Cell, Vol. 75, 1129–1136, December 17, 1993, Copyright © 1993 by Cell Press

β-D-Glucosyl-Hydroxymethyluracil: A Novel Modified Base Present in the DNA of the Parasitic Protozoan T. brucei

Janet H. Gommers-Ampt,* Fred Van Leeuwen,* Antonius L. J. de Beer,† Johannes F. G. Vliegenthart,† Miral Dizdaroglu,‡ Jeffrey A. Kowalak,§ Pamela F. Crain,‖ and Piet Borst*

*Division of Molecular Biology
The Netherlands Cancer Institute
Plesmanlaan 121
1066 CX Amsterdam
The Netherlands
†Bijvoet Center
Department of Bio-Organic Chemistry
Utrecht University
3584 CH Utrecht
The Netherlands
‡Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, Maryland 20899
§Department of Biochemistry
‖Department of Medicinal Chemistry
University of Utah
Salt Lake City, Utah 84112

Summary

We have previously shown that the DNA of the unicellular eukaryote T. brucei contains about 0.1% of a novel modified base, called J. The presence of J correlates with a DNA modification associated with the silencing of telomeric expression sites for the variant surface antigens of trypanosomes. Here we show that J is 5-((β-D-glucopyranosyloxy)-methyl)-uracil (shortened to β-D-glucosyl-hydroxymethyluracil), a base not previously found in DNA. We discuss putative pathways for the introduction of this base modification at specific positions in the DNA and the possible contribution of this modification to repression of surface antigen gene expression.

Introduction

An unusual kind of DNA modification is present in African trypanosomes. The existence of this modification was deduced from the incomplete cleavage of some PstI and PvuII restriction enzyme recognition sites (Bernards et al., 1984; Pays et al., 1984). The modification is unusual, because the incompletely cut sites are only found in and around inactive subtelomeric variant-specific surface glycoprotein (VSG) genes. Actively transcribed telomeric VSG genes or chromosome-internal VSG genes are completely digested (Bernards et al., 1984; Pays et al., 1984). An additional peculiarity is that the degree of modification at a given site increases with the length of the telomeric repeat segment (Bernards et al., 1984). The modification is present in bloodstream-form trypanosomes that undergo antigenic variation but not in the insect-form (procyclic) trypanosomes that do not vary their surface coat (Pays et al., 1984; Gommers-Ampt et al., 1991). This suggests that the modification is involved in the repression of subtelomeric expression sites for VSG genes. These are only used in bloodstream trypanosomes, and usually only one is active at a time. How the other expression sites are repressed is still unclear (reviewed by Van der Ploeg, 1987; Cross, 1990; Borst, 1991). As the DNA modification can distinguish inactive expression sites from the active expression site, it could play a key role in the repression (Bernards et al., 1984). The nature of the modification and its putative role in gene expression are thus of special interest for understanding antigenic variation in trypanosomes.

Initial attempts to identify an unusual nucleoside in trypanosome DNA by standard high pressure liquid chromatography analysis were unsuccessful (Crozatier et al., 1988). Using a more sensitive postlabeling technique combined with a two-dimensional (2D) separation, however, an unusual 5′-nucleotide, called pdJ, was detected. pdJ comprised 0.1% of the nucleotides of bloodstream-form trypanosome DNA and was enriched in DNA fractions with a high ratio of (sub)telomeric to total DNA (Gommers-Ampt et al., 1991). pdJ was undetectable (<0.0002%) in insect-form trypanosome DNA. These are properties expected for the postulated subtelomeric modified nucleotide in trypanosome DNA. In 2D thin-layer chromatography (TLC), pdJ did not comigrate with any of the known DNA nucleotides (Gommers-Ampt et al., 1991), suggesting that it is a novel DNA component. Using a combination of biochemical and chemical analytical techniques, we have established that pdJ is β-D-glucosyl-hydroxymethyldUMP. The evidence for this structure is presented in this paper.

Results

Purification of dJp by Liquid Chromatography

For analytical purposes, relatively large amounts of purified nucleotide were required. Since these could not easily be obtained from the TLC sheets on which we had initially detected 5′-nucleotide (pdJ), we developed a liquid chromatography procedure for the purification of 3′-nucleotide (dJp). Enzymatic hydrolysates of DNA from either bloodstream-form trypanosomes (containing J) or insect-form trypanosomes (not containing J) were fractionated on a reversed-phase column, as shown in Figure 1A. ^{32}P-postlabeling analysis of fractions collected from the bloodstream-form eluate showed that dJp comigrated with dTp. The dTp-containing fractions (and the closely migrating dGp-containing fractions) were therefore chromatographed on a Mono Q anion exchange column. Figure 1B shows that this resulted in a minor peak eluting at 13 min that was only obtained with DNA hydrolysates of bloodstream-form trypanosomes. Postlabeling analysis of the peak fractions and of a blank control sample (see Experimental Procedures) confirmed that this peak contains dJp and that no other nucleotides are present in these fractions (data not shown). In this way, 10 nmol of dJp was purified from about 2.5 mg of genomic bloodstream-form DNA.

