2B_4O_7$) was used as boron-10 carrier. In 1968 Sweet's co-worker Hiroshi Hatanaka returned to Japan and began to continue BNCT in Tokyo using thermal neutron beams of low intensity and ^{10}B-enriched $Na_2B_{12}H_{11}SH$ as a boron carrier. Hatanaka obtained remarkable results. Among his patients were several long term survivors. Although Hatanaha's experiments were far from ideal and stirred some criticism in medical community concerning consistency of data collection and procedures, he and his former boss Sweet can be regarded as pioneers of BNCT for brain tumors. The strong limitation in using thermal neutrons in BNCT is their low energy and rapid attenuation in tissues. In the result they do not penetrate the depth required for BNCT without severe irradiation of surface tissues [11]. The epithermal neutron beams with energy range 1-10,000 eV provide better penetration of the tissues than thermal neutrons and produce thermal neutrons at the desired depth in tissue due to moderating effects of the tissue itself. Such neutrons are relatively non destructive, seldom being captured by H or N atoms. They can be produced by using moderators to slow down fast neutrons [12]. The epithermal neutron beams are now undergoing careful and broad investigations on their radiobiological effects [13].

The greatest limitation to the vast use of BNCT is its dependence on nuclear reactors as a source of neutrons. In recent years several different means of generating neutrons were considered including californium-252 [14], spallation sources [15], and low-energy accelerators [16]. At present it seems technically feasible to obtain epithermal neutron beams suitable for BNCT from other sources than nuclear reactors. It certainly broaden the application of BNCT to clinical practice.

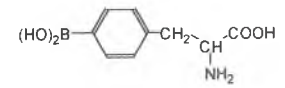
ad 2). The key to the successful development of BNCT was and remains the availability of nontoxic boron target compounds that can selectively accumulate in tumor cells. The discouraging effects obtained by Farr and Sweet [3] in their pioneer experiments with sodium borate, resulted in part probably from the fact that although the concentration of boron in tumor was three times that found in normal brain, the concentration in blood was even higher. Systematic studies for tumor-selective boron compounds were started by A.H. Soloway and his group in early 1960s [17]. Of the first importance was finding the reliable method which would allow the selection of the investigated compounds. It was achieved by using mice bearing a subcutaneously transplanted ependymoblastoma [18]. As the result of these studies sodium borocaptate $Na_2B_{12}H_{11}SH$ known as BSH appeared the most promising. It has been used since 1968 by H. Hatanaka in Japan as a boron carrier for BNCT with remarkable result [19]. It stimulated the development of boron chemistry and renewed the interest in BNCT in USA and Europe. The mechanism of action of BSH is still unknown. BSH is easily oxidised to the corresponding disulfide $B_{12}H_{11}SSB_{12}H_{11}{}^{4-}$ and its oxide $B_{12}H_{11}SS(O)B_{12}H_{11}$ [20, 21]. It has been found that the disulfide derived from BSH is taken up by tumor cells even better than BSH itself [22]. Recent experiments suggest that BSH accumulates extracellularly and that radiation damage that causes tumor death is extracellular [23]. The experimental facts accumulated with the use of BSH

for BNCT show the complexity of the problem connected with the delivery of a proper amount of boron-10 to the tumor site. The search for ^{10}B rich compounds is rapidly expanding over the vast areas of inorganic and organic chemistry. In this article only few examples will be presented for the obvious reasons.

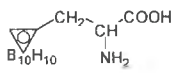
The first amino acid-boron analogues were borobetaine™, Me_3NBH_2COOH, and boroglycine™, H_3NBH_2COOH, prepared by Spielvogel [24, 25]. Spielvogel obtained a large number of boron containing amino acids, peptides, boron analogues of phosphonates, heterocyclic amine derivatives, nucleosides, and borono phosphate oligonucleotides [26]. Many of these compounds exhibited valuable biological properties. Boron analogues of amino acids and peptides are soluble in water as well as in most organic solvents. This amphiphilic character may be useful for the transport of these compounds across cell membranes. Moreover, they can cross cell membranes without a hydrolysis of the amide bond. The profound effect of the replacement of carbon by boron atom is on the pK_a value of the free amino acids. Boron analogue of glycine has a carboxyl group pK_a even 8.3, compared to 2.4 for glycine. It certainly has the effect on the hydrogen bonding properties. These properties may be responsible for the fact that borobetaine and boroglycine did not appear to localize in the Harding-Passey melanoma [27, 28]. Although they cross blood--brain-barrier (BBB) there is evident their rapid tumor excretion. The thorough toxicological studies of these type boron analogues have shown no or very little toxicity. Some of them were also screened for use in BNCT with very promising results. B-10 enriched *p*-boronophenylalanine (BPA) is just entering intensive clinical trials in the USA. Simultaneously the extensive studies have been performed to synthesize other amino acids containing multiple boron atoms. The interest was focused on so called carborane-containing amino acids. Carborane cage consisting of 10 boron atoms is only slightly larger than the space occupied by a benzene ring and two carbon atoms in the cage participate in delocalized bonding. Taking this into account, carboranylalanine (CBA) appears to be a boron analogue of phenylalanine. Some examples of boron containing amino acids are given in the figure 1. Boronated amino acids or as components of peptides can be considered as potential boron carriers for BNCT. Nevertheless, at the present state of knowledge, for each boron-containing amino acid or peptide, there are required detailed studies to determine their *in vivo* stability and tumor selectivity.

The chemistry of boron containing nucleic acids precursors began with the synthesis of 5-dihydroxyboryl uracil [29] and 5-dihydroxyboryl-2′-deoxyuridine (DBDU) [30]. Since then several boron derivatives of nucleic acid's bases and nucleosides have been synthesized [31]. A special effort was taken for the synthesis of carborane substituted ribose and base components of nucleosides because each molecule of carborane derivative contains 10 boron atoms. Some characteristic examples of this type of compounds are presented in the figure 2. 2′-*O*-(1,2-carboran-1-yl-methyl)uridine (CBU-2′) [32] and its 3′5′ isomers were efficiently taken up by F 98 glioma cells *in vitro* [33] but the concentration was higher in the cytoplasm than in the cell nucleus. The synthetical achievements of these type of compounds highly exceeded the possibility of their physiological evaluation which is now under way in the leading laboratories throughout the world.

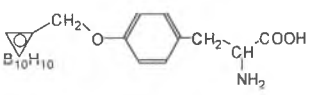
Spielvogel prepared cyanoborane adducts of 2′-deoxynucleosides by an exchange reaction of tri-



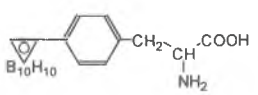
p-boronophenylalanine
BPA



carboranylalanine
CBA

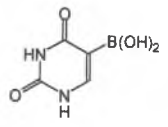


O-(o-carboran-1-yl-methyl)-tyrosine
CBT

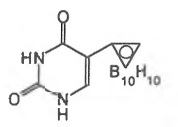


p(o-carboran-1-yl)-phenylalanine
CBPA

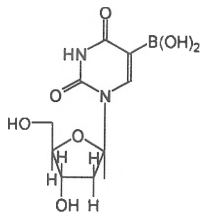
Fig. 1.



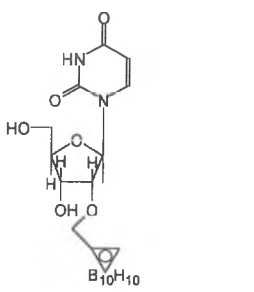
5-(dihydroxyboryl)uracil



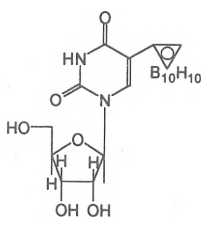
5-carboranyluracil



5-dihydroxyboryl-2′-deoxyuridine
(DBDU)

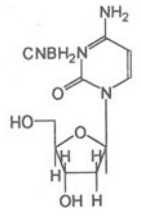


2′-O-(1,2-carboran-1-yl-methyl)uridine
(CBU-2′)

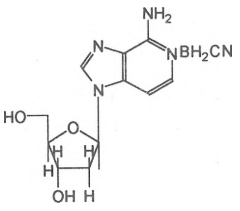


5-carboranyluridine

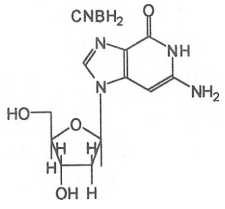
Fig. 2.



2'-deoxycytidine-N^3-cyanoborane



2'-deoxyadenine-N^1-cyanoborane



2'-deoxyguanosine-N^7-cyanoborane

Fig. 3.

phenylphosphine-cyanoborane (Ph_3PBH_2CN) with 3',5'-*O*-protected nucleosides [34] (Fig. 3).

These borane adducts of 2'-deoxynucleosides triphosphates can be incorporated into DNA during the polymerase chain reaction (PCR) and this may simplify PCR sequencing [35]. These boronated nucleosides are hydrolytically stable and preliminary results show that some of them possess antitumor activity. Yet, it remains to be tested whether they will localise in the neoplasm and will be useful for BNCT.

In 1982 it was found that 2-thiouracil (TU) can be incorporated into tumor in the amount as high as 300 μg per g of tumor. This is equivalent to 26 μg B per g of tumor if only one boron atom was incorporated in each molecule of TU [36]. This finding has stimulated the synthesis of numerous boron derivatives of TU. Some examples are presented in the figure 4.

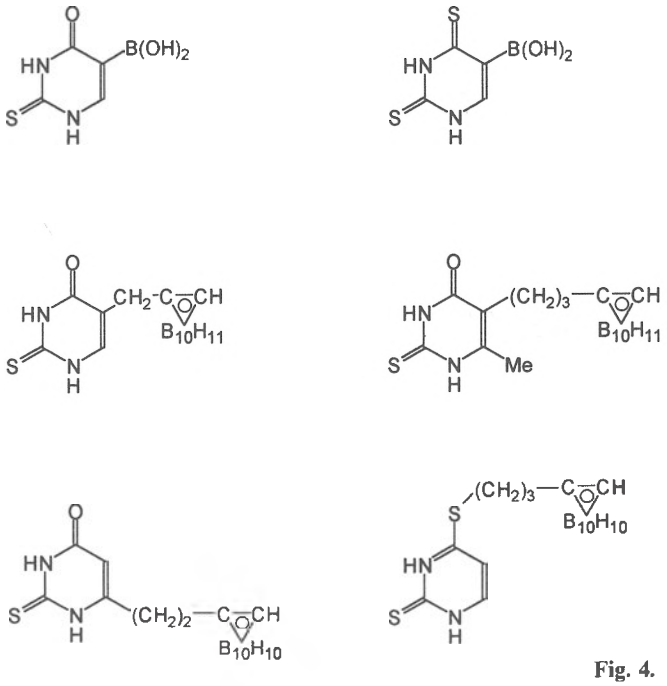
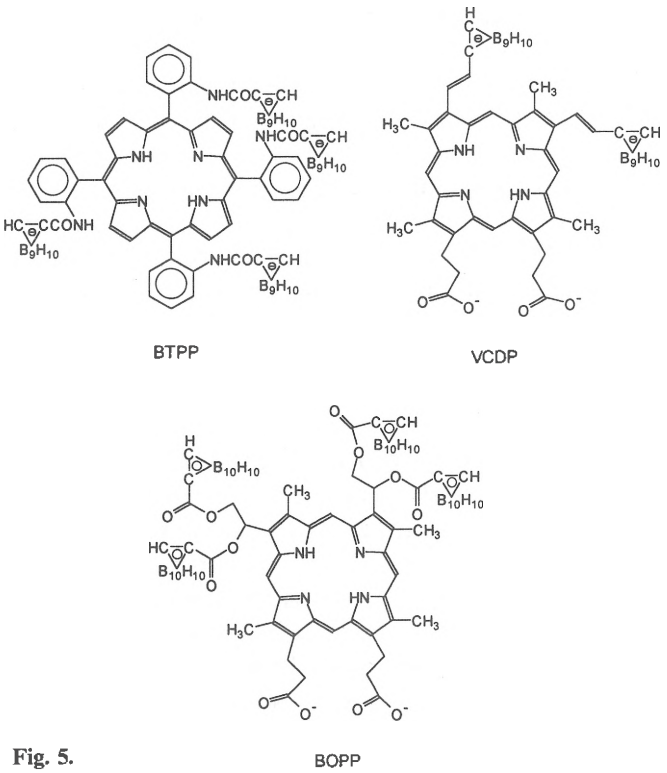


Fig. 4.

The evaluation of the selective uptake of this type of compounds by murine melanoma models showed that 5-borohydroxy TU was the most effective species giving boron concentration 30 μg B per g of tumor with tumor-to-blood boron concentrations ratios 4:1 to 7:1 [37]. The carborane derivatives of TU appeared to be too hydrophobic and were too slowly taken up by a tumor. In general the efficacy of boron containing derivatives of 2-thiouracil still remains to be proved but already obtained results are very promising.



BTPP VCDP BOPP

Fig. 5.

The propensity of porphyrin and related heterocyclic compounds to localize in the solid tumors has been utilized as the means of delivery of boron to tumor cells. The proper carborane derivatives based on tetraphenylporphyrin (BTPP) [38], hematoporphyrin (VCDP) and (BOPP) [39] were synthesized and investigated as the boron carriers for BNCT. These compounds are presented in figure 5. Above shown compounds were evaluated with a variety of murine tumor models. It appeared that BOPP provided the highest boron uptake by tumor combined with the long retention times. BOPP is localized in glioma at concentrations greater than 20 μg B per g of tumor, while the amount of boron required to achieve tumor necrosis by BNCT is about 2-3 μg B per g of tumor. The boron carriers for BNCT have to pass extremely rigid physiological tests before being accepted for clinical investigations. It is the reason that from the numerous already synthesized boron compounds only few are now being tested clinically as the boron agents for BNCT. They are shown in the figure 6.

ad 3). The third basic factor for the successful use of BNCT is the precise knowledge of tumor's location,

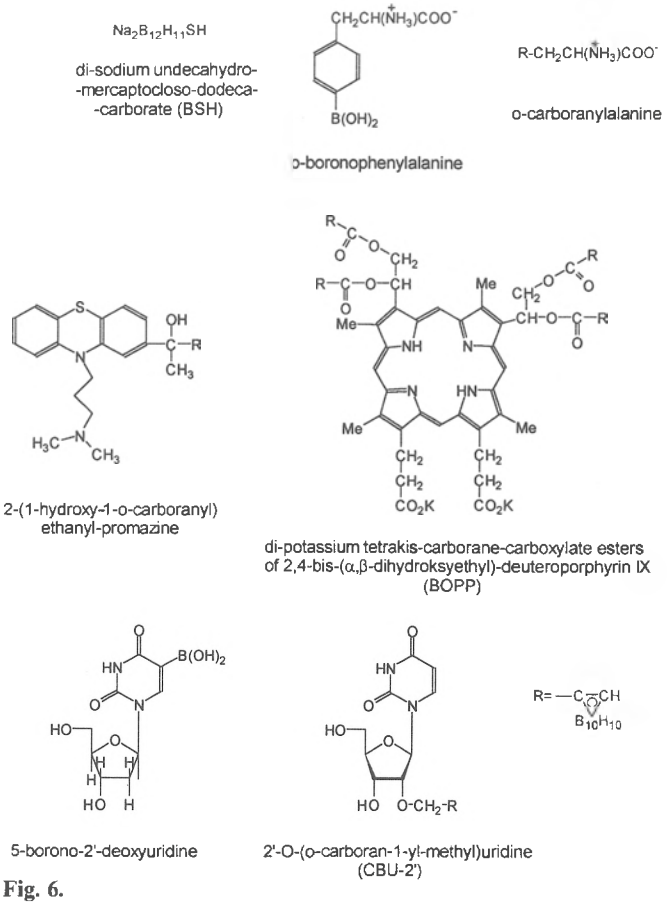


Fig. 6.

and the concentration of ^{10}B in tumor, blood and normal tissue. Of the first importance it is not only to know the actual concentration of ^{10}B in tumor but also a rate of its excretion. Since the first experiments run by Soloway *et al.* [18] the enormous progress has been attained in this field. Below, the main directions of the research are exemplified.

The weighed samples taken at a biopsy are dissolved by oxidative procedures and boron analysis is accomplished by inductive coupled or direct current plasma--atomic emission spectroscopy (ICP-AES) [40] and (DCP-AES) [41]. These methods can be extended to microscopic tissue samples of 1 mg obtained by needle biopsy. Recently, as a noninvasive estimation of ^{10}B concentrations in tumor ^{10}B magnetic resonance imaging (MRI) has been employed. This method already has proved to be very promising [42] and opening the possibility for rapid analysis involving ^{10}B carrier as the signal generating species. Another noninvasive method involves radiolabelling of ^{10}B compound coupled with single photon emission computer tomography (SPECT) to visualize the compound in tumor [43].

IV. Conclusions

It is a dream of all oncologists to destroy malignant cells one by one without doing any harm to the surrounding normal tissues. Potentially BNCT offers the means to make this dream real. For about 40 years this method has had its ups and downs but the steady hard earned progress has been observed. The final goal is worthy to combine the mutual efforts of oncologists, chemists, biologists, physicists and many scientists of different but connected specialities to eliminate danger of death by specific kind of tumor, up to now incurable.

Acknowledgment
This work was supported by KBN grant No. 023/P05/95/08

References

1. Kliegel W (1980) In: Boron in Biology, Medicine and Pharmacy. Springer-Verlag, New York
2. Locher GL (1936) *Am J Roentgenol Radium Ther* **36**: 1-13
3. Farr LE, Sweet WH, Robertson JS, Foster CG, Locksley HB, Sutherland DL, Mendelsohn ML, Stickley EF (1954) *Am J Roentgenol Radium Ther Nucl Med* **71**: 279
4. Barth RF, Soloway AH, Fairchild RG, Burger RM (1992) *Cancer* **70**: 2995-3007
5. Hawthorne MF (1993) *Angew Chem Int Ed* **32**: 950-984
6. Goudgaon NM, El-Kattan GF, Schinazi RF (1994) *Nucleosides & Nucleotides* **13**: 849-880
7. Taylor HJ (1935) *Proc Roy Soc London Ser A* **147**: 873
8. Utsumi H, Ichihashi M, Kobayashi T, Elkind MR (1989) *Pigment Cell Res* **2**: 337-342
9. Rofstad E (1986) *Acta Radiol* **25**: 1-10
10. Leibel SA, Sheline GE (1987) *J Neurosurg* **66**: 1-22
11. Mishima Y, Ichihashi M, Hatta S, Honda C, Yamamura K, Nakagawa T (1989) *Pigment Cell Res* **2**: 226-234
12. Brugger RM, Less T (1986) *Niigata, Japan: Nishimura Co Ltd* 110-116
13. Moss RL, Stecher RA, Asmussen F, Huiskamp R, Devit L, Mijnheer B (1992) In: Allen BG, Moore D, Harrington B (eds) Progress in Neutron Capture Therapy for Cancer. Plenum Press, New York, pp 7-12
14. Maruyama Y (1991) *Nuclear Science Applications* **4**: 29-43
15. Grusell E, Conde H, Larsson B, Ronnqvist T, Sornsuntisook D, Crawford J (1990) In: Haley OK, Bernard JA, Zamenhof RG (eds) Neutron Beam Design, Development and Performance for Neutron Capture Therapy. Plenum Press, New York, pp 249-258
16. Wang CK, Blue TE, Gahbauer R (1990) *Nuclear Technol* **84**: 93-104
17. Soloway AH, Wright RL, Messer JR (1961) *J Farmacol Exp Ther* **134**: 117-122
18. Entzian W, Soloway AH, Raju R, Sweet WH, Brownell GL (1966) *Acta Radiol* **5**: 95-100
19. Hatanaka H (1975) *J Neurol* **209**: 81-94
20. Wellum GR, Tolpin EI, Soloway AH, Kaczmarczyk A (1977) *Inorg Chem* **16**: 2120
21. Tolpin EI, Wellum GR, Berley SA (1978) *Inorg Chem* **17**: 2867
22. Slatkin DN, Joel DD, Fairchild RG, Micca PL, Nawrocky KM, Laster BH, Coderre JA, Finkel CC, Poletti CE, Sweet WH (1989) *Basic Life Sci* **50**: 179
23. Fairchild RG, Kahl SB, Laster BH, Kalef-Ezra J, Popenoe EA (1990) *Cancer Res* **50**: 4860-4865
24. Spielvogel BF, Wojnowich L, Das MK, McPhail AT, Hargrave KD (1976) *J Am Chem Soc* **98**: 5702-5703
25. Spielvogel BF, Das MK, McPhail AT, Onan KD, Hall IH (1980) *J Am Chem Soc* **102**: 6343-6344
26. Spielvogel BF (1994) *Phosphorus, Sulfur and Silicon* **87**: 267-276
27. Hall IH, Gilbert CJ, McPhail AT, Morse KW, Hasset K, Spielvogel BF (1985) *J Pharm Sci* **74**: 755-758
28. Spielvogel BF, Sood A, Hall IH, Fairchild RG, Micca PL (1989) *Oncology* **165**: 123-125
29. Liao TK, Podrebarac EC, Cheng CC (1964) *J Am Chem Soc* **86**: 1869
30. Schinazi RF, Prusoff WH (1978) *Tetrahedron Lett*: 4981
31. Wyzlic IM, Tjarks W, Soloway AH, Anisuzzaman AKM, Feng-Guang Rong, Barth RF (1994) *Int*

J Radiat Oncol Biol Phys **28**: 1203-1213
32. Anisuzzaman AKM, Alam F, Soloway AH (1990) *Polyhedron* **9**: 891
33. Liu L, Barth RF, Soloway AH, Anisuzzaman AKM, Alam F, Tjarks W, Zha XH, Morrison GH (1991) *Proc Am Assoc Cancer Res* **34**: 2418
34. Sood A, Spielvogel BF, Shaw BR (1989) *J Am Chem Soc* **111**: 9234-9235
35. Porter K, Briley D, Huang F, Sood A, Spielvogel BF, Shaw BR (1993) *Genome Sequencing and Analysis Conference V*, Hilton Head Island, SC, October 23, Abstracts
36. Fairchild RG, Packer S, Greenberg D, Prantika S, Brill AB, Fand I, McNally WP (1982) *Cancer Res* **42**: 5126-5132
37. Tjarks W, Gabel D (1991) *J Med Chem* **34**: 315-319
38. Kahl SB, Joel DD, Finkel GC, Micca PL, Nawrocky MM, Coderre JA, Slatkin DN (1989) *Basic Life Sci* **50**: 325
39. Kahl SB, Koo MS (1992) In: Allen BJ, Moore DE, Harrington BV (eds) Proc Fourth Int Symp Neutron Capture Therapy of Cancer. Plenum, New York
40. Tamat SR, Moore DE, Allen BJ (1987) *Anal Chem* **59**: 2161-2164
41. Barth RF, Adams DM, Soloway AH, Mechetner EB, Alam F, Anisuzzaman AKM (1991) *Anal Chem* **63**: 890-893
42. Kabalka GW, Davis M, Bendel P (1988) *Magn Reson Med* **8**: 231
43. Paxton RJ, Beatty BG, Hawthorne MF, Varadarajan A, Wiliams LE, Curtis FL, Knobler CB, Beatty JD, Shively JE (1991) *Proc Natl Acad Sci USA* **88**: 3387-3391

Mammalian DNA polymerases

JANUSZ A. SIEDLECKI[1], RADOSŁAWA NOWAK[2]

Contents:

Abbreviations used: CRE — cATP response element; CS — Cockayne's syndrome; CSA, CSB — CS group A or B complementing; ERCC — excision repair cross-complementing rodent repair deficiency; MNNG — *N*-methyl-*N*-nitro-*N*-nitrosoguanidine; PCNA — proliferating cell nuclear antigen; RF — replication factor; XP — xeroderma pigmentosum; XPA...XPG—XP group A.....G complementing.

I. Introduction

DNA synthesis is the most important process during the cell cycle. It takes place not only during the replication of genomic and organelle DNAs but also is important for maintenance of genome integrity during DNA repair and recombination. All these processes are catalyzed by multiprotein complexes. Some of the proteins participating in DNA replication and repair complexes are known, their detail function remains, however, uncertain. The key part of these complexes falls to DNA polymerases.

Five species of DNA polymerases termed α, β, γ, δ and ε have been purified from mammalian cells. Background information on these enzymes has been thoroughly reviewed [1-7]. In this article we would like to summarize the current status of the mammalian DNA polymerases.

II. Subunit composition

The subunits' composition of mammalian DNA-dependent DNA polymerases is presented in table 1. Except of DNA polymerase β, all remaining DNA polymerases are high molecular mass enzymes.

DNA polymerase α was the first isolated mammalian DNA polymerase [8]. So far the enzyme was isolated from a wide range of sources (see review [1, 2, 5, 6]). It is composed of a core catalytic subunit of approximately 160-185 kDa, primase subunit containing two polypeptides (60 kDa and 50 kDa), and of 70-kDa polypeptide, which function is to couple the DNA polymerase α/primase complex to the cellular initiation/elongation machinery [9].

DNA polymerase β is the smallest DNA polymerase (see review [1, 2, 4]). The enzyme is a single polypeptide with 335 residues of molecular mass 39 kDa with no associated exonuclease or proofreading activity. The

[1,2] Department of Molecular Biology, The Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, P. Findera 101, 02-784 Warszawa, Poland

Table 1.
Mammalian DNA polymerases

	DNA polymerase				
	α	δ	ε	β	γ
Previous names	α	δ, δ, δI	δ, δ, δII, δ*	β	mt
Protein structure					
native form	> 250	170	256	38	160-300
catalytic core	165	125	215	38	45 or 125-160
other subunits	70, 60, 50	48	55	lack	35, 47
Organelles localization	nucleus	nucleus	nucleus	nucleus	mitochondrium
Associated activities					
3′ → 5′ exonuclease	−	+	+	−	+
primase	+	−	−	−	−
Properties					
processivity	+ +	+	+ + +	+	+ + +
processivity with PCNA	NE	+ + +	NE	NE	NE
fidelity	+ + +	+ + +	+ + +	+	+ + +
Inhibitors					
aphidicolin	strong	strong	strong	lack	lack
N-ethylmaleimid	strong	strong	strong	lack	strong
Butylphenyl-dGTP	strong	100 × weaker	100 × weaker	lack	lack
dideoxy NTP	lack	lack	lack	strong	weak
Suggested biological role					
genomic DNA replication	yes	yes	yes	yes/?	no
organelles DNA replication	no	no	no	no	yes
repair	yes	yes/?	yes/?	yes	no/?
recombination	unknown	unknown	yes/?	yes	unknown

NE — no effect

protein is composed of two distinctive domains of 8-kDa N-terminal and 31-kDa C-terminal, which are connected by a protease hypersensitive region [10, 11]. The 8-kDa fragment has template-binding activity and the 31-kDa fragment contributes nucleotidyl transfer activity [11]. The hydrodynamic properties show that the two domains appear folded onto each other and this folding may be a prerequisite to create enzyme-substrate complex [12]. Crystal structure reveals that 31-kDa-domain has three subdomains (fingers, palm and thumb) arranged in U shape reminiscent of other polymerase structures [13, 14]. Structural and mutagenesis studies of DNA polymerase β, Klenow fragment of *E. coli* polymerase I, HIV-1 reverse transcriptase and bacteriophage T7 RNA polymerase shown a common nucleotidyl transfer mechanism [15].

DNA polymerase δ was discovered in 1976 as a unique DNA polymerase containing 3′ → 5′ proofreading exonuclease activity. The enzyme was first isolated from erythroid hyperplastic bone marrow of rabbit [16]. Since then it has been purified from a variety of sources (see review [1, 2, 5, 6]). The enzyme contains two subunits, a 125-kDa catalytic polypeptide and an associated 48-kDa polypeptide of unknown function.

The biggest DNA polymerase, first discovered as so-called PCNA-independent DNA polymerase δ is now called DNA polymerase ε (see review [1, 6, 7]). The data strongly suggest that the DNA polymerase ε is a single 258-kDa polypeptide. This polypeptide contains a protease-sensitive region that makes it cleavable into two parts, one of which is catalytically active [17-19]. This shorter form of DNA polymerase ε, named DNA polymerase ε*, is able to replicate a primed M13 DNA efficiently. Moreover, there are suggestions that this proteolysis might not be merely an artefact of isolation and both forms of DNA polymerase ε could have some different biological functions [7].

DNA polymerase γ was first called mitochondrial DNA polymerase. The enzyme was isolated from several eukaryotic organisms (see review [1, 2]). There are substantial differences in the subunits' composition of the enzyme from different organisms. In a chick embryo the enzyme is a tetramer of identical subunits about 47 kDa each. DNA polymerase γ isolated from *Drosophila melanogaster* has two subunits. A core enzyme with the polymerase activity of about 125 kDa, and polypeptide 35 kDa of unknown function. The enzymes isolated from other sources have also the core subunits over 105 kDa. It is not clear, however, if the core subunits are always homotetramers of a smaller polypeptide. Most of the enzyme preparations are associated with proofreading exonucleolytic activity,

which resides in a separate polypeptide. So far it was not established if the proofreading activity is derived from a separate gene product or from proteolysis of a single gene product [20].

III. Gene structure and gene expression

Except mitochondrial DNA polymerase γ, the cDNA encoding DNA polymerases have now been cloned directly or by PCR amplification techniques. The cDNA for core DNA polymerase α was cloned for human KB cells [21], HeLa cells [22] and for mouse FM3A cells [23]. The 5433-nucleotides long cDNA of KB cells DNA polymerase α contains a single open reading frame coding for protein with 1462 amino acids with calculated molecular mass 165 kDa. The promoter of this gene is highly GC rich, and does not contain the TATA box, but has CCAAT sequences on the opposite strand. The further promoter analysis reveals the 248 bp sequence, which is necessary for serum inducible expression. Within this promoter region there are consensus sequences similar to SP1-, AP1-, AP2- and E2F-binding sites [24]. The catalytic polypeptide of human polymerase α has been mapped to the short arm of the chromosome X at Xp21.3 to Xp22.1 [21]. The primase subunits were localized only for mouse. The genes coding for both primase polypeptides, p49 and p58, were localized to the mouse chromosome 1 and 10, respectively [25]. Northern hybridization studies for the human DNA polymerase α reveal a single 5.8-kb transcript, which is sufficient to encode a protein of 165-180 kDa [21, 22]. Studies of the human DNA polymerase α steady-state in quiescent cells stimulated to proliferation demonstrate over 20-fold increase of the level of this message. The level of α-pol mRNA is also higher in transformed cells than in normal ones. These data strongly suggest that the α polymerase mRNA level, like its enzymatic activity, positively correlates with cell proliferation and transformation. This also means that the DNA polymerase α gene expression is mainly regulated on transcriptional level [21, 26]. However, there are data showing that the enzymatic activity is also regulated on post-transcriptional level by the cell cycle correlated phosphorylation [27]. From the above one can conclude that p34 kinase, the key mitotic regulatory kinase, phosphorylates the catalytic polypeptide. The cell cycle dependent phosphorylation was observed for the 70-kDa subunit of DNA polymerase α [24, 27].

The human DNA polymerase δ gene is localized on chromosome 19. The cDNA of this polymerase has length of 3.5 kb, and encodes a protein of 1107 amino acids with calculated molecular mass of 124 kDa [28, 29]. The mouse enzyme cDNA encodes for 1105 amino acids. Their amino-acids sequence homology is 92% [30]. Northern blot analysis shows the presence of a single 3.4 kb transcript in both mouse and human [28-30]. In the cells stimulated to proliferation this transcript is induced with a time-course similar to the induction of DNA polymerase δ activity. The expression pattern of δ-pol mRNA throughout the cell cycle in proliferating cells and during the stimulation of quiescent cells is quite similar to that reported for DNA polymerase α and is consistent with the involvement of this enzyme in DNA replication [28].

The cDNA of the catalytic polypeptide of HeLa DNA polymerase ε encodes for a protein of 2257 amino-acids residues with a calculated molecular mass of 258 kDa [31]. The trypsin cleavage creates N-terminal catalytic polypeptide of 122 kDa of DNA polymerase ε* (amino acids from 1 to 1138) [31, 32]. A 136-kDa non-catalytic C-terminal domain of the enzyme (amino acids from 1139-2257) has a conserved cysteine cluster known as the zinc finger motif (amino acids from 2125 to 2220) and a highly acidic region (residues 1918-1948) [31-33]. The non-catalytic part of the human DNA polymerase ε is similar to the 30-34-kDa subunit of the yeast DNA polymerase ε what suggests that this acidic region is necessary for activation of transcription, very likely through interaction with other proteins [31, 33, 34]. Northern analysis reveals a single mRNA of 7.5 kb [31]. This mRNA is constitutively expressed throughout the cell cycle in proliferating cells. In quiescent cells the level of this transcript is very low but increases dramatically upon the stimulation [31, 33]. This expression pattern of DNA polymerase ε similar to DNA polymerase α and δ, strongly suggests that this enzyme may be a part of DNA replication machinery [33, 34].

DNA polymerase β is a much simpler protein than other eukaryotic cellular DNA polymerases. At the moment, it is much more known about its primary structure, gene structure and control of its expression than about the other DNA polymerases. The genes for human and mouse DNA polymerase β were mapped for both organisms to chromosome 8 (for human to 8p12). The sequence comparison of DNA polymerase β from both sources indicates that the enzyme has been highly conserved (90% homology of cDNA) [35]. There is also significant similarity between amino-acid sequences of DNA polymerase β and terminal deoxynucleotidyltransferase [36]. Organization of the entire human DNA polymerase β gene has been reported [37]. The gene spans 33 kb and contains 14 exons that range form 50 to 233 bp. The sequence analysis of the PCR product of β-pol mRNA from human cell lines indicates that there is an alternative splicing of the exon II during processing of the precursor RNA. Northern hybridization studies reveal that in human cells there is only one mRNA specie of 1.4 kb (deletion of 58 b of exon II cannot be seen during the regular Northern analysis). Similar studies in rat cells reveal four mRNA species of 1.4, 2.2, 2.5 and 4.0 kb. As we have shown [38] these high molecular mass transcripts arise from an alternative polyadenylation mechanism.

The β-pol gene expression is mainly regulated on transcriptional level, although, the inactivation of DNA polymerase β by *in vitro* phosphorylation with PKC has also been shown [39]. During the cell cycle the expression of β-pol gene is considered as a constitutive [40], but a small increase of the level of mRNA (about twice) during rat liver regeneration has been observed [41]. Similarly, the induction of this gene expression was found during tissues proliferation [41] as well as in mitogen-stimulated human lymphocytes [42]. Transcriptional activation of the β-pol gene by DNA-damaging agents is also observed in the cells after the exposure to MNNG [43] and in cells resistant to cisplatin [44].

The promoter for the human and mouse DNA polymerase β genes has been also well characterized [43, 45]. The upstream sequences contain no TATA or CCAAT elements but do contain so-called GC boxes known as the transcription factor Spl-binding consensus sequences. This promoter also consists of palindromic sequence identical to the binding site for activating transcription factor ATF/CREB. Many results indicate that the ATF/CREB transcriptional activator is a key regulatory protein for the human and mouse β-pol gene expression. Apart from the above-mentioned mechanism, yet another one exists in rodents. It was shown for rat and mouse that the DNA polymerase β gene is expressed in a tissue-specific manner [46, 47]. It is highly expressed in testis and in young brain, but in other adult organs the level of β-pol mRNA is low. This cell type-specific expression is regulated by two silencer elements, which have been found in the upstream sequence of the β polymerase gene of mouse [48].

IV. Physiological roles

Despite a long time of research the biological functions of DNA polymerases still remain uncertain. The model systems for mammalian DNA replication, especially for SV40 DNA replication, suggested that DNA polymerases α and δ are the sole polymerases required for DNA replication [49, 50]. The newest data suggest also involvement of DNA polymerase ε in this process [51, 52]. Apart from DNA polymerases, at least two multiprotein complexes named replication factor A (RF-A) and replication factor C (RF-C), and at least one additional auxiliary protein called proliferating cell nuclear antigen (PCNA) are involved in the replication process.

The replication factor A is a multisubunits single-stranded DNA binding activity. It contains three polypeptides of 70, 34 and 11 kDa. The 70-kDa protein is probably involved in DNA unwinding. The 34-kDa polypeptide is a phosphoprotein, which is highly phosphorylated in the late G1 and early S phase of the cell cycle. This protein stimulates of DNA synthesizing activity of DNA polymerase α. RF-A is required for initiation and elongation steps.

The replication factor C is also multisubunits complex with primer/template-binding activity. It also contains DNA-dependent ATPase activity. RF-C has three subunits: 140, 41 and 37 kDa. The 140-kDa protein contains a template/primer-binding activity. The ATP-binding activity is carried by the 41-kDa protein.

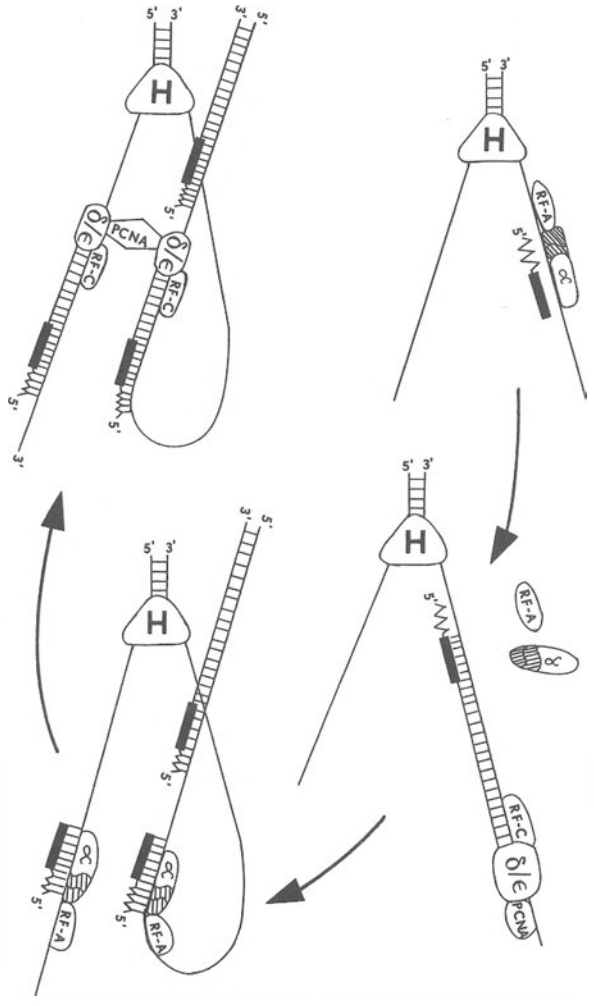


Fig. 1. Hypothetical scheme for replication of leading and lagging strands.
ΛΛΛΛΛΛ RNA primer; ▬ DNA PRIMER;
H, helicase-topoisomerase activity

PCNA is a 36-kDa protein that enhances processivity of DNA polymerase δ.

Figure 1 presents a proposed model of events during DNA replication. The starting sequences (*ori* in SV40 model system) have to be unwound by combined action of topoisomerase and helicase (H on Fig. 1). Created single stranded DNA of the leading and lagging strand is covered by RF-A (see above). The primase activity of DNA polymerase α/primase complex synthesizes a short primer on the lagging strands. Next, polymerase α uses this primer to synthesize the nascent DNA fragment. During the synthesis of the short fragment of DNA, the RF-C and PCNA interact with the lagging strand. This binding prevents poly-

merase α from further DNA synthesis. The 3′ end of nascent DNA is now recognized by DNA polymerase δ or ε, which takes over the DNA synthesis on the lagging strand. In that moment DNA polymerase α/primase dissociates from that complex and moves to the next priming site on the lagging strand. The synthesis cycle starts again. Simultaneously, DNA synthesis starts on the leading strand. The lagging strand creates a loop and the further DNA synthesis on both strands can go just like in the same direction. The synthesis on the leading strand is also initiated by DNA polymerase α/primase complex and, after forming RF-C and PCNA proteins complex, the synthesis is continued by DNA polymerase δ or ε [7, 53-55]. In this model the DNA polymerase α/primase role is limited only to the synthesis of primer DNA. The synthesis on both strands, the lagging and leading, is carried out by DNA polymerase δ or ε. So far it is not established if DNA polymerase δ or ε is responsible for DNA synthesis on the lagging or leading strand. However, the involvement of DNA polymerase δ in the replication of the leading strand is preferred [7, 54, 55]. On the other hand, in SV40 model both strands are synthesized by DNA polymerase δ [51], so it cannot be excluded that both strands are replicated solely by DNA polymerase δ. The role of DNA polymerase ε in replication is now the topic of research in many laboratories. The link of DNA polymerase ε to the replication machinery in the S phase checkpoint was confirmed [33].

In spite of the insufficient knowledge about the structure of DNA polymerase γ its biological function is well documented. This enzyme is required for the replication of mitochondrial and chloroplast DNA. Probably this is the only DNA polymerase involved in the replication of circular DNA in organelles. It is also plausible that DNA polymerase γ is able to replicate each circular extrachromosomal DNA in the cell.

DNA polymerase β is a highly conserved enzyme in vertebrates. The β-pol gene is expressed in all cells, therefore, is seen as a "housekeeping" gene. Deletion of the promoter and the first exon of the DNA polymerase β gene in the mouse germ line results in a lethal phenotype [56]. Based on the enzymatic studies with differential inhibitors and on the ability of this enzyme to fill small gaps [57-60] and nicks (R. Nowak, unpublished results), the involvement of DNA polymerase β to the DNA repair was suggested. It was also shown that DNA polymerase β can substitute for *E. coli* DNA polymerase I during DNA repair synthesis [61] and DNA replication as well [62]. The induction of the expression of DNA polymerase β during the rat liver regeneration and in stimulated human lymphocytes suggests that the enzyme is also associated with DNA replication, either by screening the DNA template just before replication or by increased repair of the newly synthesized DNA strand [41, 42, 63]. The induction of the β-pol gene expression after treatment of the cells with DNA-damaging agents such as MNNG was also observed [43]. The resistance to cisplatin during the chemotherapy is connected with an increase of DNA repair and the induction of DNA polymerase β steady-state transcript [44]. All these facts strongly suggest the essential role of DNA polymerase β in DNA repair synthesis. On the other hand, it was revealed that for example in UV-induced repair synthesis the other polymerases also participate [1, 7, 64].

Apart from DNA polymerase β, α/δ or ε in human at least twelve gene products are required for nucleotide excision repair. These are: (a) XPA and XPE, the presumed function of them is photoproduct binding; (b) XPD, XPB, CSA and CSB, which have helicase activity and/or transcription coupling function (these products are involved in transcription-dependent repair); (c) XPC, perhaps ATPase involved in transcription-independent repair; (d) ERCC1, ERCC4, ERCC11, XPF and XPG with the presumed function of nuclease activity [65-69].

Specially interesting are the studies that imply the involvement of DNA polymerase β in post-recombinational DNA synthesis. It is known that the meiotic DNA synthesis is registered only in pachytene spermatocytes where the recombination events during the meiosis takes place. Two different laboratories have shown an extremely high level of β-pol mRNA in pachytene spermatocytes [46, 47, 70].