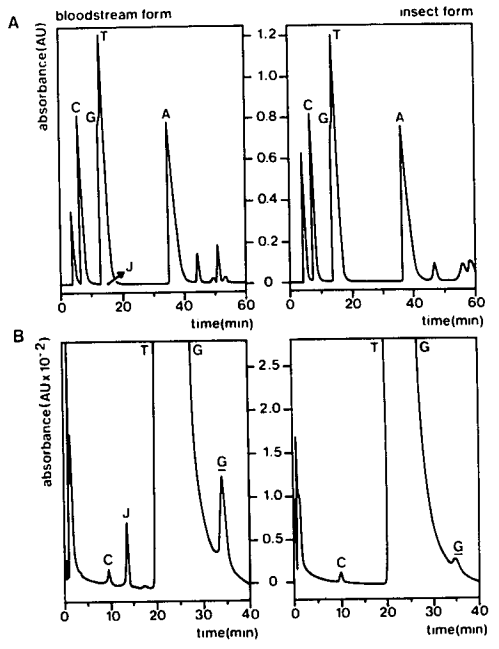

Figure 1. Purification of dJp Using Liquid Chromatography

Elution profiles shown were measured at 260 nm.

(A) The fractionation of dNps, obtained after enzymatic hydrolysis of total genomic DNA of either bloodstream-form (left) or insect-form (right) T. brucei, on a reversed-phase column. Chromatography conditions were as described in Experimental Procedures. The retention time of dJp was determined by postlabeling and 2D TLC of fractions as described in Experimental Procedures. The dJp-containing fractions were further fractionated by anion exchange chromatography.

(B) The elution profiles on an anion exchange column of bloodstream-form hydrolysate fractions derived from the reversed-phase column (left panel) and the corresponding fractions from the reversed-phase eluate of the insect-form hydrolysate (right panel). Elution conditions were as described in Experimental Procedures. The identity of dJp was confirmed by ^{32}P-postlabeling and 2D TLC analysis. N represents the deoxyribonucleotide, N a ribonucleotide.

Purified dJMP Behaves as a HydroxymethyldUMP Derivative

We have previously shown that dJp has the chemical properties of a pyrimidine rather than of a purine nucleotide (Gommers-Ampt et al., 1991). To test whether J is a C or a U analog, pdJ was exposed to nitrous acid (see Experimental Procedures). Conditions that resulted in complete deamination of pdC had no effect on pdJ (data not shown), indicating that J, like U, lacks an amino group sensitive to nitrous acid. In Table 1 the spectral characteristics of column-purified dJp are compared with those of dGp, dAp, dTp, dCp, and hydroxymethyldUp (HOMedUp), all derived from the Trypanosoma brucei DNA hydrolysate and chromatographed on the same columns as dJp. The presence of hydroxymethyluracil (HOMeU) (initially called V) in T. brucei DNA has been reported before (Gommers-Ampt et al., 1991, 1993). The results in Table 1 show that the ultraviolet (UV) absorbance spectrum of dJp is virtually identical at pH 5.6 and pH 3.5. This confirms that J is a U analog since the protonation of an exocyclic amino group results in a difference in UV spectra at pH 3.5 versus pH 5.6 (illustrated by the comparison of dCp with dTp and

Table 1. UV Absorbance Characteristics of dJp and Marker Nucleotides

dNp	G	A	T	C	HOMeU	J
pH 5.6						
λ_{max} (nm)	252	260	267	271	263	264
A_{250}/A_{260}	1.15	0.80	0.66	0.82	0.70	0.71
A_{280}/A_{260}	0.68	0.14	0.64	0.97	0.52	0.52
pH 3.5						
λ_{max} (nm)	252	258	267	278	263	263
A_{250}/A_{260}	1.14	0.85	0.68	0.56	0.73	0.71
A_{280}/A_{260}	0.68	0.18	0.65	1.73	0.52	0.51

Data were obtained by measuring UV absorbance spectra of dNps eluting from the reversed-phase column at pH 5.6 and from the anion exchange column at pH 3.5.

HOMedUp in Table 1). Table 1 also shows that the spectrum of dJp is similar to that of HOMedUp both at pH 5.6 and pH 3.5. This is neither due to identity of J and HOMeU (Gommers-Ampt et al., 1991) nor to contamination of dJp with HOMedUp, as the latter was undetectable in the purified dJp fractions. It rather suggests that J is a HOMeU derivative with a substituent that does not contribute to the UV absorbance.

Mass Spectrometry of Purified dJp

A sensitive way to check whether J is indeed a HOMeU analog is the use of gas chromatography mass spectrometry (GCMS) with selected-ion monitoring of the derivatized base (Dizdaroglu, 1991). For this purpose, dJp was hydrolyzed, and the released base J was trimethylsilylated and analyzed by GCMS with selected-ion monitoring. If J consisted of HOMeU with an acid-labile substituent, the hydrolysis would result in the release of HOMeU. We monitored a number of characteristic ions, including those at m/z 358 and m/z 343, since these are the most intense ions in the mass spectrum of the trimethylsilyl derivative of HOMeU (Dizdaroglu, 1984). Both the retention time and the partial mass spectrum obtained with dJp were identical to those of the trimethylsilyl derivative of authentic HOMeU, consistent with the presence of HOMeU in the dJp sample (data not shown). The blank of the isolation procedure did not contain any HOMeU. Since no HOMedUp was detectable in the purified dJp preparation either, these results suggest that J had been converted into HOMeU during the formic acid hydrolysis and therefore that J could indeed be a HOMeU analog.