V. Conclusions

As one can see our understanding of mammalian DNA polymerase, DNA replication and DNA repair is greatly advanced. However, our knowledge about DNA metabolism processes is still far of complete. It is necessary to elucidate the interaction between all replicative and repair proteins at the molecular level. This will lead us to the understanding how the eukaryotic cell proliferation is controlled and how the maintenance of the genome integrity is achieved.

Note added in proofs. In recently published studies (Sobol M *et al.* (1996) *Nature* **379**: 183-186) it was shown that β-polymerase functions *in vivo* specifically in base-excision repair (BER), and cannot be substituted by any other DNA polymerase. This enzyme is a rate-limiting factor in the BER pathway *in vivo*.

References

1. Wang TS-F (1991) *Annu Rev Biochem* **60**: 513-552
2. Fry M, Loeb LA (1986) In: DNA polymerases. CRC Press, Boca Raton, Fl
3. Perrino FW, Loeb LA (1990) *Mutation Res* **236**: 289-300
4. Wilson SH, Abbotts J, Widen S (1988) *Biochim Biophys Acta* **949**: 149-157
5. Bambara RA, Jessee CB (1991) *Biochim Biophys Acta* **1088**: 11-24
6. Blow J (1989) *Trends Genet* **5**: 134-136
7. Hubscher U, Thommes P (1992) *Trends Biochem Sci* **17**: 55-58

8. Bollum FJ, Potter VR (1957) *J Am Chem Soc* **79**: 3603-3604
9. Collins KL, Russo AAR, Tseng BY, KellyTJ (1993) *EMBO J* **12**: 4555-4566
10. Kumar A, Wilden SG, Williams KR, Kedar P, Karpel RL, Wilson SH (1990) *J Biol Chem* **265**: 2124-2131
11. Kumar A, Abbotts J, Karawaya EM, Wilson SH (1990) *Biochemistry* **29**: 7156-7159
12. Kim SJ, Lewis MS, Knutson JR, Porter DK, Kumar A, Wilson SH (1994) *J Mol Biol* **244**: 224-235
13. Davies II JF, Almassy JR, Hostomska Z, Ferre RA, Hostomsky Z (1994) *Cell* **76**: 1123-1133
14. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J (1994) *Science* **264**: 1891-1903
15. Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut SH (1994) *Science* **264**: 1930-1935
16. Byrnes JJ, Downey KM, Black VL, So AG (1976) *Biochemistry* **15**: 2817-2823
17. Kesti T, Syvaoja JE (1991) *J Biol Chem* **266**: 6336-6341
18. Kesti T, Frantti H, Syvaoja JE (1993) *J Biol Chem* **268**: 10238-10245
19. Niranjanakumari S, Gopinathan KP (1993) *J Biol Chem* **268**: 15557-15564
20. Longley MJ, Mosbaugh DW (1991) *J Biol Chem* **266**: 24702-24711
21. Wong SW, Wahl AF, Yuan P-M, Arai N, Pearson BE, Arai K-I, Korn D, Hunkapiller MW, Wang TS-F (1988) *EMBO J* **7**: 37-47
22. SenGupta DN, Kumar P, Żmudzka BZ, Coughlin S, Vishwanatha JK, Robey FA, Parrott C, Wilson SH (1987) *Biochemistry* **26**: 956-963
23. Miyazawa H, Izumi M, Tada S, Takada R, Masutani M, Ui M, Hanaoka F (1993) *J Biol Chem* **268**: 8111-8122
24. Pearson BE, Nasheuer H-P, Wang TS-F (1991) *Mol Cell Biol* **11**: 2081-2095
25. Adler DA, Tseng BY, Wang TS-F, Disteche CM (1991) *Genomics* **9**: 642-646
26. Wahl AF, Geis AM, Spain BH, Wong SW, Korn D, Wang TS-F (1988) *Mol Cell Biol* **8**: 5016-5025
27. Nasheuer H-P, Moore A, Wahl AF, Wang TS-F (1991) *J Biol Chem* **266**: 7893-7928
28. Yang C-L, Chang L-S, Zhang P, Hao H, Zhu L, Toomey NL, Lee MYWT (1992) *Nucleic Acids Res* **20**: 735-745
29. Chung DW, Zhang J, Tan CK, Davie EW, So AG, Downey KM (1992) *Proc Natl Acad Sci USA* **88**: 11197-11201
30. Cullmann G, Hindges R, Berchtold MW, Hubscher U (1993) *Gene* **134**: 191-200
31. Kesti T, Franti H, Syvaoja JE (1993) *J Biol Chem* **268**: 10238-10245
32. Kesti T, Syvaoja JE (1991) *J Biol Chem* **266**: 6336-6341
33. Navas TA, Zhou Z, Elledge SJ (1995) *Cell* **80**: 29-39
34. Uitto L, Halleen J, Hentunen T, Hoyhtya M, Syvaoja JE (1995) *Nucleic Acids Res* **23**: 244-247
35. Sen Gupta DN, Żmudzka BZ, Kumar P, Cobianchi F, Skowroński J, Wilson SH (1986) *Biochem Biophys Res Commun* **136**: 341-347
36. Matsukage A, Nishikawa K, Ooi T, Yoshihiko S, Yamaguchi M (1987) *J Biol Chem* **262**: 8960-8962
37. Chyan YJ, Ackerman S, Shepherd NS, McBride OW, Widen SG, Wilson SH, Wood GT (1994) *Nucleic Acids Res* **22**: 2719-2725
38. Konopiński R, Nowak R, Siedlecki JA (1995) *Acta Biochim Polon* **42**: 253-258
39. Tokui T, Inagaki M, Nishizawa K, Yatani R, Kusagawa M, Ajiro K, Nishimoto Y, Date T, Matsukage A (1991) *J Biol Chem* **266**: 10820-10824
40. Żmudzka BZ, Fornance A, Collins J, Wilson SH (1988) *Nucleic Acids Res* **16**: 9587-9596
41. Nowak R, Siedlecki JA, Kaczmarek L, Żmudzka BZ, Wilson SH (1989) *Biochim Biophys Acta* **1008**: 203-207
42. Suzuki H, Mencgazzi M, De Prati AC, Ogura T, Esumi H, Matsukage A, Libonati M (1991) *Biochem Biophys Res Commun* **181**: 623-628
43. Narayan S, Beard WA, Wilson SH (1995) *Biochemistry* **34**: 73-80
44. Chu G (1994) *J Biol Chem* **269**: 787-790
45. Yamaguchi M, Hayashi Y, Matsuoka S, Takahashi T, Matsukage A (1994) *Eur J Biochem* **221**: 227-237
46. Hirose F, Hotta Y, Yamaguchi M, Matsukage A (1989) *Exp Cell Res* **181**: 169-180
47. Nowak R, Woszczyński M, Siedlecki JA (1990) *Exp Cell Res* **191**: 51-56
48. Yamaguchi M, Hayashi Y, Furukawa K, Nishimoto Y, Matsukage A (1991) *Jpn J Cancer Res* **82**: 72-81
49. Waga S, Stillman B (1994) *Nature* (Lond) **369**: 207-212
50. Waga S, Bauer G, Stillman B (1994) *J Biol Chem* **269**: 10923-10934
51. Morrison A, Araki H, Clarck AB, Hamatake RK, Sugino A (1990) *Cell* **62**: 1143-1151
52. Morrison A, Sugino A (1993) *Prog Nucleic Acid Res Mol Biol* **46**: 93-120
53. Weiser T, Gassman M, Thommes P, Ferrari E, Hafkemeyer P, Hubscher U (1991) *J Biol Chem* **266**: 10420-10428
54. Podust VN, Georgaki A, Strack B, Hubscher U (1992) *Nucleic Acids Res* **20**: 4159-4165
55. Turchi JJ, Bambara RA (1993) *J Biol Chem* **268**: 15136-15141
56. Gu H, Marth JD, Orban PC, Mossmann H, Rajewsky K (1994) *Science* **265**: 103-106
57. Siedlecki JA, Szyszko J, Pietrzykowska J, Zmudzka B (1980) *Nucleic Acids Res* **8**: 361-375
58. Wang TS-F, Korn D (1980) *Biochemistry* **19**: 1782-1790
59. Nowak R, Kulik J, Siedlecki JA (1987) *Acta Biochim Polon* **34**: 205-215
60. Singhal RK, Wilson SH (1993) *J Biol Chem* **268**: 15906-15911
61. Ohnishi T, Yuba S, Date T, Utsumi H, Matsukage A (1990) *Nucleic Acids Res* **18**: 5673-5676
62. Sweasy JB, Loeb LA (1992) *J Biol Chem* **267**: 1407-1410
63. Mcnegazzi M, Carcereri de Prati A, Ogura T, Columbano A, Ledda-Columbano GM, Libonati M, Esumi H, Suzuki H (1992) *FEBS Lett* **310**: 135-138
64. Wang Z, Wu X, Friedberg EC (1993) *Mol Cell Biol* **13**: 1051-1058
65. Sancar A (1994) *Science* **266**: 1954-1956
66. Cleaver JE (1994) *Cell* **76**: 1-4
67. Matsuda T, Saijo M, Kuraoka I, Kobayashi T, Nakatsu Y, Nagai A, Enjoji T, Masutani C, Sugasawa K, Hanaoka F, Yasui A, Tanaka K (1995) *J Biol Chem* **270**: 4152-4157
68. Tanaka K, Wood RD (1994) *Trends Biochem Sci* **19**: 83-86
69. Walicka M (1994) *Post Biochem* **40**: 77-85 (in Polish)
70. Nowak R, Siedlecki JA (1991) *Mol Biol Reports* **15**: 25-31

Butylphenyl-dGTP as a structural probe of B family DNA polymerases

JAMES M. STATTEL[1],
GEORGE E. WRIGHT[2]

Contents:

Abbreviations used: pol — DNA polymerase; **pol** — DNA polymerization activity; **exo** — 3′ to 5′ exonuclease activity; BuPdGTP — N^2-(*p*-*n*-butylphenyl)-2′-deoxyguanosine 5′--triphosphate; BVdU — 5-(2-bromovinyl)-2′-deoxyuridine; araNTP — arabinonucleoside 5′-triphosphate; ddNTP — 2′, 3′-dideoxynucleoside 5′-triphosphate; PFA — phosphonoformate; PAA — phosphonoacetate.

I. Introduction

I-1. DNA-dependent DNA polymerases

DNA-dependent DNA polymerases are complex enzymes which bind multiple substrates and many possess multiple catalytic functions. The most extensively studied enzyme, *E. coli* DNA polymerase I (pol I), binds sequentially to partially double-stranded oligodeoxyribonucleotide primer:template and 2′-deoxyribonucleoside 5′-triphosphates (dNTP's) in a polymerization (**pol**) complex resulting in incorporation of the nucleoside 5′-monophosphate to the primer with release of pyrophosphate (PP_i). Pol I also possesses a 3′ to 5′ exonuclease (**exo**) activity that cleaves nucleoside 5′-monophosphates from double- or single-stranded DNA, and a 5′ to 3′ exonuclease activity.

[1, 2] Department of Pharmacology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655 USA

I-2. Structure of DNA-dependent DNA polymerases

The X-ray crystal structure of a partially proteolyzed fragment (Klenow fragment) of pol I, retaining **pol** and 3′ to 5′ **exo** functions [1], revealed a tertiary structure resembling an open, curled right hand that can accommodate double-stranded DNA (Fig. 1A). Genetic and chemical methods have suggested that amino-acid residues in the "palm" and "fingers" are involved in catalysis, while residues in the thumb are possibly involved in DNA synthetic processivity (the ratio k_{cat}/k_{off} of a polymerase for primer:template). The 3′ to 5′ **exo** site, located **ca.** 20 Å from the **pol** active site, was identified by its binding to dTMP. Attempts to co-crystallize the Klenow fragment with its substrates have yielded structures in which primer:template was bound in an unusual conformation [2], and in which a dNTP molecule was bound in a complex that was unlikely to be catalytically relevant [3].

Recently, the X-ray structure of crystals of eukaryotic DNA polymerase β (pol β) in a complex with primer:template and a substrate analog has been reported [4]. This structure represents the most catalytically relevant picture to date of a DNA-dependent DNA polymerase and shows that acidic residues in the palm domain are critical for divalent cation binding, ensuring the correct positioning of the incoming nucleoside triphosphate (Tab. 1). Pol β, possessing only **pol** activity and little primary sequence homology to other DNA polymerases, has nonetheless suggested a common catalytic mechanism of dNTP incorporation from comparison of structural motifs of pol β with those of pol I, HIV reverse transcriptase and T7 RNA polymerase [5-7].

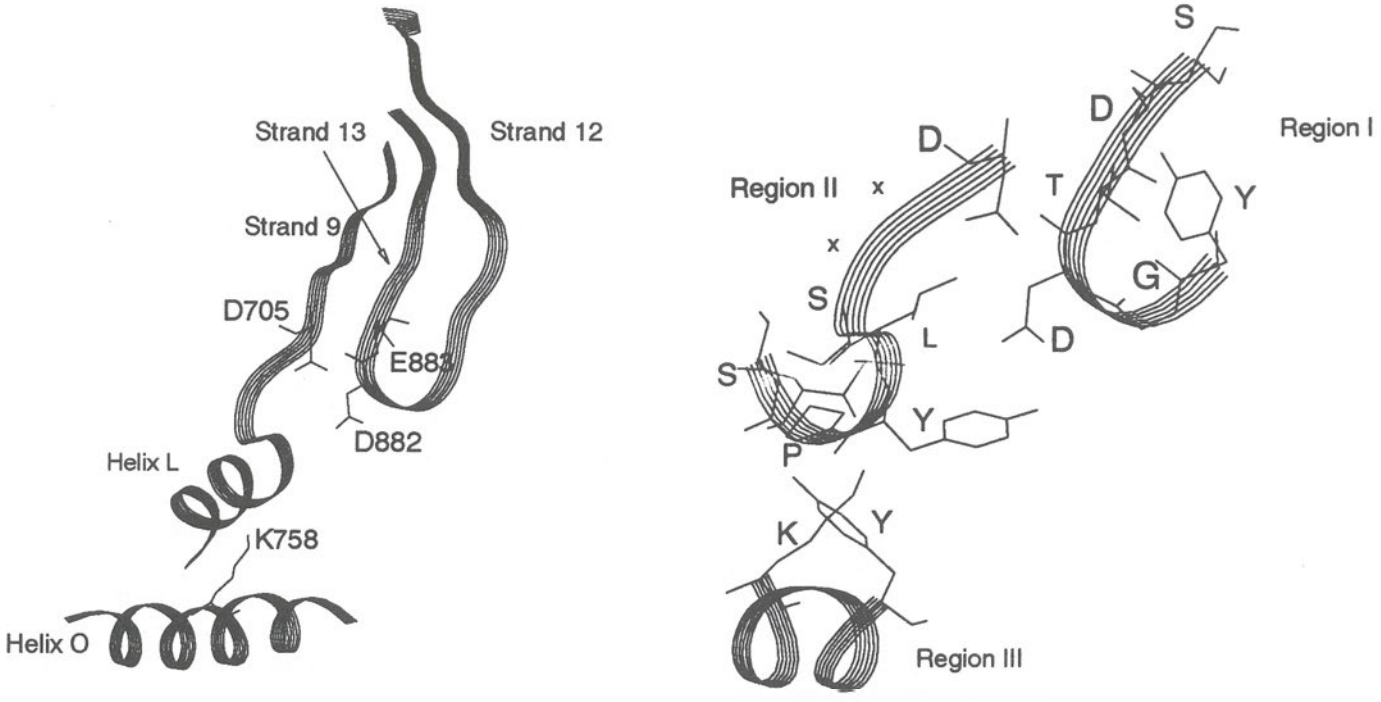


Fig. 1. A. Representation of part of the **pol** site of *E. coli* DNA polymerase I. Coordinates and secondary structure nomenclature were derived from [1]. B. Hypothetical representation of the **pol** site of B family DNA polymerases. The model is based on homologies with the corresponding sequences of *E. coli* pol I. Figures were obtained using Biosym molecular modelling sofware (Biosym) displayed on an Iris Indigo computer.

Table 1.

Conservation of acidic amino-acid residues critical for DNA polymerase function

Family	Enzyme	Residues		
A	pol I	D705	D882	E883
B	HSV1 pol	D717	D886	D888
	ϕ29 pol	D249	D456	D458
	T4 pol	D408	D618	D620
	pol α	D860	D1002	D1004
D	pol β	D256	D190	D192

II. B family DNA polymerases

II-1. Sequence similarity

Pol I and pol β represent two of the four families (A and D, respectively) of DNA polymerases, based upon amino-acid sequence homology [8]. The largest family, the B family, consists of animal cell DNA polymerases α, δ and ε, the corresponding enzymes from yeast, herpesvirus DNA polymerases, T-even bacteriophage DNA polymerases, and other small phage DNA polymerases such as ϕ29 pol and PRD1 pol. Three regions of homology in the N-terminal half of B family enzymes have been implicated in **exo** activity [9]. In addition, B family polymerases contain up to seven regions of significant sequence homology in the C-terminal region suggested to be in the **pol** domain based, in part, on effects of amino-acid mutations in these regions derived from drug resistant organisms or by site-directed mutagenesis. In particular, conserved regions I, II and III (Tab. 3-5) contain residues important for primer:template binding and dNTP catalysis.

A hypothetical model of the **pol** site of the B family enzymes is shown in figure 1B. This model is based on one derived from attempts to fit secondary structure elements and mutation data of HSV1 pol to the structure of *E. coli* pol I [10]. Although primary sequences of pol I and B family polymerases are generally unrelated, the secondary structure of the active sites may be similar, as emphasized in the comparison of figures 1A and 1B. In particular, the triad of acidic residues (Tab. 1) spanning regions I and II of the B family and α helices 9, 12 and 13 of pol I have their counterparts in the structure of the pol β complex [4].

II-2. Inhibitors

Elucidation of some structural features of **pol** active sites has been accomplished by using specific inhibitors that have unique action against B family polymerases. These inhibitors include dNTP derivatives that have been modified in the base, such as N^2-(*p*-*n*-butylphenyl)-2′-deoxyguanosine 5′-triphosphate (BuPdGTP) and E-5-(2-bromovinyl)-2′-deoxyuridine (BVdU), or in the sugar, such as arabinonucleoside 5′-triphosphates (araNTPs), 2′,3′-dideoxynucleoside 5′-triphosphates (ddNTPs), and acyclonucleotides such as acyclovir triphosphate. dNTP analogs functioning as inhibitors/substrates of DNA polymerases are reviewed in reference [11]. Pyrophosphate analogs such as phosphonoformate (PFA) [12] and dNTP mimics such as aphidicolin [11] also display selective inhibition of B family polymerases.

Table 2.

Inhibitory characteristics of BuPdGTP on B family polymerases

Enzyme	BuPdGTP (K_i, μM)	Is it a substrate?	Ref.
pol α	0.001-0.005	no	[13-15]
T4 pol	0.82	yes	[17]
ϕ29 pol	20	yes	[16]
HSV1 pol	52	yes	[a]

a, Knopf CH, personal communication.

Of particular interest are the varied responses of B family polymerases to BuPdGTP (Tab. 2). BuPdGTP was initially synthesized as a potent and selective inhibitor of pol α [13, 14]. Because BuPdGTP has been found to display a wide range of potencies and multiple inhibition mechanisms on B family polymerases, careful analysis of the effects of mutations in conserved regions of B family polymerases is yielding insight into several levels of inhibitor:enzyme interactions.

III. Butylphenyl-dGTP: its role in study of B family DNA polymerases

III-1. Mechanism of action

BuPdGTP was found to inhibit DNA synthesis by pol α competitively with dGTP and with $K_i = 1$ nM [14]. Subsequent analysis of the mechanism of inhibition revealed that BuPdGTP action on pol α does not involve polymerization of the dNMP moiety (Tab. 2). A synthetic primer:template designed to incorporate dGTP as the next required nucleotide did not show primer extension in the presence of 50 μM BuPdGTP and pol α when analyzed by polyacrylamide gel electrophoresis (PAGE) [15].

BuPdGTP inhibits the **pol** activity of ϕ29 pol and HSV1 pol with much lower potencies (i.e. competitive K_i's = 20 and 52 μM, respectively) [16] and (Knopf CH, personal communication). T4 pol is intermediate in sensitivity to BuPdGTP, with a competitive K_i of 0.82 μM [17]. All three enzymes, however, incorporate BuPdGTP when dGTP is the next required nucleotide specified by the template [15, 16] and (Knopf CH, personal communication). This finding is curious, considering the probable similarity in active site structure between these and α polymerases. Incorporation of BuPdGTP by T4 pol, a process that has been studied in detail [17], leads to a modified primer:template that has higher affinity to the enzyme than the parent inhibitor. Indeed, a synthetic 3′-BuPdG-primer:template was a potent inhibitor of T4 pol [17].

Systematic mutational analyses of conserved regions in B family polymerases have revealed much information about the function and possible secondary structure of amino-acid residues that comprise the **pol** active site. Effects of mutations on dNTP and primer:template binding, dNTP incorporation/mis-insertion rates and inhibitor sensitivity, particularly to BuPdGTP, aphidicolin, and pyrophosphate analogs, have been determined, and some of the results are summarized here.

III-2. Region I

Region I (YGDTDS; Tab. 3) contains two of the three conserved acidic residues, separated by an invariant threonine, proposed to be necessary for DNA catalysis. The "di-acidic" motif can be found in all DNA polymerase families [8]; similar motifs (YMDD, YGDD and YVDD) are found in viral reverse transcriptases and RNA-directed RNA polymerases [18]. Mutational studies in Region I of numerous B family polymerases have determined that this region is involved in substrate recognition and catalysis, specifically acting as a magnesium binding domain. None of the characterized drug resistant HSV1 mutants mapped to the **pol** gene have involved mutations in Region I [19, 20]. However, site-directed HSV1 virus mutations, e.g. G885R, D886N, T887K, and D888A, yielded enzymes with negligible **pol** activity [19]. A separate study found that the alteration G885A eliminated **pol** function, while F891C and F891Y mutant viruses were PAA and acyclovir resistant, but aphidicolin hypersen-

Table 3.

Region I of B family polymerases: phenotypes of site-directed and drug resistant mutants

consensus:		x	Y	G	D	T	D	S	x	x
pol	α (999):	I	Y	G	D	T	D	S	I	M
	mutants:		F		N	S	N	T		
	pol:		–		–	–	i	–		
	BuPdGTP:		r		r	r	?	s		
	Aph:		s		r	r	?	s		
ϕ29	(453):	I	Y	C	D	T	D	S	I	H
	mutants:		F	G	G	P	G			
	pol:		–	–	–	i	i			
	exo:		wt	wt	wt	wt	wt			
HSV1	(883):	I	Y	G	D	T	D	S	I	F
	mutants:			R, A	N	K	A			C, Y
	pol:			i, i	i	i	i			–
	PAA:			?, ?	?	?	?			r
	ACV:			?, ?	?	?	?			r
	Aph:			?, ?	?	?	?			hs
PRD1	(425):	L	Y	C	D	T	D	S	I	I
	mutant*:				Y			T		
	pol:				i			i		

Legend: Aph, aphidicolin; ACV, acyclovir; PAA, phosphonoacetate. Inhibitors: s, normosensitive; r, resistant; hs, hypersensitive; ?, unknown.

Activities: wt, like wild type; –, less than wild type; i, inactive.

*: PRD1 pol double mutant (D428Y/S431T).

sitive [20]. A site-directed double mutant in this region in PRD1 pol, D428Y/S431T, also resulted in virus that lost **pol** activity [21].

In ϕ29 [22, 23], the Y454F mutant polymerase could initiate protein-primed DNA synthesis as well as the wild type enzyme, but had reduced **pol** activity. The C455G and D456G mutants had reduced abilities to initiate synthesis from a protein primer and reduced ability to replicate M13 DNA. The T457P mutant had no initiation or **pol** capabilities, possibly resulting from changes in the secondary structure of the active site. The D458G mutant was also unable to initiate or sustain DNA synthesis. All mutated ϕ29 pols had wild type **exo** activity [22, 23].

Mutations in Region I of pol α were shown to cause marked changes in inhibitor sensitivity [24]. The Y1000F and S1005T mutants had reduced **pol** activities, and both enzymes exhibited wild type sensitivity to aphidicolin, but were 30-fold and 3-fold resistant to BuPdGTP, respectively. D1002N and T1003S mutants exhibited barely detectable **pol** activity and 10-25 fold resistance to both BuPdGTP and aphidicolin. The D1004N mutant had no detectable **pol** activity.

III-3. Region II

Region II (DxxSLYPSI; Tab. 4) contains the third conserved acidic residue of the catalytic triad important to **pol** activity. Site-directed mutagenesis of this residue in the PRD1 pol gene produced a phage that had less than 1% of wild type plaque forming ability [21]. Nucleic acid sequencing of drug resistant mutants of HSV1 pol showed that the mutant S724N was resistant to PAA and acyclovir, but hypersensitive to aphidicolin [25].

Several mutants of T4 pol have been studied in detail [26]. The L412M mutant is deficient in **pol/exo** switching and hypersensitive to PAA. The I417V mutant has enhanced **exo** activity, but exhibits wild-type sensitivity to PAA. Both L412M and I417V mutants have sensitivities to BuPdGTP comparable to that of wild type T4 pol [27]. The mutants incorporate the inhibitor with similar efficiencies, but are less tightly bound to the 3′-BuPdGMP-modified primer:template than the wild type enzyme [27].

Mutational analyses of ϕ29 pol showed that certain residues in region II are absolutely critical for synthetic activity and template binding [16, 28]. The Y254F mutant had reduced **pol** activity on protein-primed DNA and M13 DNA in the presence of Mg^{2+}, but polymerized both types of DNA at wild type levels in the presence of Mg^{2+}. The Y254F mutant was deficient in protein-primed **pol** initiation due to a reduced affinity for Me^{2+}-dATP, the initiating nucleotide. Also, Y254F pol was 100-fold more sensitive to BuPdGTP than the wild type enzyme.

Among additional mutants of ϕ29 pol in region II S252G, L253V and P255S mutants retained 60-70% of wild type **pol** activity and were able to translocate normally [28]. The S252R mutant had no detectable polymerase activity when Mg^{2+} was used as divalent cation, but recovered 10% activity when Mn^{2+} was used. The S252G, S252R, and P255S mutants showed reduced ability to bind template:primer; the S252G and P255S mutants were both hypersensitive to PAA.

Side-directed mutants in region II of human pol α have been extensively characterized [29, 30]. The mutations D860A, D860S, D860N, Y865S and Y865F did not significantly affect the enzymes' ability to bind primer:template or their processivity. Mutants D860A, Y865S and Y865F had much higher K_m values for dNTPs, however. The D860S and Y865S mutants were two-fold and ten-fold resistant to aphidicolin, respectively, while the D860A and D860N mutants were ten-fold hypersensitive to aphidicolin. Mutant Y865F exhibited wild type sensitivity to aphidicolin. The mutants D860A, D860S and D860N were four-fold to eight-fold resistant to BuPdGTP, while mutants Y865S and Y865F were 200-fold resistant to

Table 4.

Region II of B family polymerases: phenotypes of site-directed and drug resistant mutants

consensus		D	x	x	S		L	Y		P	S	I
pol	α (860):	D	F	N	S		L	Y		P	S	I
	mutants:	A, S, N			A, T			S, F			A, T	
	K_m(dNTP):	+, +, +			wt, wt			+, +			wt, wt	
	K_m(p:t):	wt, wt, wt			?, wt			wt, wt			+, wt	
	BuPdGTP:	r, r, r						hs, hs				
	Aph:	hs, r, hs						r, wt				
φ29	(249):	D	V	N	S		L	Y		P	A	Q
	mutants:				G, R		V	F		S		
	pol (Mg^{2+}):				−, i		−	−		−		
	pol (Mn^{2+}):				−, −			wt				
	K_m(Me^{2+}:dATP):							−				
	K_m(p:t):				−, −		−			−		
	BuPdGTP:							hs				
	PAA:				hs					hs		
T4	(408):	D	L	T	S	L	Y	P		S	I	I
	mutants:					M						V
	pol/exo switching:					−						
	exo											+
	PAA:					hs						wt
	BuPdGTP:					wt						wt
	p:t binding:					−						−
HSV1	(717):	D	F	A	S		L	Y		P	S	I
	mutant:										N	
	PAA:										r	
	ACV:										r	
	Aph:										hs	

Legend: Aph, aphidicolin; ACV, acyclovir; PAA, phosphonoacetate. Inhibitors: s, normosensitive; r, resistant; hs, hypersensitive. Activities: wt, like wild type; +, greater than wild type; −, less than wild type; i, inactive; p:t, primer:template.

BuPdGTP. These results confirm that Asp-860 (Tab. 1) is important in binding substrates, and suggest that Tyr-865 is also critical for dNTP binding.

Conversely, mutation of Ser-863 and Ser-867 to either Ala or Thr produced enzymes that had wild type levels of K_m for dNTPs, k_{cat}, and processivity, although the S867A mutant had a thirty-fold higher K_m for primer:template. Use of DNA footprint experiments with 2′,3′-dideoxyNTPs indicated that the 3′-OH of the primer terminus is the structural feature recognized by Ser-867 [30]. This conclusion is based on the observation that the S867A mutant had an increased capacity to extend a mis-paired primer terminus.These results suggested that the OH group of Ser-867 interacts with the primer during DNA synthesis [30].

III-4. Region III

Region III (KLxxNSxYG; Tab. 5) is a site of several drug-resistant mutations in the B family DNA polymerases. The Lys residue in this region is invariant in B family polymerases and is analogous to Lys-758 of Pol I (Fig. 1A) [18, 31, 32]. In Pol I, Lys-758 appears to be directly involved in dNTP binding and template translocation [31]. Its importance in B family polymerases is suggested by the finding that K340H, K340D and K340E mutants of PRD1 pol had no detectable **pol** activity, but had wild type **exo** activity

Table 5.

Region III of B family polymerases: phenotypes of site-directed and drug resistant mutants

consensus:		K	L	x	x	N	S	x	Y	G
pol	α (950):	K	L	T	A	N	S	M	Y	G
	mutants:	A, R, N								
	BuPdGTP:	hs, hs, hs								
	PAA:	hs, hs, r								
φ29	(383):	K	L	M	L	N	S	L	Y	G
	mutants:								S, F	
	pol (Mg^{2+}):								−, i	
	pol (Mn^{2+}):								wt, wt	
	BuPdGTP:								hs, hs	
HSV1	(811):	K	V	Y	C	N	S	V	Y	G
	mutant:			M						
	PAA:			r						
	ACV:			r						
	Aph:			hs					r	
PRD1	(340):	K	L	I	L	N	S	S	Y	G
	mutant:	H, D, E								
	pol:	i, i, i								

Legend: Aph, aphidicolin; ACV, acyclovir; PAA, phosphonoacetate. Inhibitors: s, normosensitive; r, resistant; hs, hypersensitive. Activities: wt, like wild type; +, greater than wild type; −, less than wild type; i, inactive.

[32]. Mutation of this residue in pol α results in enzymes that are hypersensitive to BuPdGTP but with variable responses to PAA (Wang TS-F, 1994, personal communication). The K950A and K950N mutants are 100-fold hypersensitive to BuPdGTP, but they are fifteen-fold hypersensitive and two-fold resistant to PAA, respectively (Wang TS-F, 1994, personal communication).

In HSV1, sequencing of drug resistant mutants involving region III revealed that the V813M mutant was resistant to araA, PAA and acyclovir, but hypersensitive to BVdU and aphidicolin, while the T821M mutant, in a residue just past region III, was aphidicolin resistant [25]. The HSV1 pol residue Tyr-818, when mutated to Gly, produces mutant viruses that are aphidicolin-resistant in marker rescue experiments [33].

In φ29 pol, mutant Y390F had greatly reduced **pol** activity, both in protein-primed and primed M13 DNA replication, when Mg^{2+} was used as divalent cation [16]. Use of Mn^{2+} restored **pol** activity to near wild type levels. Mutant Y390S had slightly reduced protein-primed and M13 DNA replicative activity, which was restored to wild type levels when Mn^{2+} was used. Both the Y390S and Y390F mutants were hypersensitive to BuPdGTP. These results suggest that Tyr-390 is critical for template-dependent dNTP binding [16].

IV. Hypothetical structure of B family DNA polymerases

Information obtained from mutational changes in conserved regions of B family polymerases can be used to predict a model for the overall structure of the **pol**

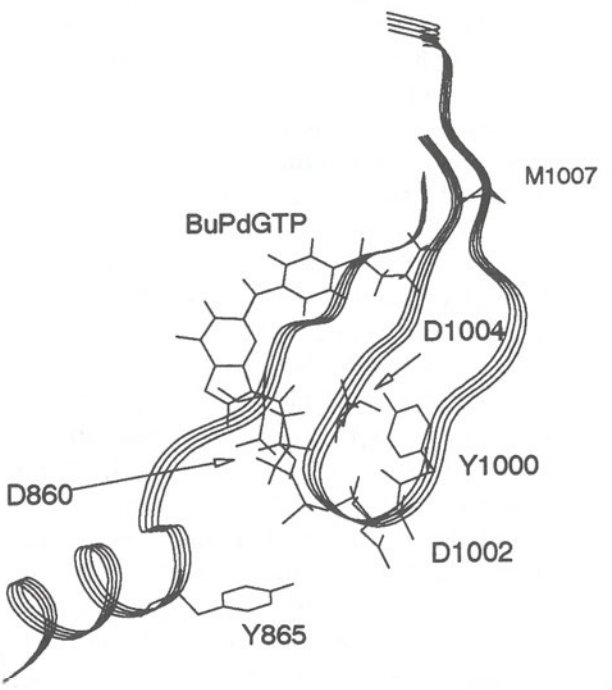


Fig. 2. A model of the interaction of BuPdGTP with the **pol** site of DNA polymerase α. See text for details.

site. Specifically, our consideration of how BuPdGTP interacts with the **pol** site to produce such varied mechanistic responses within the highly conserved B family polymerases has led to a model of the **pol** site, which is detailed for pol α in figure 2. This model is based upon the backbone secondary structure of pol I and on predicted amino-acid identities with pol I and pol β [1, 4-7]. The three aspartate residues in the model are positioned to coordinate the two divalent cations which, in turn, position the incoming dNTP.

We note the large number of generally conserved aromatic residues in this region of the B family polymerases (Tab. 3-5). The moderate sensitivity of most of the enzymes in the B family to BuPdGTP may result from interaction of the *p*-butylphenyl group with one or more of these residues. Interestingly, one such residue in pol α is non-aromatic, i.e. Met-1007 (Fig. 2). If the secondary structure of region I is indeed in an anti-parallel β-sheet conformation as depicted, we propose that the Met-1007:*p*-butylphenyl interaction may be partly responsible for the unusually high affinity of BuPdGTP for pol α compared to all other B family enzymes, and is perhaps related to the inability of pol α to incorporate BuPdGTP. The effect of a site-directed mutation of Met-1007 in pol α to an aromatic residue on BuPdGTP sensitivity might support this hypothesis. Conversely, the analogous residue in T4 pol, Tyr-623, could be converted to Met and the resulting polymerase analyzed for increased affinity and/or loss of ability to incorporate BuPdGTP. Experiments with the Y623M mutant of T4 pol are underway in this laboratory.

Acknowledgments

The authors thank Dr. John K. Young for assistance in computer modelling and graphics. The authors are grateful to collaborators for testing of inhibitors and sharing of information. Work in the authors' laboratory was supported by National Institutes of Health grant GM21747.

References

1. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA (1985) *Nature* (Lond) **313**: 762-766
2. Beese LS, Derbyshire V, Steitz TA (1993) *Science* **260**: 352-355
3. Beese LS, Friedman JM, Steitz TA (1993) *Biochemistry* **32**: 14095-14101
4. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J (1994) *Science* **264**: 1891-1903
5. Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut J (1994) *Science* **264**: 1930-1935
6. Steitz TA, Smerdon SJ, Jäger J, Joyce CM (1994) *Science* **266**: 2022-2025
7. Pelletier H (1994) *Science* **266**: 2025-2026
8. Ito J, Braithwaite DK (1991) *Nucleic Acids Res* **19**: 4045-4057
9. Bernad A, Blanco L, Lázaro JM, Martin G, Salas M (1989) *Cell* **59**: 219-228
10. Lindborg B (1992) *Antiviral Chem & Chemother* **3**: 223-241
11. Wright GE, Brown NC (1990) *Pharmac Ther* **47**: 447-497
12. Öberg B (1989) *Pharmac Ther* **40**: 213-285
13. Wright GE, Dudycz LW (1984) *J Med Chem* **27**: 175-181
14. Khan NN, Wright GE, Dudycz LW, Brown NC (1984) *Nucleic Acids Res* **12**: 3695-3705
15. Khan NN, Wright GE, Brown NC (1991) *Nucleic Acids Res* **19**: 1627-1632
16. Blasco MA, Lázaro JM, Bernad A, Blanco L, Salas M (1992) *J Biol Chem* **267**: 19427-19434
17. Khan NN, Reha-Krantz LJ, Wright GE (1994) *Nucleic Acids Res* **22**: 232-237
18. Delarue M, Poch O, Tordo N, Moras D, Argos P (1990) *Protein Eng* **3**: 461-467
19. Dorsky DI, Crumpacker CS (1990) *J Virology* **64**: 1394-1397
20. Marcy AL, Hwang CB, Ruffner KL, Coen DM (1990) *J Virology* **64**: 5883-5890
21. Jung G, Leavitt MC, Schultz M, Ito J (1990) *Biochem Biophys Res Commun* **170**: 1294-1300
22. Bernad A, Lázaro JM, Salas M, Blanco L (1990) *Proc Natl Acad Sci USA* **87**: 4610-4614
23. Bernad A, Blanco L, Salas M (1990) *Gene* **94**: 45-51
24. Copeland WC, Wang TS-F (1993) *J Biol Chem* **268**: 11028-11040
25. Gibbs JS, Chiou HC, Bastow KF, Cheng YC, Coen DM (1988) *Proc Natl Acad Sci USA* **85**: 6672-6676
26. Reha-Krantz LJ, Nonay RL, Stocki S (1993) *J Virology* **67**: 60-66
27. Wright GE, Reha-Krantz LJ, Khan NN, manuscript in preparation
28. Blasco MA, Lázaro JM, Blanco L, Salas M (1993) *J Biol Chem* **268**: 24106-24113
29. Dong Q, Copeland WC, Wang TS-F (1993) *J Biol Chem* **268**: 24163-24174
30. Dong Q, Copeland WC, Wang TS-F (1993) *J Biol Chem* **268**: 24174-24182
31. Pandey VN, Kaushik N, Modak MJ (1994) *J Biol Chem* **269**: 13258-13265
32. Zhu W, Leavitt MC, Jung G, Ito J (1994) *Biochim Biophys Acta* **1219**: 260-266
33. Hall JD, Wang Y, Pierpont J, Berlin MS, Rundlett SE, Woodward S (1989) *Nucleic Acids Res* **17**: 9231-9244

Chemotherapy of human immunodeficiency virus (HIV) infection based on chemotherapeutic intervention with early steps of the virus replicative cycle

ERIK DE CLERCQ*

Contents:

Abbreviations used: *Enzyme and inhibitors*: RT — reverse transcriptase; NNRTI — non-nucleoside reverse transcriptase inhibitor. *Viruses*: CMV — cytomegalovirus; FIV — feline immunodeficiency virus; HBV — human hepatitis B virus; HSV — herpes simplex virus; MAIDS — murine AIDS virus; MSV — murine sarcoma virus; RSV — respiratory syncytial virus; SIV — simian immunodeficiency virus; VZV — vesicular stomatitis virus. *Other abbreviations*: ANP — acyclic nucleoside phosphonate; α-APA — α-anilinophenylacetamide; ATA — aurintricarboxylic acid; AZT — 3′-azido-2′,3′-dideoxythymidine; bis(POM) — bis(pivaloyloxymethyl); DDC — 2′,3′-dideoxycytidine; DDI — 2′,3′-dideoxyinosine; D4T — 2′,3′-didehydro-2′,3′-dideoxythymidine; FPMPA — 9-(3-fluoro-2-phosphonylmethoxypropyl)adenine; FPMPDAP, PMEDAP and PMPDAP — 2,6-diaminopurine counter parts of FPMPA, PMEA and PMPA, respectively; HEPT — 1-(2-hydroxyethoxymethyl)-6-(phenylthio)thymine; MOI — multiplicity of infection; PAPS — polyacetal polysulfate; PAS — polyanetholesulfonate; PMEA — 9-(2-phosphonylmethoxyethyl)adenine; PMPA — 9-(2-phosphonylmethoxypropyl) adenine; PVS — polyvinylsulfonate; PSS — polystyrenesulfonate; 3TC — (-)2′,3′-dideoxy-3′-thiacytidine; TIBO — tetrahydroimidazo(4,5,1-*jk*)(1,4)benzodiazepin-2(1*H*)-one; TSAO — 2′,5′-bis-*O*-(*tert*-butyldimethylsilyl)-3′-spiro-5″-(4″-amino-1″,2″-oxathiole-2″,2″-dioxide)-pyrimidine or -purine.

I. Introduction

The replicative cycle of human immunodeficiency virus (HIV) can be interrupted at several stages: **(i)** virus-cell binding, by polyanionic substances (polysulfates, -sulfonates, -carboxylates and -oxometalates); **(ii)** virus-cell fusion, by plant lectins, negatively charged albumins and triterpene (i.e. betulinic acid) derivatives; **(iii)** virus uncoating, by bicyclams; **(iv)** reverse transcription, by both substrate (i.e. dideoxynucleoside and acyclic nucleoside phosphonate) analogues and non-substrate analogues (i.e. non-nucleoside reverse transcriptase inhibitors (NNRTIs)); **(v)** integration; **(vi)** DNA replication; **(vii)** DNA → mRNA transcription; **(viii)** mRNA translation, by antisense oligonucleotides and ribozymes; **(ix)** maturation, by proteolysis, myristoylation and glycosylation inhibitors; and, finally, **(x)** assembly/release (budding). Here I will focus on the chemotherapeutic interventions at the first four steps, being virus adsorption, fusion, uncoating and reverse (RNA → DNA) transcription.

II. Virus-cell binding

Adsorption of HIV to the host cell depends on the interaction between the viral envelope glycoprotein gp120 and the cellular CD4 receptor (and/or heparan sulfate at the cell surface), and this interaction can be blocked by various polyanionic substances that interact with either gp120 or CD4, or both, and that belong to any of the following classes [1]: polysulfates [2, 3], polysulfonates, polycarboxylates or polyoxometalates. Within each of these four classes of compounds we have found several congeners to inhibit HIV replication in cell culture at concentrations that are at least 1,000-fold, if not 10,000-fold, below the cytotoxicity threshold. This includes **(i)** the **polysulfates** PAPS (polyacetal polysulfate) [4], O-acylated heparin and dermatan sulfate derivatives [5], modified cyclodextrin sulfates [6], and sulfated polysaccharides derived from seaweeds (i.e. *Nothogenia fastigiata* [7] and *Aghardhiella tenera* [8]; **(ii) polysulfonates**, i.e. suramin, PVS (polyvinylsulfonate), PSS (polystyrenesulfonate), PAS (poly-anetholesulfonate) [9] and bis-naphthalenedisulfonates [10]; **(iii) polycarboxylates**, i.e. ATA (aurintricarboxylic acid) [11] and derivatives thereof (i.e.