To obtain the molecular weight of intact dJp, we turned to electrospray ionization mass spectrometry, which is especially suitable for the analysis of highly polar charged molecules (Smith et al., 1990). The results obtained for dJp are shown in Figure 2. A prominent peak of m/z 499 was observed and assigned to the $(M-H)^-$ ion; the $(M-H)^-$ ion for HOMedUp would occur at m/z 337, but is absent. A molecular weight of 500.2 for dJp is therefore calculated from the centroided ion profile (inset in Figure 2). dJp must contain an even number of nitrogen atoms because its molecular weight is an even number. Subtracting the weight of the HOMedUp partial structure (derived from the

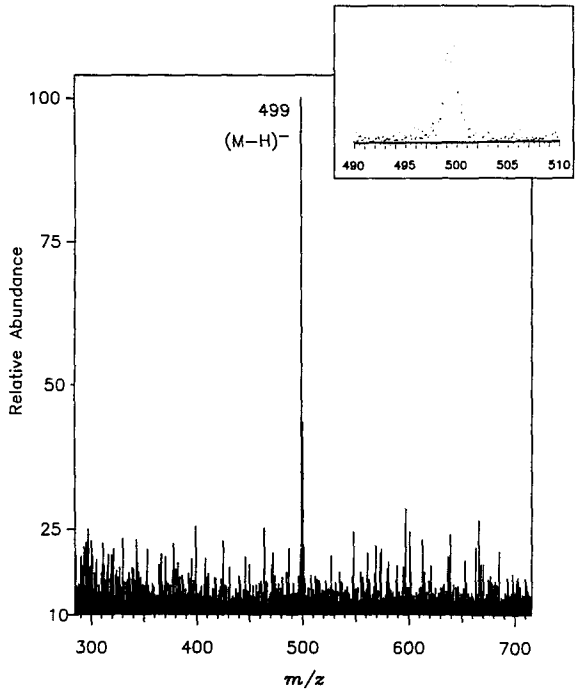

Figure 2. Partial Electrospray Mass Spectrum of Nucleotide dJp
(Inset) Ion profile from m/z 490 to m/z 510 from which a value of 499.2 was derived for the $(M-H)^-$ ion of dJp.

GCMS spectrum) from the molecular weight of dJp gives a value of 163 Da for the labile base substituent, consistent with the value of a hexose fragment. Identical results were obtained with a marker nucleotide glucosyl-HOMedUp, derived by deamination from glucosyl-HOMedCp, as described below.

pdJ Comigrates with β-Glucosyl-HOMedUMP

The experiments presented thus far suggest that J is HOMeU with a substituent that does not affect the UV spectrum and that is attached by a bond sensitive to formic acid hydrolysis (for GCMS). These properties resemble those of glucosylated hydroxymethylcytosine (HOMeC) (Lichtenstein and Cohen, 1960; Wyatt and Cohen, 1953). Moreover, the molecular weight obtained for intact dJp is consistent with that of glucosylated HOMedUMP. We therefore tested whether J could be hexosylated HOMeU by comparing the chromatographic behavior of pdJ with that of authentic α- and β-glucosyl-HOMedU-5′-monophosphate, prepared by nitrous acid deamination of α- and β-glucosylated HOMedC-5′-monophosphate, isolated from T4 DNA. In T4 DNA, the base C is completely replaced by HOMeC with a glucose in either the α (70%) or the β (30%) configuration attached to the hydroxygroup (Lehman and Pratt, 1960). 2D TLC analyses show that α-glucosyl-HOMedUMP, β-glucosyl-HOMedUMP, and pdJ comigrate both on unmodified cellulose (Figure 3) and polyethyleneimine (PEI) cellulose (data not shown), suggesting that J is α- or β-glucosyl-HOMeU or an isomeric hexosyl-HOMeU.

pdJ Is a Substrate for Galactosyl Transferase

To analyze the nature of the neutral hexose attached to the HOMe group in J, we tested whether pdJ is a substrate for galactosyl transferase (UDP-galactose:D-glucose-4-β-galactosyl transferase, EC 2.4.1.22). This enzyme is known to transfer galactose from a UDP-galactose donor to the C4-linked hydroxygroup of D-glucose, with a preference for β-glucosides rather than α-glucosides (Schanbacher and Ebner, 1970). In the marker nucleotides α- and β-glucosyl-HOMedCMP (and the α- and β-glucosyl-HOMedUMP derived from this), the C4-linked hydroxygroup of glucose should be available for the reaction, since the C1-linked hydroxygroup is linked to the base (see Lehman and Pratt, 1960). A diglucosylated HOMedCMP marker nucleotide was isolated from phage T2, in which it replaces 6% of C (Lehman and Pratt, 1960). ^{32}P-labeled dJMP and control nucleotides were incubated with galactosyl transferase, and the reaction products were size separated by polyacrylamide (24%) gel electrophoresis. Figure 4 shows that the glucose-containing marker nucleotides can serve as substrates for galactosyl transferase, resulting in a reaction product with a mobility comparable to that of the marker nucleotide containing two glucose groups (righthand-most lane). The glucosylated U analogs comigrate with their corresponding C analogs because of their mass difference of only 1 g/mol. A slightly different migration was observed, however, for the α versus the β configuration of glucosyl-HOMedUMP or glucosyl-HOMedCMP and also of their galactosylated reaction products. The β-glucosides are indeed better substrates for the enzyme (Figure 4, lanes 4 and 8, showing a complete conversion) than the α glucosides (lanes 4, 8, 10, and 12, only partly converted). HOMedUMP is not a substrate, as expected (Figure 4, lane 2). Lanes 5 and 6 in Figure 4 show that pdJ comigrates with β- rather than with α-glucosyl-HOMedUMP and that it is completely converted by galactosyl transferase into a product comigrating with