* Rega Institute for Medical Research, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

hybrid molecules of ATA with cholesterol), and polyhydroxycarboxylates; and (**iv**) **polyoxometalates** [12] of either the Keggin type (i.e. $K_6BGaW_{11}O_{39}$), or Dawson type (i.e. $K_{12}H_2P_2W_{12}O_{48}$) or Keggin sandwich type (i.e. $K_{13}Ce(SiW_{11}O_{39})_2$). These polyanionic substances not only inhibit HIV infection by virtue of their interference with virus binding to the cells, but they also inhibit syncytium formation between virus-infected ($gp120^+$) cells and uninfected ($CD4^+$) cells. Consequently, the polyanionic substances protect $CD4^+$ cells against destruction that would otherwise result from their fusion with $gp120^+$ cells. The polyanionic substances may be expected to prevent HIV infection by blocking both virus-to-cell and cell-to-cell spread, and, in this sense they may be useful as vaginal microbicides in the prevention of heterosexual HIV transmission. Moreover, polyanionic substances have been shown to inhibit various viruses other than HIV: i.e. herpesviruses (herpes simplex virus (HSV), cytomegalovirus (CMV)), myxoviruses (influenza A, respiratory syncytial virus (RSV)), rhabdoviruses (vesicular stomatitis virus (VZV)), arena-, toga, flavi- and bunyaviruses, and this broad-spectrum antiviral activity considerably widens the scope of their therapeutic applicability.

III. Virus-cell fusion

The following compounds seem to be specifically targeted at the **virus-cell fusion process** (since they are highly inhibitory to syncytium formation without markedly affecting virus binding to the cells: mannose-specific **lectins** [13] from orchid species *Cymbidium* hybrid, *Epipactis helleborine* and *Listera ovata*; succinylated [14] and aconitylated [15] **albumins** (obtained through derivatization of human serum albumin with succinic anhydride or *cis*-aconitic anhydride); and triterpene (i.e. **betulinic acid**) derivatives [16]. Whereas the plant lectins are inhibitory to HIV-1, HIV-2 and other enveloped viruses such as CMV and influenza, the succinylated/aconitylated albumins, and the betulinic acid derivatives, are only (or mainly) active against HIV-1, but not HIV-2 or other (retro)viruses. Even within the HIV-1 family, some strains have been found to be insensitive to the inhibitory effects of the betulinic acid derivatives, which indicates that these compounds must interfere in a highly specific manner with the virus-cell fusion process.

IV. Virus fusion/uncoating

While originally discovered as an "impurity" in a commercially available preparation of cyclam, the **bicyclam** derivatives now rank among the most potent and selective inhibitors of HIV replication that have ever been described. The first "lead" compound, 1,1′-propylenebis(1,4,8,11-tetraazacyclotetradecane), showed anti-HIV activity in the 0.1-1 micromolar range [17]. Through lead optimization bicyclam derivatives, i.e. 1,1′-[1,4-phenylenebis(methylene)]-bis(1,4,8,11-tetraazacyclotetradecane) [18], have been obtained that show anti-HIV activity in the 1-5 nM concentration range, without cytotoxicity at 0.5 mM, which means a selectivity index of 10^5 or higher. The bicyclams are equally active against HIV-1 and HIV-2. They represent the first compounds reported to interfere with the **fusion/uncoating process** of HIV. This unique mode of action was suggested by "time of addition" experiments, whereby the cells are initially infected at a high multiplicity of infection (MOI) and the compounds are added at varying times, i.e. 0, 1, 2, 3, ... up to 24 hours after infection. For the bicyclams, the stage of interaction appeared to be the viral uncoating event, and this hypothesis was further corroborated by uncoating assays in which ribonuclease sensitivity of radiolabeled virions (containing [5-^{3}H]uridine labeled RNA) was monitored: the bicyclams were found to protect RNA from degradation by ribonuclease, as could be expected if virion uncoating, that is dissociation of viral proteins from the viral RNA, was blocked [19]. The bicyclams also inhibit HIV-induced syncytium formation, but only at concentrations that are about 100-fold higher than those required for inhibition of HIV replication. This means that they may not only be inhibitory to viral uncoating *per se*, but also interfere with the virus-cell fusion process that actually precedes, and is associated with, viral uncoating. As a putative target for the interaction of the bicyclams we have recently identified a highly conserved portion in the proximity of the V3 and V4 loops of the gp120 glycoprotein.

V. Reverse transcription

V-1. Substrate analogues

Of all possible target molecules within the HIV replicative cycle, the **reverse transcriptase** has attracted most attention. The well established anti-HIV agents 3′-azido-2′,3′-dideoxythymidine (AZT), 2′,3′-dideoxyinosine (DDI), 2′,3′-dideoxycytidine (DDC), 2′,3′-didehydro-2′,3′-dideoxythymidine (D4T) and (−)2′,3′-dideoxy,3′-thiacytidine (3TC) interact, following their intracellular phosphorylation to the corresponding 5′-triphosphates, with the substrate binding site of the HIV reverse transcriptase (RT). We have discovered that **acyclic nucleoside phosphonate** (ANP) analogues [20], i.e. PMEA [9-(2-phosphonylmethoxyethyl)adenine], FPMPA [9-(3-fluoro-2-phosphonylmethoxypropyl)adenine], PMPA [9-(2-phosphonylmethoxypropyl)adenine] [21] and their 2,6-diaminopurine counterparts PMEDAP, FPMPDAP and PMPDAP are selective and potent inhibitors of the replication of HIV-1, HIV-2 and other retroviruses: i.e. murine sarcoma virus (MSV), murine AIDS (MAIDS) virus, simian immunodeficiency virus (SIV), feline im-

munodeficiency virus (FIV) [22] and visna/meadi virus (sheep) [23]. They are also inhibitory to hepadnaviruses (i.e. human hepatitis B virus (HBV) and duck hepatitis virus [24, 25]) by virtue of their inhibitory effect on hepatitis B virus DNA synthesis [26]. The ANP analogues can as such be taken up by the cells; they are phosphorylated intracellularly to their diphosphate derivatives which then act as competitive inhibitors/alternate substrates in the HIV RT reaction. If incorporated into the growing DNA chain, the ANP analogues terminate the chain growth and thus act as DNA chain terminators.

The ANP analogues rank among the most efficacious *in vivo* inhibitors of retrovirus infections that have ever been described. They are markedly more active than AZT in inhibiting MSV infection in mice, FIV infection in cats and SIV infection in monkeys [27]. They are also effective against visna virus infection in sheep (lambs) [28] and duck hepatitis B virus infection in ducks [24, 25]. PMEA has also proved inhibitory to HIV replication in humans, and the bis(pivaloyloxymethyl) ester of PMEA (bis(POM) PMEA), which has better oral bioavailability than PMEA itself is now being pursued as an oral therapeutic modality for the treatment of both HIV and HBV infections. A unique property of PMEA and all ANP analogues is that they afford a prolonged antiviral response, lasting for several days or even weeks. This long-lasting antiviral activity can be attributed to the long half-life of the ANP metabolites.

V-2. Non-substrate analogues

An entirely new lead [29, 30] for the chemotherapy of HIV-1 infection is represented by the **TIBO** [tetrahydroimidazo(4,5,1-*jk*) (1,4)-benzodiazepin-2(1*H*)-one] [31], **HEPT** [1-(2-hydroxyethoxymethyl)-6-(phenylthio)thymine] derivatives (i.e. 6-benzyl-1-ethoxymethyl-5-isopropyluracil [32]), **TSAO** [2′,5′-bis-*O*-(*tert*-butyldimethylsilyl)-3′-spiro-5″-(4″-amino-1″,2″-oxathiole-2″,2″-dioxide)-pyrimidine or -purine] [33-36] and **α-APA** (α-anilinophenylacetamide) derivatives [37]. The most active congeners inhibit the replication of HIV-1 at nanomolar concentrations that are 10^5-fold lower than the cytotoxic concentrations, thus achieving a selectivity index of 10^5. Neither TIBO, nor HEPT, or TSAO or α-APA inhibit the replication of viruses other than HIV-1. Through "time of addition" experiments, it has become clear that TIBO, HEPT, TSAO and α-APA must interfere with the reverse transcription process. They do so by a mechanism of action that is totally different from that of the dideoxynucleoside analogues (AZT, DDI, DDC, D4T or 3TC) or acyclic nucleoside phosphonates (PMEA, FPMPA, PMPA), which interact with the substrate (dNTP) binding site of HIV-1 reverse transcriptase [38, 39]. TIBO and its congeners interact allosterically with a non-substrate binding site of the **HIV-1 reverse transcriptase** and thus inactivate the enzyme non-competitively with regard to the natural substrate. The molecular coordinates of the RT binding site for the TIBO class of compounds, which are now collectively referred to as "non-nucleoside reverse transcriptase inhibitors" (NNRTIs), have become increasingly clear, primarily through the identification of HIV-1 RT mutations that confer resistance to the NNRTIs, combined with elucidation of the three-dimensional structure of HIV-1 RT [40].

VI. HIV-1 resistance to NNRTIs

HIV-1 resistance to NNRTIs readily develops upon passage of the virus in cell culture in the presence of the compounds [41]. This resistance is due to point mutations in the HIV-1 RT gene, as has been confirmed by site-directed mutagenesis. Hence, the following RT mutations account for resistance to TIBO, HEPT, TSAO and α-APA: 100 Leu → Ile (TIBO^r); 103 Lys → Asn (TIBO^r); 106 Val → Ala (HEPT^r); 138 Glu → Lys (TSAO^r); 179 Val → Asp (TIBO^r); 181 Tyr → Cys (TIBO^r, HEPT^r, α-APA^r); 188 Tyr → His (TIBO^r, HEPT^r) [42]. For other NNRTIs, such as the quinoxaline derivatives, resistance development is based on mutations at position 106 (Val → Ala), 181 (Tyr → Cys) and 190 (Gly → Glu) [43]. All these mutations, including the 181 Tyr → Cys mutation but with the exception of the 138 Glu → Lys mutation, must reside in the p66 subunit of HIV-1 RT to confer resistance. The 138 Glu → Lys (TSAO^r) mutation must be located on the p51 subunit to induce resistance (to TSAO) [44]. Some of these mutations, i.e. 181 Tyr → Cys, have also been shown to occur *in vivo* in the clinic, but the significance (i.e. pathogenicity, reversibility and transmissibility) of the mutant strains and their role in disease progression remains to be assessed.

As future strategies [42] to overcome, prevent or circumvent the development of HIV-1 resistance to NNRTIs, we could envisage the following: **(i)** if resistance develops to one NNRTI, switch to another NNRTI to which the virus has remained sensitive; **(ii)** combine drugs with mutually antagonistic resistance patterns, i.e. drugs that lead to resistance mutations which counteract each other (i.e. TIBO (181 Tyr → Cys) with AZT (215 Thr → Tyr)); **(iii)** use from the beginning sufficiently high ("knocking out") concentrations of the NNRTIs, so as to completely suppress virus replication and thus prevent any virus breakthrough; and **(iv)** in combining strategies **(ii)** and **(iii)**, use from the start sufficiently high concentrations of different compounds in combination.

HIV-1 RT mutations that impart resistance to some NNRTIs not necessarily lead to cross-resistance to other NNRTIs: for example, the TSAO^r (138 Glu → Lys) mutant remains fully sensitive to various other NNRTIs, including TIBO, HEPT, α-APA and quinoxaline derivatives [45]. The widest cross-resis-

tance is conferred by the 181 Tyr → Cys mutation. The 181 Tyr → Cys mutant can, under continuous pressure of the NNRTIs, mutate further to the highly resistant 181 Cys → Ile mutant, but, even so, it retains marked sensitivity to some of the HEPT derivatives [46].

If the NNRTIs are used from the beginning at sufficiently high concentrations, which for some of the NNRTIs (i.e. quinoxaline and thiocarboxanilide derivatives) do not have to be in excess of 0.1-0.5 μg/ml [43, 47]), they are able to completely suppress ("knock out") virus replication [48, 49], and thus prevent the breakthrough of any virus, whether resistant or not. If two or more of the NNRTIs are combined from the start (strategy **iv**), virus breakthrough can be prevented for a much longer time, and at much lower drug concentrations than if the compounds are used singly. This has been demonstrated with combinations of TSAO and thiocarboxanilide derivatives with one another, and with combinations of these NNRTIs with 3TC [50].

VII. Conclusion

Several HIV inhibitors (for example, bicyclams and NNRTIs) have been found to inhibit HIV replication at nanomolar concentrations that are ⩾ 100,000-fold lower than the concentrations required to impair the viability of the uninfected host cells. When applied to the HIV-infected cells at sufficiently high, but subtoxic, concentrations, these compounds are able to completely suppress ("knock out") virus replication, so that the cell cultures become apparently "cleared" from the infection. When combined with each other, these compounds may knock out the virus, and prevent the breakthrough of virus (whether resistant or not), at even lower concentrations, and for longer time periods than if the drugs are used individually.

Acknowledgments

The original investigations are supported in part by the Biomedical Research Programme of the European Community, the Belgian Nationaal Fonds voor Wetenschappelijk Onderzoek, the Belgian Fonds voor Geneeskundig Wetenschappelijk Onderzoek, the Belgian Geconcerteerde Onderzoeksacties and the Janssen Research Foundation. I thank C. Callebaut for excellent editorial assistance.

References

1. De Clercq E (1993) In: Morrow WJW, Haigwood NL (eds) HIV Molecular Organization, Pathogenicity and Treatment. Elsevier Science Publishers, Amsterdam, pp 267-292
2. De Clercq E (1993) In: Yalpani M (ed) Carbohydrates and Carbohydrate Polymers, Analysis, Biotechnology, Modification, Antiviral Biomedical and Other Applications. ATL Press, pp 87-100
3. Witvrouw M, Desmyter J, De Clercq E (1994) *Antiviral Chem Chemother* **5**: 345-359
4. Witvrouw M, Schols D, Andrei G, Snoeck R, Ikeda S, Pauwels R, Van Schepdael A, Arnout J, Claes P, Desmyter J, De Clercq E (1992) *Antiviral Chem Chemother* **3**: 351-360
5. Bârzu T, Level M, Petitou M, Lormeau J-C, Choay J, Schols D, Baba M, Pauwels R, Witvrouw M, De Clercq E (1993) *J Med Chem* **36**: 3546-3555
6. Otake T, Schols D, Witvrouw M, Naesens L, Nakashima H, Moriya T, Kurita H, Matsumoto K, Ueba N, De Clercq E (1994) *Antiviral Chem Chemother* **5**: 155-161
7. Damonte E, Neyts J, Pujol CA, Snoeck R, Andrei G, Ikeda S, Witvrouw M, Reymen D, Haines H, Matulewicz MC, Cerezo A, Coto CE, De Clercq E (1994) *Biochem Pharmacol* **47**: 2187-2192
8. Witvrouw M, Este JA, Mateu MQ, Reymen D, Andrei G, Snoeck R, Ikeda S, Pauwels R, Bianchini NV, Desmyter J, De Clercq E (1994) *Antiviral Chem Chemother* **5**: 297-303
9. Mohan P, Schols D, Baba M, De Clercq E (1992) *Antiviral Res* **18**: 139-150
10. Baba M, Schols D, Mohan P, De Clercq E, Shigeta S (1993) *Antiviral Chem Chemother* **4**: 229-234
11. Cushman M, Wang P, Stowell JG, Schols D, De Clercq E (1992) *J Org Chem* **57**: 7241-7248
12. Yamamoto N, Schols D, De Clercq E, Debyser Z, Pauwels R, Balzarini J, Nakashima H, Baba M, Hosoya M, Snoeck R, Neyts J, Andrei G, Murrer BA, Theobald B, Bossard G, Henson G, Abrams M, Picker D (1992) *Mol Pharmacol* **42**: 1109-1117
13. Balzarini J, Neyts J, Schols D, Hosoya M, Van Damme E, Peumans W, De Clercq E (1992) *Antiviral Res* **18**: 191-207
14. Jansen RW, Molema G, Pauwels R, Schols D, De Clercq E, Meijer DKF (1991) *Mol Pharmacol* **39**: 818-823
15. Jansen RW, Schols D, Pauwels R, De Clercq E, Meijer DKF (1993) *Mol Pharmacol* **44**: 1003-1007
16. Mayaux J-F, Bousseau A, Pauwels R, Huet T, Hénin Y, Dereu N, Evers M, Soler F, Poujade C, De Clercq E, Le Pecq J-B (1994) *Proc Natl Acad Sci USA* **91**: 3564-3568
17. De Clercq E, Yamamoto N, Pauwels R, Baba M, Schols D, Nakashima H, Balzarini J, Debyser Z, Murrer BA, Schwartz D, Thornton D, Bridger G, Fricker S, Henson G, Abrams M, Picker D (1992) *Proc Natl Acad Sci USA* **89**: 5286-5290
18. De Clercq E, Yamamoto N, Pauwels R, Balzarini J, Witvrouw M, De Vreese K, Debyser Z, Rosenwirth B, Peichl P, Datema R, Thornton D, Skerlj R, Gaul F, Padmanabhan S, Bridger G, Henson G, Abrams M (1994) *Antimicrob Agents Chemother* **38**: 668-674
19. De Clercq E (1992) *Int J Immunother* **8**: 115-123
20. De Clercq E (1991) *Biochem Pharmacol* **42**: 963-972
21. Balzarini J, Holy A, Jindrich J, Naesens L, Snoeck R, Schols D, De Clercq E (1993) *Antimicrob Agents Chemother* **37**: 332-338
22. Hartmann K, Balzarini J, Higgins J, De Clercq E, Pedersen NC (1994) *Antiviral Chem Chemother* **5**: 13-19
23. Thormar H, Balzarini J, Holy A, Jindrich J, Rosenberg I, Debyser Z, Desmyter J, De Clercq E (1993) *Antimicrob Agents Chemother* **37**: 2540-2544
24. Heijtink RA, Kruining J, de Wilde GA, Balzarini J, De Clercq E, Schalm SW (1994) *Antimicrob Agents Chemother* **38**: 2180-2182
25. Balzarini J, Kruining J, Heijtink R, De Clercq E (1994) *Antiviral Chem Chemother* **5**: 360-365
26. Yokota T, Konno K, Shigeta S, Holy A, Balzarini J, De Clercq E (1994) *Antiviral Chem Chemother* **5**: 57-63
27. Naesens L, Balzarini J, De Clercq E (1994) *Rev Med Virol* **4**: 147-159
28. Thormar H, Georgsson G, Pálsson PA, Balzarini J, Naesens L, Torsteinsdóttir S, De Clercq E (1995) *Proc Natl Acad Sci USA* **92**: 3283-3287
29. De Clercq E (1993) *Med Res Rev* **13**: 229-258
30. De Clercq E (1994) *Ann NY Acad Sci* **724**: 438-456
31. Pauwels R, Andries K, Debyser Z, Kukla MJ, Schols D, Breslin HJ, Woestenborghs R, De-

smyter J, Janssen MAC, De Clercq E, Janssen PAJ (1994) *Antimicrob Agents Chemother* **38**: 2863-2870
32. Baba M, Shigeta S, Yuasa S, Takashima H, Sekiya K, Ubasawa M, Tanaka H, Miyasaka T, Walker RT, De Clercq E (1994) *Antimicrob Agents Chemother* **38**: 688-692
33. Balzarini J, Pérez-Pérez MJ, San-Félix A, Schols D, Perno C-F, Vandamme A-M, Camarasa MJ, De Clercq E (1992) *Proc Natl Acad Sci USA* **89**: 4392-4396
34. Balzarini J, Pérez-Pérez M-J, San-Félix A, Camarasa M-J, Bradhurst IC, Barr PJ, De Clercq E (1992) *J Biol Chem* **267**: 11831-11838
35. San-Félix A, Velázquez S, Pérez-Pérez MJ, Balzarini J, De Clercq E, Camarasa MJ (1994) *J Med Chem* **37**: 453-460
36. Alvarez R, Velázquez S, San-Félix A, Aquaro S, De Clercq E, Perno C-F, Karlsson A, Balzarini J, Camarasa MJ (1994) *J Med Chem* **37**: 4185-4194
37. Pauwels R, Andries K, Debyser Z, Van Daele P, Schols D, Stoffels P, De Vreese K, Woestenborghs R, Vandamme A-M, Janssen CGM, Anné J, Cauwenbergh G, Desmyter J, Heykants J, Janssen MAC, De Clercq E, Janssen PAJ (1993) *Proc Natl Acad Sci USA* **90**: 1711-1715
38. Balzarini J, Hao Z, Herdewijn P, Johns DG, De Clercq E (1991) *Proc Natl Acad Sci USA* **88**: 1499-1503
39. Balzarini J, Holy A, Jindrich J, Dvorakova H, Hao Z, Snoeck R, Herdewijn P, Johns DG, De Clercq E (1991) *Proc Natl Acad Sci USA* **88**: 4961-4965
40. Tantillo C, Ding J, Jacobo-Molina A, Nanni RG, Boyer PL, Hughes SH, Pauwels R, Andries K, Janssen PAJ, Arnold E (1994) *J Mol Biol* **243**: 369-387
41. Balzarini J, Karlsson A, Pérez-Pérez M-J, Vrang L, Walbers J, Zhang H, Öberg B, Vandamme A-M, Camarasa M-J, De Clercq E (1993) *Virology* **192**: 246-253
42. De Clercq E (1994) *Biochem Pharmacol* **47**: 155-169
43. Balzarini J, Karlsson A, Meichsner C, Paessens A, Riess G, De Clercq E, Kleim J-P (1994) *J Virol* **68**: 7986-7992
44. Jonckheere H, Taymans J-M, Balzarini J, Velázquez S, Camarasa M-J, Desmyter J, De Clercq E, Anné J (1994) *J Biol Chem* **269**: 25255-25258
45. Balzarini J, Karlsson A, Vandamme A-M, Pérez-Pérez M-J, Zhang H, Vrang L, Öberg B, Bäckbro K, Unge T, San-Félix A, Velázquez S, Camarasa M-J, De Clercq E (1993) *Proc Natl Acad Sci USA* **90**: 6952-6956
46. Balzarini J, Karlsson A, Sardana VV, Emini EA, Camarasa M-J, De Clercq E (1994) *Proc Natl Acad Sci USA* **91**: 6599-6603
47. Balzarini J, Pérez-Pérez M-J, Velazquez S, San-Félix A, Camarasa M-J, De Clercq E, Karlsson A (1995) *Proc Natl Acad Sci USA*, **92**: 5470-5474
48. Balzarini J, Karlsson A, Pérez-Pérez M-J, Camarasa M-J, De Clercq E (1993) *Virology* **196**: 576-585
49. De Clercq E (1994) *Int J Immunother* **10**: 145-158
50. Balzarini J, Pelemans H, San-Félix A, Camarasa M-J, De Clercq E, Karlsson A (1996) Submitted for publication

Herpesvirus-encoded thymidine kinase in chemotherapy

TOMASZ OSTROWSKI[1],
JOLANTA NOWOTNA[2],
BOŻENNA GOLANKIEWICZ[3]

Contents:

Abbreviations used: HSV — herpes simplex virus; VZV — varicella zoster virus; EBV — Epstein-Barr virus; CMV — human cytomegalovirus; TMP — thymidine monophosphate; TDP — thymidine diphosphate; ACV — acyclovir; ACV-TP — acyclovir triphosphate; GCV — ganciclovir; BVDU — brivudine; PG — N^2-phenylguanine; PhdG — N^2-phenyl-2′-deoxyguanosine; mCF$_3$PG — N^2-[m-(trifluoromethyl)phenyl]guanine; HBPG — 9-(4-hydroxybutyl)-N^2-phenylguanine; HBG — 9-(4-hydroxybutyl)guanine; S-BVDU — (E)-5-(2-bromovinyl)-2′-deoxy-4′-thiouridine.

[1,2,3] Institute of Bioorganic Chemistry of the Polish Academy of Sciences, ul. Z. Noskowskiego 12/14, 61-704 Poznań, Poland

I. Introduction

I-1. Herpesvirus diseases in man [1]

Viruses of the family Herpesviridae, widely disseminated in nature in most of animal species, demonstrate a remarkable ability to remain latent in their natural hosts. Six herpesviruses have been isolated so far from humans. Herpes simplex viruses, type 1 (HSV-1) and type 2 (HSV-2) and varicella zoster virus (VZV) establish latent infections in nerve tissue. About 80% of human population have antibodies to HSV-1 but only 20% show any symptoms. HSV-1 manifests itself primarily in cold sores (labialis), a type of mouth ulcer (gingivostomatitis), and eye disease (keratitis). HSV-2 accounts for most cases of genital infections. Both HSV-1 and HSV-2 may cause encephalitis. VZV causes chickenpox, an infection which is usually mild in children, but severe when contracted by an adult. On reactivation, a latent VZV causes shingles, an extremely painful condition. Epstein-Barr virus (EBV), which remains latent in B lymphocytes, causes glandular fever and mononucleosis. Human herpes virus (HHV-6) is associated with roseola. Human cytomegalovirus (CMV) in most cases is asymptomatic. Many people have antibodies to cytomegalovirus, but never exhibit any manifestation of the illness. However, in immunologically deficient patients (e.g. transplantation, cancer and HIV positive patients), CMV infections give rise to severe pneumonial and colitis conditions and in the case of CMV retinitis, blindness. Activation of latent herpetic viruses in immunocompromised patients is a general phenomenon which causes severe virus infections and strong demand for an effective therapy.

I-2. Herpesvirus-encoded enzymes

Herpesviruses share a significant biological property. They specify a large array of enzymes involved in nucleic acid metabolism (e.g. thymidine kinase, thymidylate synthase, dUTPase, ribonucleotide reductase, etc.), DNA synthesis (e.g., DNA polymerase helicase, primase), and, possibly, processing of proteins (e.g. protein kinase), although the exact array of enzymes may vary somewhat from one herpesvirus to another [2]. Out of that variety of enzymes, the herpetic thymidine kinases (TK-s) have been those most intensively studied. Together with their cellular counterparts, they have become increasingly important for analyses of the regulation of gene expression in eukaryotic cells, as primary targets for antiherpetic drugs and for genetic engineering of recombinant virus vectors and vaccines [3].

Research on the properties, functions, sequence analysis, origin and evolution of viral thymidine kinase and TMP kinase and nucleoside diphosphate kinase has already been reviewed [3, 4]. The following review will focus on the themes relevant to therapy and drug development.

II. The role of viral TK in the development of selective antiviral and antitumor agents

II-1. Substrate specificity of human and viral TK-s

Thymidine kinase (TK) (ATP : thymidine : 5′-phosphotransferase; EC 2.7.1.21) is the crucial enzyme in the pyrimidine salvage pathway. This enzyme catalyzes the transfer of the γ phosphate from ATP to the 5′ OH of thymidine to give thymidine monophosphate (TMP). Viral TK induction in HSV-infected cells was discovered already in 1963 by Kit and Dubbs [5], who predicted that the TK is encoded in the viral genome. This was subsequently proved to be correct by different approaches: genetic analysis of the TK locus, gene transfer experiments, synthesis *in vitro*, and gene cloning [6-10].

Unlike cellular thymidine kinases, corresponding viral enzymes are highly tolerant to the structural deviations from their natural substrates: ATP as a phosphate donor and thymidine as a phosphate acceptor. While the vertebrate cytosol TK-s efficiently utilize only ATP and dATP as phosphate donors, herpesvirus TK enzymes use CTP and GTP as well as ATP [3, 11, 12]. Additionally, the viral TK-s have significant activity as thymidylate kinases and are able to catalyze a second phosphoryl transfer with the formation of thymidine diphosphate (TDP) as well as TMP. Whereas the substrate specificity of cellular TK is essentially limited to thymidine, the viral TK-s are general deoxypirymidine kinases [11, 13, 14]. Moreover, they also accept as substrates purine analogues, even those carrying acyclic sugar moiety such as acyclovir [15, 16].

Substrate specificity differences between the human and viral kinases have become the basis for the development of selective antiviral agents.

II-2. Metabolic activation of nucleoside analogues by viral TK-s

Unusually broad substrate specificity of herpetic TK-s extending to purine derivatives has been found by a mere chance. Acyclovir (ACV, **1**), 9-[(2-hydroxyethoxy)methyl]guanine, the first highly potent and selective antiherpetic drug, has been discovered by antiviral screening of a series of compounds designed originally as substrates and inhibitors of adenosine deaminase [17, 18]. The degree of selectivity of ACV is unusual [15, 19-21]. It is highly active against HSV-1, HSV-2 and VZV. It is active against EBV and pseudorabies, only slightly against CMV and inactive against other DNA viruses, e.g. vaccinia, as well as RNA viruses [19]. It is not cytotoxic to the mammalian cells, in which these viruses are grown, at concentra-

tions that are a hundred times greater than those required for antiviral activity. Studies on the mechanism of action of ACV carried out by Elion *et al.* [15, 19-21] revealed a pivotal role of the viral TK in its selectivity. Acyclovir is converted by the viral TK-s specified by HSV or VZV to its monophosphate derivative. After the first phosphate has been added, diphosphorylation and triphosphorylation are catalyzed by cellular enzymes, resulting in ACV triphosphate (ACV-TP) concentrations that are 200 to 1000 times higher in HSV-infected cells than in uninfected cells [20]. Since cellular TK cannot use ACV as a substrate, very little ACV-TP is formed in uninfected cells [15, 22]. The small amount of phosphorylation that occurs in normal cells is due to a 5′-nucleotidase [23]. Acyclovir triphosphate prevents viral DNA synthesis by inhibiting the viral DNA polymerase. Because ACV-TP lacks the 3′-hydroxyl group required to elongate the DNA chain, the growing chain of DNA is terminated. In the presence of the deoxynucleoside triphosphate complementary to the next template position, the viral DNA polymerase is functionally inactivated [24]. In addition, ACV-TP is a much better substrate for viral polymerase than for cellular DNA polymerase alpha, resulting in little incorporation of ACV into cellular DNA.

Variations in the carbohydrate portion of acyclovir have resulted in a number of active analogues, e.g. ganciclovir (GCV, **2**, formerly called DHPG, BW B759U and 2′NDG) [25-27], buciclovir (BCV, **3**) [28], penciclovir (PCV, **4**) and its oral form famciclovir (FCV, **5**) [29-32]. The mechanism of action of all these compounds is similar to that of acyclovir, i.e. their antiviral activity against HSV and VZV depends on a specific phosphorylation by the virus-induced TK and subsequent selective inhibition of viral DNA polymerase by the respective acyclic nucleoside triphosphate. Unlike the rest of the compounds, ganciclovir is also active against CMV and has been licensed for the treatment of severe CMV infections. CMV does not encode a thymidine kinase. Rather, the CMV enzyme that catalyzes the initial phosphorylation of GCV in CMV-infected cells is the phosphotransferase, a protein kinase homologue, encoded by the UL97 gene [33, 34].

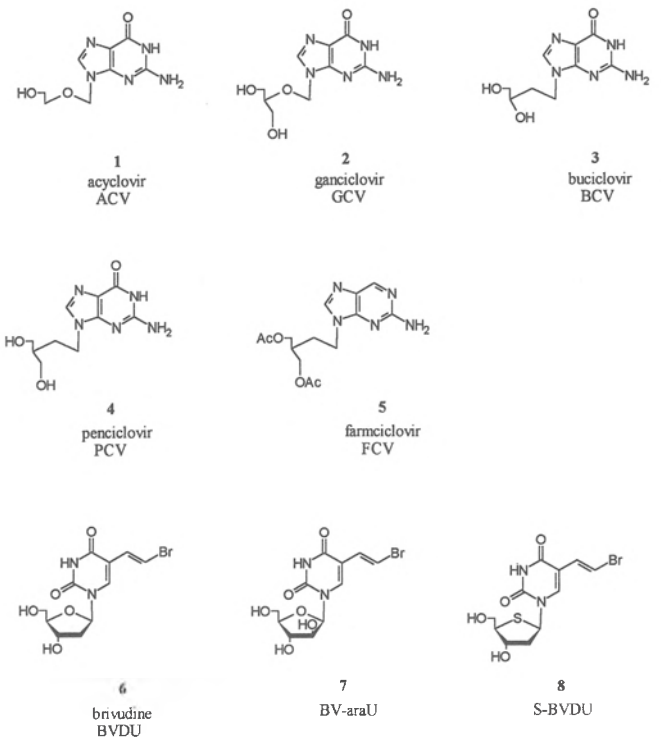
Modification of the guanine moiety of acyclovir and ganciclovir with 1,N^2-etheno bridge lowers the activity against HSV-1, HSV-2, VZV, and CMV by a factor of 10^2 or more. Further substitutions in the resulting ring enhance the antiviral activity and may render the tricyclic analogue fluorescent [35-37].

5-Ethyl-2′-deoxyuridine described by Świerkowski and Shugar in 1969 [38] is the early representative of the other group of antivirals depending for their activity on phosphorylation by the virus-encoded TK-s, namely, 5-substituted-2′-deoxyuridines. Among these analogues, (E)-5-(2-bromovinyl)-2′-deoxyuridine (brivudine, BVDU, **6**) and its congeners, e.g. 1-β-D-arabinofuranosyl (E)-5-(2-bromovinyl)uracil (BV-araU, **7**) and (E)-5-(2-bromovinyl)-2′-deoxy-4′-thiouridine (S-BVDU, **8**) have been developed by De Clercq, Walker, Descamps *et al.* as very potent and selective antiherpetic drugs [39-41].

Unlike the acyclic guanosine analogues which are converted by virus induced TK-s only to their monophosphate, BVDU is converted by some of these enzymes, such as specified by HSV-1 and VZV, also to the 5′-diphosphate.

The present knowledge regarding the specificity of HSV thymidine kinase for its 5-substituted 2′-deoxyuridine substrates has been recently reviewed. It has been concluded that the electronic characteristics of the 5-substituent of 5-substituted 2′-deoxyuridine derivative are important for the interaction with TK. All compounds which are good substrates demonstrate an attractive region above the 5 substituent [42].

Nucleosides analogues - substrates of herpetic TK-s



II-3. Inhibition of HSV TK for suppression of virus reactivation

HSV TK is not required for efficient replication of the virus in dividing cells [11]. However, TK is important for the pathogenicity of the virus in humans [43] and its reactivation from the latent state. A strategy for the control of herpetic recurrences is to interfere with the reactivation process by targeting inhibitors against virus-encoded TK. Along this line, several groups have developed HSV TK inhibitors as potential

antiviral drugs [44-46]. HSV TK inhibitors exemplified by compounds **9-12** are either guanine or thymine analogues and so they may be expected to bind at the same site as substrates of the enzyme.

Extensive studies of the Wright's group investigating the selective inhibitory activity of N^2-phenylguanine (PG) derivatives towards the HSV-1 and HSV-2 TK-s relative to the human enzyme *in vitro* and *in vivo* [46-49] resulted in two most potent inhibitors N^2-phenyl-2′-deoxyguanosine (PhdG, **9**) and N^2-[*m*-(trifluoromethyl)phenyl]guanine (*m*CF_3PG, **10**). Both compounds diminished the reactivation of latent HSV from explanted murine ganglia [49].

Very recently, 9-(4-hydroxybutyl)-N^2-phenylguanine (HBPG, **11**) has been described. Being a potent inhibitor with both good water and lipid solubility, favourable pharmacokinetics and lack of acute toxicity in mice, it is a good candidate for animal trials of the ability of TK inhibitors to prevent recurrent herpes virus infections. Similar to PhdG [46], HBPG (**11**) is a competitive non substrate inhibitor of HSV-1 TK. The inability of the enzyme to phosphorylate HBPG is interesting in view of the reported properties of a closely related compound 9-(4-hydroxybutyl)guanine (HBG) [50]. The latter, an isostere of acyclovir, is a substrate for HSV-1 TK and possesses antiherpetic activity. If it is assumed that both compounds bind viral TK-s at the same site, the presence of N^2 phenyl ring of HBPG has two consequences. First, the interaction of the phenyl ring results in a greater affinity of HPBG in comparison with HBG which indicates that the N^2 phenyl ring is a major site of inhibitor-enzyme binding. Second, interaction of the phenyl group with the enzyme blocks the phosphorylation of the hydroxy group in the 9-substituent by an unknown mechanism.

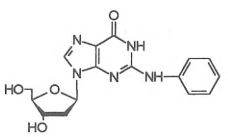
II-4. Cytostatic activity of antiherpetic drugs against HSV TK gene-transfected tumor cells

Very recent developments are extending the importance of herpetic thymidine kinases and nucleoside analogues being their substrates to additional field of therapy—treatment of localized tumors by direct *in situ* introduction of endogenous genes into proliferating tumors. It has been found that cells modified to contain HSV TK gene become sensitive to treatment with the antiviral agent ganciclovir (GCV), whereas normal cells are unaffected by this agent [51]. In an experiment designed to fight brain tumors, rats with cerebral glioma were given an intratumoral injection of murine fibroblasts that were producing a retroviral vector in which the HSV-1 TK gene had been inserted. Subsequent treatment with ganciclovir resulted in tumor regression [52, 53]. A clinical protocol named "molecular surgery" has been worked out to investigate the possibility of using this gene/chemotherapy approach in the treatment of some inoperable brain tumors in humans [54].

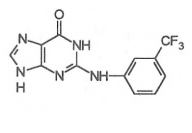
The cytotoxicity acquired by ganciclovir against tumor cells that have been transfected with HSV TK gene is not the only mechanistic aspect of tumor regression. Significant antitumor immunity which follows HSV-TK and GCV treatments suggests that the immune system plays a critical role in the sustained tumor regressions associated with these treatments [55].

Intensive studies on the differential mechanism of cytotoxic effect of various antiherpetic drugs on tumor (murine mammary carcinoma FM3A) cells transfected by the thymidine kinase genes of herpes simplex virus type 1 or type 2 are being carried out by Balzarini, De Clercq *et al.* [56-61]. In their most recent work [62], the authors have recapitulated the results. Based on the potency and mechanism of cytostatic action, the antiherpetic compounds can be divided into two different classes. One encompasses (E)-5-(2-bromovinyl)-2′-deoxyuridine (**6**) and structurally related analogues there of (i.e., the cytosine derivative (E)-5-(2-bromovinyl)-2′-deoxycytidine and the 4′-thio derivative (E)-5-(2-bromovinyl)-2′-deoxy-4′-thiouridine (**8**)). These compounds are highly cytostatic against FM3A/TK$^-$/HSV-1 TK$^+$ and FM3A/TK$^-$/HSV-2 TK$^+$ cells (50% inhibitory concentrations ranging from 0.047 to 0.001 μM) and inhibit cell proliferation by inhibiting cellular thymidylate synthase. The other class consists of the acyclic guanosine derivatives ganciclovir (**2**), buciclovir (**3**), and penciclovir (**4**). These compounds are also more inhibitory to the HSV-1 TK or HSV-2 TK gene-transfected FM3A cells than to FM3A/O or FM3A/TK$^-$ cells, but only at concentrations that are higher than the concentration at which the (E)-5-(2-bromovinyl)-2′-deoxyuridine derivatives proved to be inhibitory. These acyclic guanosine analogues appear to be targeted at the cellular DNA polymerase. (E)-5-(2-bromovinyl)-2′-deoxy-4′-thiouridine seems to be a promising candidate compound for the treatment of HSV-1 TK gene-transfected tumors *in vivo* due to its metabolic stability (i.e. resistance to hydrolysis by thymidine phosphorylase).

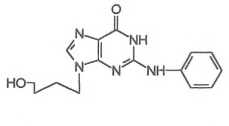
Nucleosides analogues - inhibitors of herpetic TK-s



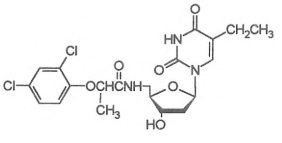
9
PhdG



10
mCF_3PG



11
HBPG



12

III. Perspectives

Acyclovir, discovered by screening, is still considered the antiviral drug of choice for treatment of HSV and VZV infections [63, 64].

The key to the rational drug design of more effective antivirals depending on virus-encoded thymidine kinase is a better understanding of the structure of active site and mode of the interaction of TK proteins with their substrates. Intensive research is carried out in that area. As up to now neither X-ray nor NMR data are available, other approaches are being developed based on site-directed mutagenesis and mutations connected with antiviral resistance. Two recent studies: one on computer-aided active-site-directed modeling of HSV-1 TK [65, 66] and the other on mutant enzymes obtained by selection from a plasmid with random sequence substitution [67] seem to open new perspectives.

Acknowledgment
This work was supported by KBN grant No. 025/P05/95/08.