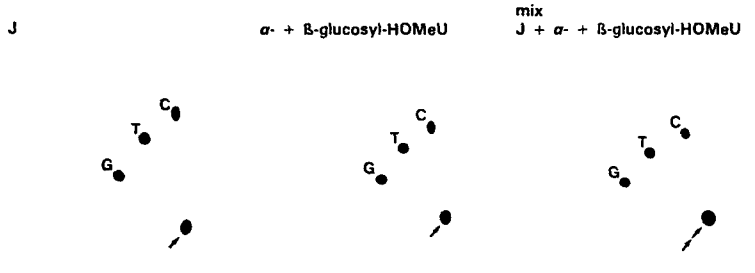

Figure 3. dJMP Comigrates with Glucosylated HOMedUMP during 2D TLC

The autoradiograms show 2D TLC separations on unmodified cellulose sheets of ^{32}P-labeled pdT, pdC, and pdG mixed with pdJ (left panel), α- and β-glucosyl-HOMedUMP (middle panel), and both pdJ and α- and β-glucosyl-HOMedUMP (right panel). Separation conditions and application of the nucleotides to the sheets were as described in Experimental Procedures.

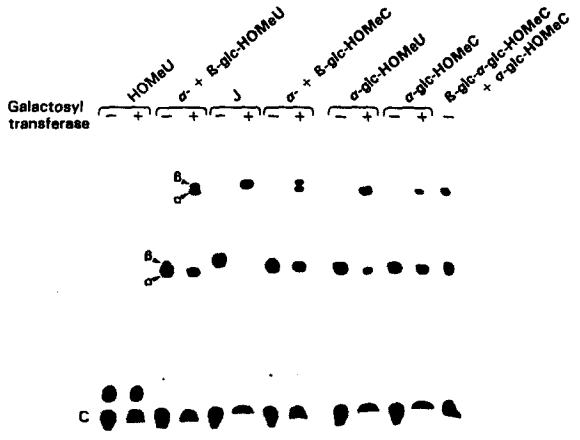

Figure 4. pdJ Is a Substrate for the Enzyme Galactosyl Transferase

pdJ is a substrate for the enzyme galactosyl transferase (UDP-galactose: D-glucose-4-β-galactosyl transferase, EC 2.4.1.22). The autoradiogram shows ^{32}P-labeled deoxyribonucleotides ([^{32}P]pdNs), size fractionated by 24% polyacrylamide gel electrophoresis. Each [^{32}P]pdN was either treated with galactosyl transferase (plus) or not (minus). The base (N) of each nucleotide is indicated above the lanes. In each sample, [^{32}P]pdC (C) was mixed in prior to electrophoresis as an internal reference. α and β indicate the configuration in which the glucose is glycosidically linked to the hydroxygroup of the marker DNA bases HOMeU and HOMeC. The ratio of α-glucosyl-HOMeU and α-glucosyl-HOMeC versus β-glucosyl-HOMeU and β-glucosyl-HOMeC in the lanes marked α plus β is about 4:1.

galactosylated β-glucosyl-HOMedUMP (lane 4). To confirm that the galactosyl transferase has a strong preference for glucose, a 2.5 × 10^6-fold molar excess of either glucose, galactose, or mannose was added to a reaction mixture containing pdJ and limiting amounts of galactosyl transferase. Analysis of the reaction products showed that only glucose is able to compete with pdJ for the enzyme (data not shown). These results indicate that the hexose linked in J to HOMeU is β-D-glucose.

Nuclear Magnetic Resonance Spectroscopy of Purified dJp

To verify the nature of the hexose and the mode of its attachment to the pyrimidine ring, we analyzed 7 nmol of dJp by nuclear magnetic resonance (NMR). The ^{1}H-NMR spectral data of dJp are presented in Table 2. The spectrum is fully compatible with the presence of 5-((β-glucopyranosyloxy)-methyl)-2′-deoxyuridine 3′-monophosphate (Figure 5). Although the exact nature of the pyrimidine ring could not be deduced from this spectrum, all other elements can be recognized. The signals of the deoxyribose residue are readily assigned by comparison with spectral data of dTp (Table 2) (Wood et al., 1974). A difference in p^{2}H, affecting the deuteration of the 3′-phosphate group, probably causes the slightly different chemical shift values of the dJp H2″, H4′, and H5″ signals, which are supported by the more profound upfield shift ($\Delta\delta$ = −0.07 ppm) of the H3′ resonance. The pyrimidine H6 signal appears as a singlet at 8.081 ppm. Together with its downfield shift ($\Delta\delta$ = 0.41 ppm) and the presence of two doublets with an identical coupling constant of −12.4 Hz at 4.640 ppm and 4.522 ppm, this indicates a substituted HOMe group at the 5 position of the pyrimidine ring. The chemical shift values of the substituent closely resemble those of methyl-β-D-glucopyranoside (Table 2) (Perkins et al., 1977). The anomeric proton resonates at 4.512 ppm, both its position and its coupling constant (J_{12} = 7.9 Hz)