References

1. Lapucci A, Macchia M, Parkin A (1993) *Il Farmaco* **48:** 871-895
2. Roizman B (1990) In: Fields BN, Knipe DM (eds) Fundamental Virology. Raven Press, New York, pp 844-845
3. Kit S (1985) *Microbiol Sci* **2:** 369-375
4. Gentry GA (1992) *Pharmacol Ther* **54:** 319-355
5. Kit S, Dubbs DR (1963) *Biochem Biophys Res Commun* **11:** 55-59
6. Summers WP, Wagner M, Summers EC (1975) *Proc Natl Acad Sci USA* **72:** 4081-4084
7. Wigler M, Silverstein S, Lee LS, Pelicer A, Cheng YC, Axel R (1977) *Cell* **11:** 223-232
8. Smiley JR, Steege DA, Juricek DK, Summers WP, Ruddle FH (1978) *Cell* **5:** 455-468
9. Wagner MJ, Sharp JA, Summers WC (1981) *Proc Natl Acad Sci USA* **78:** 1441-1445
10. Waldman AS, Husslein E, Milman G (1983) *J Biol Chem* **258:** 11571-11575
11. Jamieson A, Gentry GA, Subak-Sharpe JH (1974) *J Gen Virol* **24:** 465-480
12. Cheng YC (1976) *Biochim Biophys Acta* **452:** 370-381
13. Chen MS, Prusoff WH (1978) *J Biol Chem* **253:** 1325-1327
14. Cooper GM (1973) *Proc Natl Acad Sci USA* **70:** 3768-3792
15. Elion GB, Furman PA, Fyfe JA, de Miranda P, Beauchamp L, Schaeffer HJ (1977) *Proc Natl Acad Sci USA* **74:** 5716-5720
16. Fyfe JA, Keller PM, Furman PA, Miller RL, Elion GB (1978) *J. Biol Chem* **253:** 8721-8727
17. Schaeffer HJ, Gurwara S, Vince R, Bittner S (1971) *J Med Chem* **14:** 367-369
18. Elion GB (1989) *Science* **244:** 41-47
19. Schaeffer HJ, Beauchamp L, de Miranda P, Elion GB, Bauer DJ, Collins P (1978) *Nature* (Lond) **272:** 583-585
20. Elion GB (1982) *Am J Med* **73**(1A): 7-13
21. Elion GB (1983) *J Antimicrob Chemother* **12:** 9-17
22. Furman PA, de Miranda P, St Clair MH, Elion GB (1981) *Antimicrob Agents Chemother* **20:** 518-524
23. Keller PM, Mc Kee SA, Fyfe JA (1985) *J Biol Chem* **260:** 8664-8667
24. Reardon JE, Spector T (1989) *J Biol Chem* **264:** 7405-7411
25. Ogilvie KK, Cherujan UD, Radatus BK, Smith KO, Galloway KS, Kennel WL (1982) *Can J Chem* **60:** 3005-3010
26. Martin JC, Dvorak CA, Smee DF, Matthews TR, Verheyden JPH (1983) *J Med Chem* **26:** 759-761
27. Field AK, Davies ME, De Witt C, Perry HC, Lion R, Gemerhausen J, Karkas JD, Ashton WT, Johnston DBR, Tolman RL (1983) *Proc Natl Acad Sci USA* **80:** 4139-4143
28. Larsson A, Oberg B, Alenius S, Hagberg CE, Johansson NG, Lindborg B, Stening G (1983) *Antimicrob Agents Chemother* **23:** 664-670
29. Harnden MR, Jarvest RL, Bacon TH, Boyd MR (1987) *J Med Chem* **30:** 1636-1642
30. Boyd MR, Bacon TH, Sutton D, Cole M (1987) *Antimicrob Agents Chemother* **31:** 1238-1242
31. Boyd MR, Safrin S, Kern ER (1993) *Antiviral Chem Chemother* **4** (Suppl 1): 3-11
32. Vere Hodge RA (1993) *Antiviral Chem Chemother* **4:** 67-84
33. Littler E, Stuart AD, Chee MS (1992) *Nature* (Lond) **358:** 160-162
34. Sullivan V, Talarico CL, Stanat SC, Davis M, Coen DM, Biron KK (1992) *Nature* (Lond) **358:** 162-164
35. Boryski J, Golankiewicz B, De Clercq E (1988) *J Med Chem* **31:** 1351-1355
36. Boryski J, Golankiewicz B, De Clercq E (1991) *J Med Chem* **34:** 2380-2383
37. Golankiewicz B, Ostrowski T, Andrei G, Snoeck R, De Clercq E (1994) *J Med Chem* **37:** 3187-3190
38. Świerkowski M, Shugar D (1969) *J Med Chem* **12:** 533-534
39. De Clercq E, Descamps J, De Somer P, Barr PJ, Jones AS, Walker RT (1979) *Proc Natl Acad Sci USA* **76:** 2947-2951
40. Descamps J, Sehgal RK, De Clercq E, Allaudeen HS (1982) *J Virol* **43:** 332-336
41. Dyson MR, Coe PL, Walker RT (1991) *J Med Chem* **34:** 2782-2786
42. Herderwijn PAM (1994) *Antiviral Chem Chemother* **5:** 131-146
43. Coen DM (1991) *Antiviral Res* **15:** 287-300
44. Kaufman HE, Varnell ED, Cheng YC, Bobek M, Thompson HW, Dutschman GE (1991) *Antiviral Res* **16:** 227-232
45. Martin JA, Duncan JB, Hall MJ, Wong-Kai-lin P, Lambert RW, Thomas GJ (1989) *Nucleosides & Nucleotides* **8:** 753-764
46. Focher F, Hildebrand C, Freese S, Ciarrocchi G, Noonan T, Sangalli S, Brown N, Spadari S, Wright G (1988) *J Med Chem* **31:** 1496-1500
47. Focher F, Hildebrand C, Sangalli S, Sandoli D, Ciarrocchi G, Rebuzzini A, Pedrali-Noy G, Manservigi R, Wright G, Brown N, Spadari S (1989) *Methods Find Exp Clin Pharmacol* **11:** 577-582
48. Hildebrand C, Sandoli D, Focher F, Gambino J, Ciarrocchi G, Spadari S, Wright G (1990) *J Med Chem* **33:** 203-206
49. Leib DA, Ruffner KL, Hildebrand C, Schaffer PA, Wright GE, Coen DM (1990) *Antimicrob Agents Chemother* **34:** 1285-1286
50. Larsson A, Alenius S, Johansson N-G, Öberg B (1983) *Antiviral Res* **3:** 77-86
51. Moolten FL (1986) *Cancer Res* **46:** 5276-5281
52. Culver KW, Ram Z, Walbridge S, Ishii H, Oldfield EH, Blaese RM (1992) *Science* **256:** 1550-1552
53. Ram Z, Culver KW, Walbridge S, Blaese RM, Oldfield EH (1993) *Cancer Res* **53:** 83-88
54. Stone R (1992) *Science* **256:** 1513
55. Barba D, Hardin J, Sadelain M, Gage FH (1994) *Proc Natl Acad Sci USA* **91:** 4348-4352
56. Balzarini J, De Clercq E, Ayusawa D, Seno T (1985) *FEBS Lett* **185:** 95-100
57. Balzarini J, De Clercq E, Verbruggen A, Ayusawa D, Seno T (1985) *Mol Pharmacol* **28:** 581-587
58. Balzarini J, De Clercq E, Verbruggen A, Ayusawa D, Shimizu K, Seno T (1987) *Mol Pharmacol* **32:** 410-416
59. Balzarini J, De Clercq E (1989) *Methods Find Exp Clin Pharmacol* **11:** 379-389

60. Balzarini J, Bohman C, De Clercq E (1993) *J Biol Chem* **268:** 6332-6337
61. Bohman C, Balzarini J, Wigerinck P, Van Aerschot A, Herdewijn P, De Clercq E (1994) *J Biol Chem* **269:** 8036-8043
62. Balzarini J, Bohman C, Walker RT, De Clercq E (1994) *Mol Pharmacol* **45:** 1253-1258
63. Kimberlin DW, Coen DM, Biron KK, Cohen JI, Lamb RA, McKinlay M, Emini EA, Whitley RJ (1995) *Antiviral Res* **26:** 369-401
64. Ertl P, Snowden W, Lowe D, Miller W, Collins P, Littler E (1995) *Antiviral Chem Chemother* **6:** 89-97
65. Folkers G, Trumpp-Kallmeyer S, Gutbrod O, Krickl S, Fetzer J, Keil GM (1991) *J Comput-Aided Mol Des* **5:** 385-404
66. Michael M, Fetzer J, Folkers G (1994) *Eur J Biochem* **226:** 219-226
67. Munir KM, French DC, Dube DK, Loeb LA (1994) *Protein Eng* **7:** 83-89

The potential of acyclic nucleoside phosphonates as broad-spectrum antiviral agents

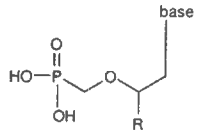
LIEVE NAESENS[1],
JAN BALZARINI[2],
ERIK DE CLERCQ[3]

Contents:

Abbreviations used: *Enzymes*: PRPP — 5-phosphoribosyl-1--pyrophosphate synthetase; RT — reverse transcriptase; TK — thymidine kinase. *Viruses*: CMV — cytomegalovirus; HSV — herpes simplex virus; VZV — varicella-zoster virus. *Other abbreviations*: ANP — acyclic nucleoside phosphonate; AZT — 3′-azido-2′,3′-dideoxythymidine; DAP — 2,6-diaminopurine; FPMP — 3-fluoro-2-phosphonylmethoxypropyl derivatives; HPMP — 3-hydroxy-2-phosphonylmethoxypropyl derivatives; PME — 2-phosphonylmethoxyethyl derivatives; PMP — 2-phosphonylmethoxypropyl derivatives; bis-(POM) — bis(pivaloyloxymethyl).

I. Introduction

The acyclic nucleoside phosphonate (ANP) analogues represent an original class of nucleotide analogues that consist of a purine or pyrimidine base, linked to an acyclic alkyl chain containing a phosphonate group. Several subgroups, each showing a different antiviral activity spectrum, can be distinguished: the 3-hydroxy--2-phosphonylmethoxypropyl (HPMP) derivatives (prototypes: HPMPA and HPMPC); the 2-phosphonylmethoxyethyl (PME) derivatives (prototype: PMEA); the 3-fluoro-2-phosphonylmethoxypropyl (FPMP) derivatives (prototype: FPMPA); and the 2-phosphonylmethoxypropyl (PMP) derivatives (prototype: (*R*)-PMPA) (Fig. 1). HPMPC (cidofovir) has been the subject of far advanced clinical trials for the treatment of cytomegalovirus (CMV) infections, whereas PMEA is being explored for its efficiency in HIV-infected individuals. Here we will review the antiviral potential of the ANP analogues, i.e. their



B	BASE	COMPOUND NAME	ANTIVIRAL ACTIVITY SPECTRUM
H	adenine	PMEA	retroviruses, herpesviruses, hepadnaviruses
CH_2F	adenine	FPMPA	retroviruses, hepadnaviruses
CH_3	adenine	PMPA	retroviruses, hepadnaviruses
CH_2OH	adenine	HPMPA	herpesviruses, adenoviruses, poxviruses
CH_2OH	cytosine	HPMPC	herpesviruses, adenoviruses, poxviruses

Fig. 1. Structural formulae and antiviral activity spectrum of ANP analogues.
PMEA — 9-(2-phosphonylmethoxyethyl)adenine; FPMPA — (*S*)-9-(3-fluoro-2-phosphonylmethoxypropyl)adenine; PMPA — (*R*)-9-(2-phosphonylmethoxypropyl)adenine; HPMPA — (*S*)-9-(3-hydroxy-2-phosphonylmethoxypropyl)-adenine; HPMPC — (*S*)-1-(3-hydroxy-2-phosphonylmethoxypropyl)cytosine.

[1,2,3] Rega Institute for Medical Research, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

antiviral efficacy *in vitro* and *in vivo*, and their pharmacological features, with a special focus on HPMPC and PMEA.

II. Mode of action and metabolism

Due to the strong anionic character, the cellular uptake of ANP analogues is rather slow and inefficient, involving a temperature-dependent, saturable and active transport system [1, 2]. Two independent studies have demonstrated that PMEA, HPMPA and HPMPC are taken up *via* an endocytosis-like process that is not influenced by known inhibitors of nucleoside transport [1, 2]. This is in accordance with the recent data of C i h l á ř *et al.* [3], who have isolated a 50-kDa plasma membrane protein, which may mediate the cellular transport of PMEA. After cellular uptake, the ANP analogues HPMPC, HPMPA, and PMEA undergo enzymatic phosphorylation, generating the diphoshorylated metabolites (designated HPMPCpp, HPMPApp and PMEApp, respectively), which are considered to be the active forms at the level of the viral DNA polymerase. The phosphorylation pathway of HPMPC is clearly distinct from that of the purine derivatives HPMPA, FPMPA and PMEA (Table 1). The latter compounds have been shown to be substrates of 5-phosphoribosyl-1-pyrophosphate (PRPP) synthetase, which can convert them directly into their diphosphorylated metabolites [4]. However, since the phosphorylating capacity of this enzyme for HPMPA, FPMPA and PMEA is at least 1000-fold lower than for the natural substrate AMP, the physiological role of PRPP synthetase in the activation of HPMPA, FPMPA and PMEA remains unclear. The phosphorylation may also proceed in two steps, each catalyzed by AMP-kinase, as suggested by the experiments of M e r t a *et al.* [5]. The purified AMP-kinase used in these studies was shown to specifically recognize the (*S*)-enantiomers of HPMPA and FPMPA, thus providing a biochemical basis for the antiviral in activity of their (*R*)-enantiomers [5]. In the case of HPMPC, the metabolites formed in cell culture are HPMPCp (the monophosphorylated form), the active form HPMPCpp and the choline adduct, HPMPCp-choline [6-8]. The long intracellular half-lives of HPMPCpp and HPMPCp-choline (17 hours and 48 hours, respectively) are supposed to form the basis for the marked antiviral efficacy of HPMPC upon infrequent dosing, as observed *in vitro* and *in vivo* [12]. Similar observations have been made for PMEA: the intracellular half-life of PMEApp being about 18 hours [9]. For HPMPC, however, the choline metabolite acts as an extra intracellular reservoir, making the intracellular accumulation of HPMPC and its metabolites much more pronounced than in the case of PMEA. Accordingly, HPMPC shows higher antiviral

Table 1.

Proposed mechanisms for intracellular uptake, activation, and antiviral activity of the ANP analogues

Compound	Cellular uptake	Enzymes that may be involved in			
		intracellular metabolism[a]	antiviral activity		cytotoxic activity
			anti-herpesvirus	anti-retrovirus	
HPMPC	endocytosis [2]	PNMP kinase [8]; cytidylyltransferase [8]	viral DNA polymerase (HSV-1, HSV-2, CMV) [7, 10]	not active	cellular DNA polymerases [7] cellular ribonucleotide reductase?
HPMPA	endocytosis [1]	AMP kinase [5]; PRPP synthetase [4]	viral DNA polymerase (HSV-1, HSV-2, CMV) [10, 14]; HSV-1 ribonucleotide reductase [15]	not active	cellular DNA polymerases [10]; purine nucleoside phosphorylase [16]; cellular ribonucleotide reductase?
PMEA	endocytosis [1]	AMP kinase [5]; PRPP synthetase [4]	viral DNA polymerase (HSV-1, HSV-2) [10]; HSV-1 ribonucleotide reductase [15]	reverse transcriptase [9]	cellular DNA polymerases [9, 10]; purine nucleoside phosphorylase [16]; cellular ribonucleotide reductase?
FPMPA	endocytosis?	AMP kinase [5]; PRPP synthetase [4]	not active	reverse transcriptase [11]	cellular DNA polymerases [11]; purine nucleoside phosphorylase?; cellular ribonucleotide reductase?

[a] Abbreviations: PNMP kinase: pyrimidine nucleoside-5'-monophosphate kinase or cytidylate kinase; AMP kinase: adenylate kinase; PRPP synthetase: 5-phosphoribosyl-5-pyrophosphate kinase.

activity than PMEA in a short-term or infrequent treatment schedule [13].

There is general consensus on the role of the viral DNA polymerase as the prime target for the antiviral action of the ANP analogues. The K_i values for inhibition of HSV-1 DNA polymerase are 1.42, 0.86, and 0.11 μM, for HPMPApp, HPMPCpp, and PMEApp, respectively [7, 10]. In the case of HIV-1 reverse transcriptase, the K_i values for PMEApp and (*RS*)-FPMPApp are 0.09 μM and 0.034 μM, respectively [9, 11]. Inhibition of cellular DNA polymerases (mainly α) occurs at much higher concentrations. For instance, the K_i of HPMPCpp for HSV-1 DNA polymerase is > 50-fold lower than the K_i for cellular DNA polymerase α [7]. Similarly, in CMV-infected cell cultures, HPMPC inhibits CMV DNA synthesis at concentrations that are 1000-fold lower than those required to inhibit cellular DNA synthesis, thus providing a biochemical basis for the high therapeutic index of HPMPC observed *in vitro* and *in vivo* [12]. The inhibition of DNA polymerase by HPMPCpp (and the diphosphorylated forms of other ANP analogues) is competitive regarding the natural substrate (i.e. dCTP *versus* HPMPCpp, or dATP *versus* HPMPApp, PMEApp and FPMPApp) [7, 9-11]. In this way, the ANP analogues may directly inhibit the incorporation of the natural deoxynucleotides. Alternatively, HPMPCpp, HPMPApp, PMEApp, or FPMPApp may serve as alternate substrates, thus being incorporated in the growing DNA chain. It has been observed that, after their incorporation, HPMPA (and possibly HPMPC) initially allow further DNA chain elongation at the 3-hydroxyl group in the alkyl side chain, but eventually lead to DNA chain termination [14]. In contrast, since PMEA and FPMPA lack a hydroxyl group, these compounds act directly as DNA chain terminators at the level of the DNA polymerase [9]. Thus it appears that the antiretroviral effect of PMEA and FPMPA is based on instant inactivation of the reverse transcriptase following direct DNA chain termination. *Vice versa*, the poor if any activity of HPMPA and HPMPC against retroviruses may be due to their inability to act as direct DNA chain terminators.

In addition to viral and cellular DNA polymerases, a number of other enzymes may be involved in the antiviral and/or cytotoxic effects of the ANP analogues. PMEApp and HPMPApp have been found to inhibit HSV-1-encoded ribonucleotide reductase [15]. However, since CMV and HIV do not encode for a virus-specific ribonucleotide reductase, the antiviral efficacy of PMEA against these two viruses should be ascribed to inhibition of other enzymes. Whether cellular ribonucleotide reductase may be the target for the cytotoxic activity of the ANP analogues at higher doses has yet to be defined. Also, the role of cellular purine nucleoside phosphorylase remains to be settled [16]. This enzyme was found to be inhibited by different purine containing ANP analogues, and their corresponding mono- and diphosphorylated metabolites [16].

III. Anti-retrovirus activity

Irrespective of the retrovirus test system used, PMEA has proven to be a remarkably active antiretroviral agent *in vitro*, the antivirally effective concentration ranging from 0.006 to 2 μg/ml [17]. Thus, PMEA must be metabolised in different cell types, i.e. fibroblast cells, as well as lymphocyte cells and monocyte-macrophages, and the active metabolite PMEApp is an inhibitor of reverse transcriptases of both human and animal retroviruses. In this respect, PMEA is clearly distinct from HIV-1-specific non-nucleoside reverse transcriptase inhibitors. Broad-spectrum antiretroviral activity is also displayed by FPMPA, (*R*)-PMPA and its 2,6-diaminopurine analogue (*R*)-PMPDAP, the latter being the most potent antiretroviral ANP compound described to date [18, 19].

The antiretroviral potential of PMEA has been evaluated in a large variety of animal models for retrovirus infections, using either mice or cats infected with oncogenic retroviruses (i.e. Moloney murine sarcoma virus, Rauscher or Friend murine leukemia virus, LP-BM5 murine aids virus complex, and feline leukemia virus), or lentivirus models (i.e. cats infected with feline immunodeficiency virus, lambs infected with visna-maedi virus, or rhesus monkeys infected with simian immunodeficiency virus) (for an overview: see [20]). Taken together, the animal data point to PMEA as a powerful antiretroviral agent, which is superior to AZT both in antiviral potency and therapeutic index [21]. PMEA may also be effective against neurologic manifestations of retrovirus infections, as indicated by its inhibitory effect on intracerebral retrovirus infections in mice [22], and its efficacy against visna-maedi virus infection in lambs [23].

IV. Anti-DNA virus activity

HPMPA and HPMPC are markedly active *in vitro* against a broad range of DNA viruses, including human herpesviruses (i.e. human herpesvirus type 1 (HSV-1) and type 2 (HSV-2), varicella-zoster virus (VZV), cytomegalovirus (CMV), Epstein-Barr virus, and human herpesvirus type 6), adenoviruses, iridiviruses (i.e. African swine fever virus), and poxviruses (i.e. vaccinia virus) (for an overview: see [24-26]). In contrast, FPMPA and (*R*)-PMPA are devoid of any marked anti-herpesvirus activity [18]. An intermediate spectrum is being displayed by PMEA and PMEDAP, which are active against human herpesviruses, but not against adenoviruses and poxviruses [25]. The marked activity of PMEA, FPMPA and (*R*)-PMPA against hepadnaviruses (i.e. human and duck hepatitis B virus) may be based on inhibition of the reverse transcription

reaction during HBV replication (as observed for other RT inhibitors such as 2',3'-dideoxycytidine) [27, 28]. The antiherpetic activity of HPMPA, HPMPC and PMEA has been demonstrated in a large number of animal models, including: mice infected by HSV-1 or HSV-2 by the intraperitoneal, intracerebral, intracutaneous, or vaginal route; mice or rats infected by murine cytomegalovirus or rat cytomegalovirus, respectively; and horses infected with equine herpesvirus by the intranasal route (for an overview: see [26]). In these experiments, drug treatment was performed either by the intraperitoneal, subcutaneous, or topical route. Most interestingly, HPMPA, HPMPC, and PMEA were found to be markedly active against thymidine kinase-deficient (TK^-) HSV-1 strains, in a rabbit model for ocular HSV-1 infection [29]. In this respect, the ANP analogues are clearly distinct from classical antiherpetic drugs such as acyclovir and brivudin, which depend on the HSV-1 thymidine kinase for their phosphorylation and antiviral activity. Overall, HPMPC emerges as a potent and selective anti-herpesvirus agent, its therapeutic index being higher than that of HPMPA, and its potency against herpesviruses exceeding that of PMEA.

V. Pharmacology

The pharmacokinetics of PMEA have been extensively studied in mice, cats and monkeys. Although one should be careful in extrapolating these data, it seems that related ANP analogues have a similar pharmacokinetic profile. The elimination of PMEA is rather rapid, with a terminal elimination half-life of 10 min, 67 min, and 2 h, in mice, cats and monkeys, respectively [30, 31]. PMEA is mainly eliminated (in unchanged form) by the kidneys, and the contribution of metabolism to the elimination of PMEA is minor [31]. Tissue distribution studies in mice point to low yet significant penetration in the brain. This observation has important implications for the use of PMEA against neurological manifestations of viral infections.

Due to their high anionic charge, the ANP analogues have a rather low oral bioavailability, i.e. 22% and 4% for HPMPC and PMEA, respectively (data obtained in cynomologus monkeys) [32, 33]. The oral absorption of PMEA can be increased by using lipohilic ester derivatives that release PMEA in the plasma. Bis(pivaloyloxymethyl)-PMEA (bis(POM)-PMEA) has recently been shown to display an oral bioavailability of 50% in mice, and a marked antiviral efficacy upon oral administration [34].

VI. Clinical experience

The safety and efficacy of HPMPC have been evaluted in a Phase I/II trial in 31 HIV-seropositive men with asymptomatic CMV infection [35]. HPMPC was administered by intravenous infusion, at a dose ranging from 0.5 to 10 mg/kg per week. At doses $\geqslant$ 3 mg/kg, a prolonged and dose-dependent anti-CMV effects was observed, as evidenced by a reduction of CMV titers in semen and urine. However, several patients developed a significant and dose-dependent nephrotoxicity, characterized by proteinuria and increased serum creatinine levels. Concomitant administration with probenecid markedly reduced the severity of the nephrotoxicity. It has been postulated that probenecid interferes with the active tubular secretion of HPMPC, thus preventing the massive accumulation of HPMPC in the tubular cells. Alternatively, accumulation of HPMPC in the kidneys could possibly be prevented by using cyclic HPMPC (the internal ester HPMPC), which has recently been shown to be considerably less nephrotoxic than HPMPC in cynomolgus monkeys [36, 37].