Table 2. Chemical Shifts and Coupling Constants of dJp and Reference Compounds

		Chemical Shift in						
Residue	Reporter Group	dJp	dTp	Methyl-β-D-Glucose	Coupling Constant	dJp	dTp	Methyl-β-D-Glucose
Deoxyribose	H1′	6.317	6.30	–	$J_{1'2'}$	6.9	7.2	–
	H2′	2.404	2.40	–	$J_{1'2''}$	6.5	6.5	–
	H2″	2.559	2.54	–	$J_{2'2''}$	−14.0	−14.3	–
	H3′	4.682	4.75	–	$J_{2'3'}$	6.9	6.8	–
	H4′	4.147	4.17	–	$J_{2''3'}$	4.0	3.7	–
	H5′	3.863	3.86	–	$J_{3'4'}$	3.9	4.0	–
	H5″	3.822	3.80	–	$J_{4'5'}$	3.3	3.5	–
					$J_{4'5''}$	4.6	4.8	–
					$J_{5'5''}$	−12.7	−12.5	–
Pyrimidine[a]	H6	8.081	7.67	–	$J_{7'7}$	−12.4	–	–
	H7	4.640	–	–				
	H7′	4.522	–	–				
Glucose	H1	4.512	–	4.374	J_{12}	7.9	–	8.1
	H2	3.270	–	3.254	J_{23}	9.2	–	9.2
	H3	3.480	–	3.483	J_{34}	9.2	–	9.4
	H4	3.373	–	3.373	J_{45}	9.4	–	9.4
	H5	3.440	–	3.458	J_{56}	2.1	–	1.7
	H6	3.899	–	3.921	$J_{56'}$	5.8	–	6.1
	H6′	3.707	–	3.726	$J_{66'}$	−12.5	–	−12.4

^{1}H chemical shifts (in parts per million) and coupling constants (in hertz) of dJp, and the reference compounds dTp (Wood et al., 1974) and methyl-β-D-glucose (Perkins et al., 1977). Chemical shifts are given in parts per million downfield from sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured in ^2H_2O relative to internal acetone (δ = 2.225 ppm; Vliegenthart et al., 1983) at 290 K and p^{2}H 7.5. (Reference compounds at 303 K [p^{2}H 5.8] and 298 K [p^{2}H 7], respectively.)

[a] The resonance at lower field was tentatively assigned to H7.

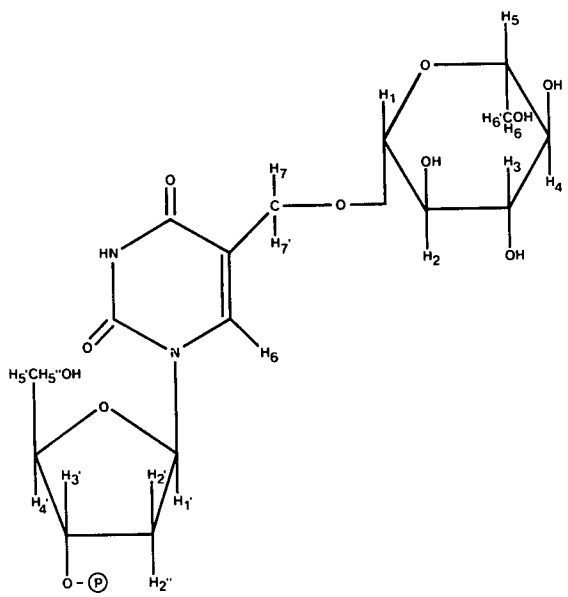

Figure 5. Structure of dJp

The structure of dJp (β-D-glucosyl-HOMedUp) indicating the numbering of the hydrogen atoms as referred to in Table 2. The H_7 and $H_{7'}$ are interchangeable.

being indicative of the β configuration (Vliegenthart et al., 1983). The strong vicinal coupling that is observed for all ring protons is characteristic for β-glucopyranose (Vliegenthart et al., 1983), and the assignment is further substantiated by the coupling constants of the HOMe group, which are in full agreement with a β-glucopyranose residue (Nishida et al., 1984). The detection of only one anomeric proton, indicative for the β configuration, shows that the C1-linked hydroxygroup of the glucose is involved in the linkage to the pyrimidine ring.

Discussion

Only three different modified bases have thus far been identified as normal constituents in eukaryotic DNA: 5-methyl-C, N^6-methyladenine, and 5-HOMeU (reviewed by Rae and Steele, 1978). Our work adds a fourth base to this list: J, also known as β-D-glucosyl-HOMeU, formally 5-((β-D-glucopyranosyloxy)-methyl)-uracil (see Figure 5 for the corresponding 3'-nucleotide dJp). Solving the structure has not been easy. Since the nucleotide represents only 0.1% of the DNA of bloodstream T. brucei, it was not feasible to obtain large amounts of purified dJp. A sensitive mass spectrometry technique for structure determination, GCMS, resulted only in the detection of HOMeU, as J was unstable during the procedure used, and electrospray ionization mass spectrometry only gave the molecular weight of dJp. Nevertheless, a combination of mass spectrometry, chromatographic procedures, NMR, and biochemistry has yielded an unambiguous structure. That J is a HOMeU analog is shown by GCMS analysis of derivatized base J and is consistent with the UV absorbance spectrum (Table 1) and the observed resistance of J to nitrous acid treatment. The molecular weight of intact dJp, determined by electrospray ionization mass spectrometry, is consistent with J being hexosylated HOMeU (Figure 2). That pdJ is β-hexosyl-HOMedUMP is shown by its comigration with β-glucosyl-HOMedUMP in 2D on both unmodified (Figure 3) and PEI cellulose TLC sheets, as well as in 24% polyacrylamide gels (Figure 4). The ability of galactosyl transferase to convert J completely into a reaction product comigrating with galactosylated β-glucosyl-HOMedUMP (Figure 4) shows that the hexose is β-D-glucose, since the enzyme is specific for the D form of glucose and reacts preferentially with β rather than with α glucosides (Schanbacher and Ebner, 1970), and that the C4-linked hydroxygroup of the glucose is free, indicating a glycosidic linkage. The NMR results in Table 2 prove that the hexose is β-glucopyranose and that the anomeric carbon (C1)-linked OH group is involved in the covalent linkage to the HOMe group of HOMedUMP. Taken together, these data unambiguously establish that J is 5-((β-D-glucopyranosyloxy)-methyl)-uracil, a base not found previously in DNA of eukaryotes, prokaryotes, viruses, or bacteriophages. Rather than using the official name, we will use the name β-D-glucosyl-HOMeU, in analogy to the commonly used name glucosyl-HOMeC.