Interestingly, one patient who received intravenous HPMPC in the above study recovered from an acyclovir-resistant rectal herpes simplex infection [35]. Similarly, the potential usefulness of HPMPC in the treatment of HSV infections that are insensitive to acyclovir or foscarnet has been shown in a recent case report on the use of topical HPMPC against mucocutaneous herpesvirus infections [38]. Also, intravitreal injection of HPMPC has proven to be beneficial in the treatment of CMV retinitis [39, 40]. The therapeutic potential of HPMPC eye drops in the treatment of eye infections caused by herpes simplex virus or by adenoviruses has already been demonstrated in animal models [29, 41], but remains to be confirmed in humans.

The tolerance and anti-HIV efficacy of intravenous PMEA have been investigated in a limited dose-escalating Phase I/II study in HIV-infected individuals [42, 43]. At a dose 1 mg/kg per day, a significant increase in CD4 cell count and a drop in HIV p24 was noted. However, several patients manifested neutropenia and moderate hepatic toxicity, as evidenced by increased hepatic transaminases. In contrast, no signs of neutropenia have been observed in HIV-infected patients who received oral bis(POM)-PMEA, at daily doses up to 500 mg [44]. The favorable oral bioavailability of bis(POM)-PMEA (40% in humans), and apparent lack of severe toxicity, point to the potential usefulness of bis(POM)-PMEA as an oral prodrug of PMEA.

VII. Conclusion

Recent clinical studies on HPMPC have demonstrated its efficacy against herpesviruses (herpes simplex virus and cytomegalovirus), including strains that are insensitive to the classical antiherpetic agents. The efficacy of HPMPC against other DNA viruses such as adenoviruses has yet to be determined. The dual antiviral activity of PMEA against both herpesviruses and retroviruses should be advantageous in the

treatment of HIV-infected individuals that suffer from opportunistic herpesvirus infections. This should become clear from ongoing clinical studies with bis(POM)-PMEA, the oral prodrug of PMEA. In addition, the highly specific antiretroviral compounds FPMPA and (*R*)-PMPA merit further investigation in HIV-infected individuals.

References

1. Palu G, Stefanelli S, Rassu M, Parolin C, Balzarini J, De Clercq E (1991) *Antiviral Res* **16**: 115-119
2. Connelly MC, Robbins BL, Fridland A (1993) *Biochem Pharmacol* **46**: 1053-1057
3. Cihlář T, Rosenberg I, Votruba I, Holý A (1995) *Antimicrob Agents Chemother* **39**: 117-124
4. Balzarini J, De Clercq E (1991) *J Biol Chem* **266**: 8686-8689
5. Merta A, Votruba I, Jindřich J, Holý A, Cihlář T, Rosenberg I, Otmar M, Herve TY (1992) *Biochem Pharmacol* **44**: 2067-2077
6. Bronson JJ, Ho HT, De Boeck H, Woods K, Ghazzouli I, Martin JC, Hitchcock MJM (1989) *Ann NY Acad Sci* **616**: 398-407
7. Ho HT, Woods KL, Bronson JJ, De Boeck H, Martin JC, Hitchcock MJM (1992) *Mol Pharmacol* **41**: 197-202
8. Cihlář T, Votruba I, Horska K, Liboska R, Rosenberg I, Holý A (1992) *Collect Czech Commun* **57**: 661-672
9. Balzarini J, Hao Z, Herdewijn P, Johns DG, De Clercq E (1991) *Proc Natl Acad Sci USA* **88**: 1499-1503
10. Holý A, Votruba I, Merta A, Černý J, Veselý J, Vlach J, Šedivá K, Rosenberg I, Otmar M, Hřebabecký H, Trávniček M, Vonka V, Snoeck R, De Clercq E (1990) *Antiviral Res* **13**: 295-312
11. Balzarini J, Holý A, Jindřich J, Dvorakova H, Hao Z, Snoeck R, Herdewijn P, Johns DG, De Clercq E (1991) *Proc Natl Acad Sci USA* **88**: 4961-4965
12. Neyts J, Snoeck R, Schols D, Balzarini J, De Clercq E (1990) *Virology* **179**: 41-50
13. Aduma P, Connelly MC, Srinivas RV, Fridland A (1995) *Mol Pharmacol* **47**: 816-822
14. Votruba I, Bernaerts R, Sakuma T, De Clercq E, Merta A, Rosenberg I, Holý A (1987) *Mol Pharmacol* **32**: 524-529
15. Černý J, Votruba I, Vonka V, Rosenberg I, Otmar M, Holý A (1990) *Antiviral Res* **13**: 253-264
16. Šedivá K, Ananiev AV, Votruba I, Holý A, Rosenberg I (1991) *Int J Purine Pyrimidine Res* **2**: 35-39
17. Balzarini J, Naesens L, Slachmuylders J, Niphuis H, Rosenberg I, Holý A, Schellekens H, De Clercq E (1991) *AIDS* **5**: 21-28
18. Balzarini J, Holý A, Jindřich J, Naesens L, Snoeck R, Schols D, De Clercq E (1993) *Antimicrob Agents Chemother* **37**: 332-338
19. Vahlenkamp TW, Deronde A, Balzarini J, Naesens L, De Clercq E, Vaneijk MJT, Horzinek MC, Egbering HF (1995) *Antimicrob Agents Chemother* **39**: 746-749
20. Naesens L, Balzarini J, De Clercq E (1994) *Rev Med Virol* **4**: 147-159
21. Balzarini J, Naesens L, Herdewijn P, Rosenberg I, Holý A, Pauwels R, Baba M, Johns DG, De Clercq E (1989) *Proc Natl Acad Sci USA* **86**: 332-336
22. Balzarini J, Sobis H, Naesens L, Vandeputte M, De Clercq E (1990) *Int J Cancer* **45**: 486-489
23. Thormar H, Georgsson G, Palsson PA, Balzarini J, Naesens L, Torsteinsdottir S, De Clercq E (1995) *Proc Natl Acad Sci USA* **92**: 3283-3287
24. De Clercq E, Holý A, Rosenberg I, Sakuma T, Balzarini J, Maudgal PC (1986) *Nature* (London) **323**: 464-467
25. De Clercq E, Sakuma T, Baba M, Pauwels R, Balzarini J, Rosenberg I, Holý A (1987) *Antiviral Res* **8**: 261-272
26. De Clercq E (1993) *Rev Med Virol* **3**: 85-96
27. Heijtink RA, Kruining J, De Wilde GA, Balzarini J, De Clercq E, Schalm SW (1994) *Antimicrob Agents Chemother* **38**: 2180-2182
28. Yokota T, Mochizuki S, Konno K, Mori S, Shigeta S, De Clercq E (1991) *Antimicrob Agents Chemother* **35**: 394-397
29. Maudgal PC, De Clercq E (1991) *Antiviral Res* **16**: 93-100
30. Naesens L, Balzarini J, De Clercq E (1992) *Clin Chem* **38**: 480-485
31. Naesens L, Balzarini J, De Clercq E (1992) *Drug Metab Disp* **20**: 747-752
32. Bischofberger N, Prisbe E, Arimilli M, Cundy K, Lee B (1993) *Antiviral Res* **20 S(1)**: 142
33. Cundy KC, Shaw JP, Lee WA (1994) *Antimicrob Agents Chemother* **38**: 365-368
34. Naesens L, Balzarini J, Bischofberger N, De Clercq E (1996) *Antimicrob Agents Chemother* **40**: 22-28
35. Lalezari JP, Drew WL, Glutzer E, Miner D, Flaherty J, Fisher PE, Cundy K, Hannigan J, Martin JC, Jaffe HS (1995) *J Inf Dis* **171**: 788-796
36. Hitchcock MJM, Lacy SA (1995) *Annu Society of Toxicology Meeting* March 6, 1995, Baltimore, MD Abstract
37. Bischofberger N, Hitchcock MJM, Chen MS, Barkhimer DB, Cundy K, Kent KM, Lacy SA, Lee WA, Li ZH, Mendel DB, Smee DF, Smith JL (1994) *Antimicrob Agents Chemother* **38**: 2387-2391
38. Snoeck R, Andrei G, Gerard M, Silverman A, Hedderman A, Balzarini J, Sadzot-Delvaux C, Tricot G, Clumeck N, De Clercq E (1994) *Clin Inf Dis* **18**: 570-578
39. Kirsch LS, Arevalo JF, De Clercq E, Chavez E, Manguia D, Garcia R, Freeman WR (1995) *Am J Ophthalmol* **119**: 466-476
40. Kirsch LS, Arevalo JF, Delapaz EC, Munguia D, De Clercq E, Freeman WR (1995) *Ophthalmology* **102**: 533-542
41. Gordon YJ, Romanowski EG, Araullo-Cruz T (1994) *Invest Ophthalmol Vis Sci* **35**: 4135-4143
42. Collier AC, Coombs RW, Nienow J, Paradise M, Yang HH, Troxel S, Boggs J, Ebeling D, Jaffe HS, Corey L (1994) *First Natl Conf Human Retroviruses* Abstract Book No 563
43. Walker RE, Vogel SN, Jaffe HS, Polis MA, Kovacs JA, Falloon J, Davey RT, Ebeling D, Cundy K, Paar D, Markowitz M, Masur H, Lane HC (1994) *First Natl Conf Human Retroviruses* Abstract Book No 522
44. Barditch-Crovo PA, Toole J, Burgee H, Wachsman M, Ebeling D, Cundy KC, Jaffe HS, Lietman PS (1995) *Antiviral Res* **26**: pp A229, abstract No 9

Symposium on structure and biological functions of nucleic acid components and their analogues, and related topics —

POLISH ACADEMY OF SCIENCES
INSTITUTE OF BIOCHEMISTRY AND BIOPHYSICS

&
UNIVERSITY OF WARSAW
DEPARTMENT OF BIOPHYSICS

SYMPOSIUM ON STRUCTURE AND BIOLOGICAL FUNCTIONS OF NUCLEIC ACID COMPONENTS AND THEIR ANALOGUES, AND RELATED TOPICS

In honour of
Professor David Shugar
on the occasion of his
80th Birthday Anniversary

Warsaw, 2 - 5 September 1995

SYMPOSIUM SPONSORS

A. Pawińskiego 5a, PL 02-106 Warszawa
phone 658 47 24, fax 658 46 36, (48)39 121623

Honorary Protectorate

The Scientific Council of the Institute
Zofia Lassota (Chairman)

Organizing Committee

Kazimierz L. Wierzchowski (Chairman)
Irena Pietrzykowska (V-chairman)
Daniela Barszcz
Tadeusz Chojnacki
Jarosław Kuśmierek
Grażyna Muszyńska
Ryszard Stolarski

Symposium Secretariat

Agata Kwinto
Institute of Biochemistry
and Biophysics, PAS
A. Pawińskiego St. 5a,
PL 02-106 Warszawa,
building B, first floor, room no. 41
phone: 658 47 53,
fax (komertel): (48) 39 121623

■ SATURDAY, SEPTEMBER 2nd

9.00 - 9.30 Opening ceremony

9.30 - 10.00 *David SHUGAR*
Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa
Untitled

10.10 - 10.40 *Helmut VORBRÜGGEN*
Pharmaceutical Research, Schering AG, Berlin
Trends and advances in nucleoside synthesis

10.45 - 11.15 *Wolfgang PFLEIDERER*
Fakultät für Chemie, Universität Konstanz, Konstanz
Recent progress in oligonucleotide synthesis

11.20 - 11.40 Coffee break

11.40 - 12.10 *Frank SEELA*
Laboratorium für Organische und Bioorganische Chemie, Universität Osnabrück, Osnabrück
Oligonucleotides with modified bases or an alternate sugar phosphate backbone

12.15 - 12.45 *Wojciech STEC*
Center of Molecular and Macromolecular Studies, Polish Academy of Sciences, Łódź
Stereocontrolled synthesis of oligo(nucleoside methanephosphate)s

12.50 - 13.20 *Bożenna GOLANKIEWICZ*
Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań
Synthetic and biological applications of tricyclic analogues of guanosine

13.30 - 15.00 Lunch

15.00 - 15.30 *Karin BIRNBAUM*
Division of Informatics, National Research Council, Ottawa
Crystal structure of some acyclonucleosides with antiviral activity, and related compounds

15.35 - 16.05 *Wolfram SAENGER*
Institut für Kristallographie, Freie Universität Berlin, Berlin
Recognition of Tet repressor by tetracycline and regulation of antibiotic resistance

16.10 - 16.40 *Wilhelm GUSCHLBAUER*
Service de Biochimie et Genetique Cellulaire, Centre d'Etudes de Saclay, Gif-sur-Yvette
"Small is beautiful": major modifications in DNA structure by small substituents or ligands

16.45 - 17.15 *Maciej WIEWIÓROWSKI*
Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań
Crystal engineering of nucleosides: new tools for studying basic processes of molecular biology

17.30 - 18.30 Welcome mixer

18.30 - 19.15 *Friedrich Cramer*
Max-Planck-Institut für Experimentelle Medizin, Göttingen
Emil Fischer's lock and key concept a hundred years ago and today

■ SUNDAY, SEPTEMBER 3rd

In the morning:
Visit to the Old City of Warsaw and the King's Castle

■ MONDAY, SEPTEMBER 4th

9.00 - 9.30 *Józef S. KWIATKOWSKI*
Institute of Physics, N. Copernicus University, Toruń
Molecular structure and IR spectra of the DNA bases and their derivatives: experiment and theory

9.35 - 10.05 *G. Victor FAZAKERLEY*
Service de Biochimie et Genetique Moleculaire, Centre d'Etudes de Saclay, Gif-sur-Yvette
Structures of mismatched base pairs in DNA

10.10 - 10.40 *Jane K. SETLOW*
Brookhaven National Laboratory, Upton
Influence of DNA configuration on mutagenesis

10.45 - 11.05 Coffee break

11.05 - 11.35 *Robert H. HAYNES*
Department of Biology, York University, North York
Genetic consequences of deoxyribonucleotide pool imbalances

11.40 - 12.10 *Celina JANION*
Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa
MMS-induced mutagenesis in *E. coli*: DNA damage and repair

12.15 - 12.45 *Ronald HANCOCK*
Centre de Recherche sur le Cancer, L'Hotel Dieu de Québec, Québec
Structure of DNA topoisomerase II

13.00 - 14.30 Lunch

14.30 - 15.00 *George E. WRIGHT*
Department of Pharmacology, Medical School, University of Massachusetts, Worcester
Nucleotide probes of DNA polymerases

15.05 - 15.35 *Alexander KRAYEWSKY*
Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow
Topological analysis of DNA polymerase active center using modified dNTP's

15.40 - 16.10 *Wojciech RODE*
Nencki Institute of Experimental Biology, Polish Academy of Sciences, Warszawa
Interaction with thymidylate synthase of 2- and/or 4-substituted analogues of 2′-deoxyuridylate and 5-fluoro-2′-deoxyuridylate

16.15 - 16.35 Coffee break

16.35 - 17.05 *Katarzyna BĘBENEK*
National Institutes of Health, National Institute of Environmental Health Sciences, Research Triangle Park
Structure-function studies of HIV-1 reverse transcriptase

17.10 - 17.40 *Nils Gunnar JOHANSSON*
Research Laboratories, Medivir AB, Huddinge
Structure-activity relationships for phosphorylation of nucleoside analogs to monophosphates by nucleoside kinases

17.45 - 18.15 *Lorenzo A. PINNA*
Dipartimento di Chimica Biologica, Università di Padova, Padova
Molecular enzymology of protein kinase CK2, a "virtuoso" in the nuclear orchestra

18.20 - 18.50 *Ryszard KOLE*
Lineberger Comprehensive Cancer Center, School of Medicine, University of North Carolina, Chapel Hill
Mechanism of splice site selection

■ TUESDAY, SEPTEMBER 5th

9.00 - 9.30 *Erik De CLERCQ*
Rega Institute for Medical Research, University of Leuven, Leuven
Nucleoside analogues as antiviral agents

9.35 - 10.05 *David C.-K. CHU*
College of Pharmacy, The University of Georgia, Georgia
Nucleosides as anti-HBV agents

10.10 - 10.40 *Yechiel BECKER*
Department of Molecular Virology, Faculty of Medicine, The Hebrew University, Jerusalem
Targeting of antiviral peptides and drugs to skin Langerhans cells for transfer to HIV in infected lymph nodes — a novel anti-AIDS strategy

10.45 - 11.05 Coffee break

11.05 - 11.35 *Krzysztof PANKIEWICZ*
Memorial Sloan-Kettering Cancer Center, New York
NAD analogues as anticancer agents

11.40 - 12.10 *Gilles GOSSELIN*
Laboratoire de Chimie Bioorganique, Université Montpellier II, Montpellier
New insights regarding the potential of the pronucleotide approach in chemotherapy

12.15 - 12.45 *Robert M. TANGUAY*
Laboratory of Cellular and Developmental Genetics, RSVS, Laval University, Ste-Foy
Tyrosine and its catabolites: from disease to cancer

12.50 - 13.20 *Ephraim KATCHALSKI-KATZIR*
Department of Membrane Research and Biophysics, Weizmann Institute of Science, Rehovot
Synthesis, physicochemical and biological properties of poly-alpha-amino acid — the simplest of protein models

13.30 - 15.00 Lunch

15.00 - 15.30 *Ven MURTHY*
Department of Biochemistry, Faculty of Medicine, Laval University, Ste-Foy
Genotypes and phenotypes of lipoprotein lipase deficiency in the French Canadian population of Québec

15.35 - 16.05 *Sergio CAFFIERI*
Department of Pharmaceutical Sciences, University of Padova, Padova
Psoralen photosensitization: role of photodamages to nucleic acid and membrane lipid components

16.45 Closing

For published lectures of the Symposium see: *Acta Biochimica Polonica* (1996) vol. 43 No. 1.

HONOROWE PROMOCJE

Z radością informujemy o uroczystościach nadania tytułu doktora Honoris Causa dwum zasłużonym członkom naszego Towarzystwa.

Profesor Jerzy Pawełkiewicz — twórca i kierownik pierwszej katedry biochemii w akademickiej uczelni rolniczej w kraju, profesor Akademii Rolniczej w Poznaniu, członek rzeczywisty PAN i członek grona założycieli Polskiego Towarzystwa Biochemicznego — otrzymał doktorat honorowy swej macierzystej uczelni, a uroczysta promocja odbyła się dnia 20 października 1995 roku.

Profesor Mariusz Żydowo, biochemik, uczeń i następca profesora Włodzimierza Mozołowskiego — od 1945 roku już obecny w Akademii Lekarskiej w Gdańsku: jako student, absolwent z indeksem o numerze 1, lekarz i doktor medycyny dyplomowany summa cum laudae, docent habilitowany i profesor w Zakładzie Biochemii, inicjator budowy i projektodawca gmachu Biochemii, Zakładów Teoretycznych Akademii, pierwszy w kraju rektor uczelni akademickich z wyboru, zasłużony członek rzeczywisty i honorowy Polskiego Towarzystwa Biochemicznego — otrzymał doktorat honorowy swej macierzystej uczelni dnia 7 grudnia 1995 roku.

W imieniu Zarządu Głównego Towarzystwa i Redakcji składam obydwu Honorowym Doktorom serdeczne Gratulacje i Życzenia.

Zofia Zielińska

Uprzejmie zawiadamiamy PT Koleżanki i PT Kolegów
o przeniesieniu siedziby Zarządu Głównego
Polskiego Towarzystwa Biochemicznego
na teren Instytutu Biologii Doświadczalnej im. M. Nenckiego
pokój 632 i 633

Obecny adres:

Polskie Towarzystwo Biochemiczne
ul. Pasteura 3, 02-093 Warszawa
Tel. bezpośredni 658 20 99
Tel. przez centralę 659 85 71 w. 352
Fax 22 53 42

Dyżury biura Zarządu odbywać się będą jak dotychczas
we wtorki w godz 12-18

Kwartalnik „Postępy Biochemii" wydawany z pomocą finansową Komitetu Badań Naukowych

Pokwitowanie dla wpłacającego

zł .....

słownie .....

wpłacający .....

imię, nazwisko, dokładny adres z kodem pocztowym

na rachunek

Polskie Towarzystwo Biochemiczne
02-093 Warszawa, ul. Pasteura 3
P.B.K. XIII/O W-wa, Al. Jerozolimskie
37 00 44-1225-139-11

stempel

Pobrano opłatę

zł. .....

podpis przyjmującego

Odcinek dla posiadacza rachunku

zł .....

słownie .....

wpłacający .....

imię, nazwisko, dokładny adres z kodem pocztowym

na rachunek

Polskie Towarzystwo Biochemiczne
02-093 Warszawa, ul. Pasteura 3
P.B.K. XIII/O W-wa, Al. Jerozolimskie
37 00 44-1225-139-11

stempel

Pobrano opłatę

zł. .....

podpis przyjmującego

Odcinek dla poczty lub banku

zł .....

słownie .....

wpłacający .....

imię, nazwisko, dokładny adres z kodem pocztowym

na rachunek

Polskie Towarzystwo Biochemiczne
02-093 Warszawa, ul. Pasteura 3
P.B.K. XIII/O W-wa, Al. Jerozolimskie
37 00 44-1225-139-11

stempel

Pobrano opłatę

zł. .....

podpis przyjmującego

For additional copies contact the
Executive Committee of the Polish
Biochemical Society.
ul. L. Pasteura 3, 02-093 Warszawa
Poland