Where Is J in DNA and How Is It Introduced?

J has the properties expected for the modified base responsible for incomplete cleavage of PstI and PvuII sites in inactive subtelomeric VSG genes from bloodstream trypanosomes: it is only present in the bloodstream phase of the trypanosome life cycle and is enriched in minichromosomes, which have a high ratio of (sub)telomeric to total DNA (Gommers-Ampt et al., 1991). Since no other nucleotide with these properties was detected in T. brucei DNA, we concluded that J is the long-sought modified base. There is more pdJ in total T. brucei DNA (0.1% of the nucleotides), however, than can be accommodated in subtelomeric PstI and PvuII sites only. Additional J could be present in (sub)telomeric sequences of DNA not yet analyzed with restriction enzymes (Bernards et al., 1984), in segments with few restriction enzyme sites, such as the (sub)telomeric repeats, or in both.

How J is introduced into DNA is not known. We expect that T will be converted into HOMeU at the DNA level by a sequence-specific enzyme, a thymine hydroxylase, followed by glucose addition by a β-glucosyl transferase, analogous to the enzyme that converts the HOMeC in phage DNA into β-glucosyl-HOMeC (Kornberg et al., 1961; Kornberg and Baker, 1992). This speculative scheme is compatible with our finding that HOMeU is present in higher levels in bloodstream-form DNA than in insect-form DNA, in which J is absent (Gommers-Ampt et al., 1991, 1993). The excess in the former may represent the precursor of J. Further support for this scheme comes from our recent unpublished experiments with insect-form T. brucei, grown in the presence of HOMedU. Trypanosomes stably transfected with the herpes simplex virus thymidine kinase gene readily incorporate exogenous HOMedU into their DNA. We have obtained transformants in which up to 20% of T has been replaced by HOMeU. These also

contain about 0.1% J. Trypanosomes grown in the absence of HOMedU contain less than 0.0002% J (Gommers-Ampt et al., 1991). These results indicate that trypanosomes can glucosylate HOMeU in DNA.

The presence of J in other organisms than African trypanosomatids cannot be excluded yet (cf. Gommers-Ampt et al., 1991). If J is only present in parasites, it may provide a new target for drug development.

What Could Be the Function of J in Trypanosome DNA?

Thus far, J has only been found in trypanosomatids undergoing antigenic variation. The modification is present in (sub)telomeric DNA and only in the bloodstream form of the trypanosome life cycle in which antigenic variation occurs. Only inactive but not active expression sites are affected by modification. A role in the silencing of subtelomeric VSG gene expression sites therefore seems plausible. Two possibilities can be envisaged: DNA modification causes telomere silencing (Bernards et al., 1984), or modification is the consequence of silencing and acts to stabilize or maintain the inactive state of all but one VSG expression site telomere, thereby preventing the formation of patchwork trypanosome coats. Such a function for J would be analogous to the function envisaged for 5-methyl-C in mammals (see Bird, 1992; Lyon, 1993).

With the structure of J in hand, it is now possible to study the function of this modified base. It should be possible to identify the genes required for J biosynthesis or recognition and to disrupt all copies of these genes, as trypanosomes only insert exogenous DNA into their genome by homologous recombination, making gene disruption relatively easy (ten Asbroek et al., 1991, 1993; Lee and Van der Ploeg, 1990; Eid and Sollner-Webb, 1991). The results of these experiments may also be relevant to gene silencing by DNA modification in other organisms.

Experimental Procedures

Trypanosomes

The trypanosomes used belong to strain 427 of T. brucei brucei. Bloodstream-form trypanosomes were grown and isolated as described in Gommers-Ampt et al. (1991). Procyclic (insect-form) trypanosomes were grown in the semidefined medium described by Brun and Schönenberger (1979)

Isolation and Enzymatic Hydrolysis of DNA

Total genomic trypanosome DNA was isolated as described by Bernards et al. (1981) and resuspended in 2 mM Tris–HCl (pH 7.6). The DNA was enzymatically hydrolyzed to deoxyribonucleoside 3′-monophosphates (dNps) as described by Gommers-Ampt et al. (1991). Prior to liquid chromatography, the reaction mixture was heated at 100°C for 5 min and centrifuged to remove precipitated material that would interfere with subsequent purification.

Liquid Chromatography

dNps, derived from enzymic hydrolysis of 100 μg of total genomic T. brucei DNA, were separated on a Supelco LC-18-S reversed-phase high performance liquid chromatography column as described by Gommers-Ampt et al. (1993). The elution time of dJp was determined by ^{32}P-postlabeling and by 2D TLC analysis of collected fractions as described (Gommers-Ampt et al., 1993). Fractions containing dTp (and dJp in the bloodstream-form eluate) were chromatographed on a Mono Q anion exchange column (Pharmacia) as described (Gommers-Ampt et al., 1993), and dJp-containing fractions were collected. As a blank control sample for subsequent GCMS, fractions derived from the Mono Q column loaded with reversed-phase column-derived nucleotides from insect-form T. brucei and corresponding to the retention time of dJp (13 min) were used. The UV absorbance spectrum of each eluted nucleotide was measured with a photodiode array detector (Waters 990). Since dJp comigrated with dTp on the C18 column (Figure 1A) when total genomic hydrolysate was used, it was rechromatographed on the C18 column after isolation from the Mono Q column (pH 3.5) to obtain its spectrum at pH 5.6.

Isolation and Purification of ^{32}P-Labeled Deoxyribonucleoside 5′-Monophosphates

^{32}P-postlabeling analysis of total genomic DNA and separation of the resulting ^{32}P-labeled nucleotides in 2D on PEI cellulose TLC sheets were modified by Gommers-Ampt et al. (1991) from Gupta et al. (1982). The genomic DNAs used were from T. brucei bloodstream-form variants 1.8 and 118a′, Bacillus subtilis bacteriophage H1 (Gommers-Ampt et al., 1991), Escherichia coli bacteriophage T4 (Sigma), and E. coli bacteriophage T2 (Sigma) for the isolation of dJMP, dTMP, dGMP, and dCMP; HOMedUMP; α-glucosyl-HOMedCMP and β-glucosyl-HOMedCMP; and α-glucosyl-HOMedCMP and β-glucosyl-α-glucosyl-HOMedCMP, respectively. All the modified nucleotides could be separated from the normal nucleotides as well as from each other on PEI cellulose (data not shown). The ^{32}P-labeled glucosylated HOMedUMPs were obtained by nitrous acid deamination of their corresponding cytidine analogs derived from phages T2 and T4 as described below. For the isolation of [^{32}P]pdU, deoxyuridine 3′-monophosphate (Sigma) was ^{32}P-postlabeled and chromatographed in 1D on PEI cellulose. The nucleotides of interest were excised from the sheets, washed three times with methanol, and eluted from the sheet with 500 μl of 0.2 M ammonium formate (pH 3.5). After brief centrifugation, the supernatant was collected and dried in vacuo. Residual ammonium sulphate, used for elution in the second dimension, was removed by Sephadex G10 exclusion chromatography in 50 mM ammonium formate (pH 3.5).

2D TLC

Labeled nucleotides were separated by 2D TLC using either unmodified cellulose sheets (Merck) or PEI-impregnated sheets (Polygram Cel 300 PEI, Macherey Nagel). Elution conditions were as described by Gommers-Ampt et al. (1991). Prior to application to the sheets, the purified nucleotides were dried and resuspended in either 7 mM bicine NaOH (pH 9.6), 20 mM NaAc (pH 5.3) (for the unmodified cellulose) or 18 mM NaAc (pH 7.2) (for the PEI sheets).

Chemical Deamination of Nucleotides

Nitrous acid deamination was done according to Wyatt and Cohen (1953). After the reaction, the nucleotide was purified from the salt using Sephadex G10 exclusion chromatography, with 50 mM ammonium formate (pH 3.5) as elution buffer. The identity of the marker nucleotide-glucosylated HOMedUMP (obtained after chemical deamination of the corresponding C analog derived from bacteriophage T4 DNA) was verified with thermospray mass spectrometry. The high pressure liquid chromatography–purified C analog before deamination and the U analog after deamination yielded the expected molecular weights of 499 and 500, respectively (data not shown).

Acid Hydrolysis, Trimethylsilylation, and GCMS

Acid hydrolysis, trimethylsilylation, and GCMS were performed essentially as described previously (Gajewski et al., 1990). In brief, the nucleotides (0.25 nmol) were lyophilized and then hydrolyzed with formic acid (88%) in evacuated and sealed tubes at 140°C for 30 min. After hydrolysis, the samples were lyophilized and trimethylsilylated with 0.1 ml of a mixture of N,O-bis(trimethylsilyl)trifluoroacetamide and acetonitrile (2:1 [v:v]) by heating for 30 min at 130°C. Analysis of derivatized samples was performed by using a mass-selective detector interfaced to a gas chromatograph (both from Hewlett Packard). From each sample, 5 μl was injected into the injection port (kept at 250°C) of the gas chromatograph. A split ratio of 1:10 was used, resulting in 2.5 pmol of hydrolyzed and derivatized nucleotide going through the GC column. A fused silica capillary column (12.5 mm × 0.2 mm) coated

with cross-linked 5% phenylmethylsilicone gum phase (film thickness, 0.33 μm; Hewlett Packard) was used for the separation, with helium as the carrier gas. The temperature of the column was increased from 150°C to 260°C at a rate of 2°C per minute after 2 min at 150°C. Selected ion monitoring was performed in the electron ionization mode at 70 eV. The temperature of the ion source of the mass spectrometer was kept at 250°C.

Electrospray Ionization Mass Spectrometry of dJp

The electrospray mass spectrum of dJp was obtained on a Vestec 201 mass spectrometer (Vestec, Houston, Texas) controlled by a Vector One data system (Teknivent Corporation, St. Louis, Missouri). Instrument conditions were as follows: −2.0 kV (needle), −400 V (nozzle), −10 V (repeller); ion current, −0.10 μA; source temperature, 50°C. A solution of dJp (approximately 4 pmol/μl in MeOH–H_2O [9:1]) was continuously infused into the instrument at a rate of 1 μl/min using a Harvard 22 syringe pump. Ten 7 s scans were averaged to obtain the spectrum shown in Figure 2. The ion profile from m/z 490 to m/z 510 (inset in Figure 2), from which the $(M-H)^-$ value of dJp was measured, was obtained from a single 10 s scan acquired in the data system calibration mode.

Polyacrylamide (24%) Gel Electrophoresis

The dried nucleotides were dissolved in 5 μl of loading dye (95% formamide, 10 mM EDTA, 0.1% bromophenolblue, 0.1% xylene cyanol), heated for 2 min at 80°C, and applied to a 24% polyacrylamide (acrylamide:bisacrylamide, 38:2), 7 M urea, 1 × TBE (90 mM Tris, 90 mM boric acid, 2 mM EDTA [pH 8.0]) gel. Electrophoresis was performed at 65 W for 2 hr.

Galactosyl Transferase Treatment

The galactosyl transferase (UDP-galactose:D-glucose-4-β-galactosyl transferase, EC 2.4.1.22) reactions contained 1 pmol of [^{32}P]pdN, 4 mM UDP-galactose (Sigma), 50 mM glycine (pH 8.5), 5 mM $MnCl_2$, and 0. 2 U of galactosyl transferase (Sigma) in a total volume of 10 μl and were incubated for 2.5 hr at 37°C. From each sample as well as from unreacted [^{32}P]pdNs, a volume equivalent to 15 cps was mixed with an equal amount of [^{32}P]pdC (15 cps), dried in vacuo, dissolved in 5 μl of formamide-loading dye (as described above), and subjected to 24% polyacrylamide gel electrophoresis as described above. For the competition experiment, 0.4 pmol of [^{32}P]pdJ was incubated as described above but with 0.02 U of galactosyl transferase with or without 1 μmol of hexose. Each sample (2 μl) was applied on a PEI TLC sheet, which was eluted in 1D as described (Gommers-Ampt et al., 1991), allowing the separation of nucleotides containing two hexoses from those having one hexose substituent.

^{1}H-NMR Spectroscopy

Prior to ^{1}H-NMR spectroscopic analysis, dJp was repeatedly exchanged in ^2H_2O (99.9%; Merck isotopes) with intermediate lyophilization. The sample was finally dissolved in 99.96% ^2H_2O (p^2H 7.5). The resolution-enhanced 1D 600 MHz ^{1}H-NMR spectrum was recorded at a Bruker AM-600 spectrometer (Bijvoet Center, Department of NMR Spectroscopy, Utrecht University, Utrecht, The Netherlands), operating at a probe temperature of 290 K to avoid disturbance by the HOD resonance. Chemical shifts are given relative to sodium-4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured indirectly to internal acetone in ^2H_2O (δ = 2.225 ppm) (Vliegenthart et al., 1983).

Acknowledgments

We thank Dr. J. A. McCloskey (Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah), Dr. G. van de Werken and G. J. ten Hove (National Institute of Public Health and Environmental Protection, Bilthoven, The Netherlands), Dr. N. M. M. Nibbering (Institution for Mass Spectrometry, Amsterdam University, Amsterdam), Dr. D. H. van den Eijnden (Department of Medical Chemistry, Free University, Amsterdam), Dr. C. Vink (The Netherlands Cancer Institute, Amsterdam), and B. Penninkhof (Department of Bio-Organic Chemistry, Utrecht University, Utrecht, The Netherlands) for helpful discussions; Dr. W. Bitter, P. A. Blundell, Dr. M. J. L. Ligtenberg, Dr. R. H. A. Plasterk, Dr. G. Rudenko, Dr. A. H. Schinkel, and Dr. M. C. Taylor for critical reading of the manuscript; and G. J. Stroomberg for technical assistance. This work was supported by grants from the Netherlands Foundation for Chemical Research, with financial support of the Netherlands Organisation for Scientific Research (NWO), and from the National Institutes of Health (GM 29812).

Received July 29, 1993; revised September 27, 1993.

References

Bernards, A., Van der Ploeg, L. H. T., Frasch, A. C. C., Borst, P., Boothroyd, J. C., Coleman, S., and Cross, G. A. M. (1981). Activation of trypanosome surface glycoprotein genes involves a duplication–transposition leading to an altered 3′ end. Cell *27*, 497–505.

Bernards, A., Van Harten-Loosbroek, N., and Borst, P. (1984). Modification of telomeric DNA in *Trypanosoma brucei*: a role in antigenic variation? Nucl. Acids Res. *12*, 4153–4170.

Bird, A. (1992). The essentials of DNA methylation. Cell *70*, 5–8.

Borst, P. (1991). Molecular genetics of antigenic variation. Immunol. Today *12*, A29–A33.

Brun, R., and Schönenberger, M. (1979). Cultivation and *in vitro* cloning of procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Acta Trop. *36*, 289–292.

Cross, G. A. M. (1990). Cellular and genetic aspects of antigenic variation in trypanosomes. Annu. Rev. Immunol. *8*, 83–110.

Crozatier, M., De Brij, R. J., Den Engelse, L., Johnson, P. J., and Borst, P. (1988). Nucleoside analysis of DNA from *Trypanosoma brucei* and *Trypanosoma equiperdum*. Mol. Biochem. Parasitol. *31*, 127–132.

Dizdaroglu, M. (1984). The use of capillary gas chromatography–mass spectrometry for identification of radiation-induced DNA base damage and DNA base-amino acid cross-links. J. Chromatograph. *295*, 103–121.

Dizdaroglu, M. (1991). Chemical determination of free radical-induced damage to DNA. Free Radical Biol. Med. *10*, 225–242.

Eid, J. E., and Sollner-Webb, B. (1991). Stable integrative transformation of *Trypanosoma brucei* that occurs exclusively by homologous recombination. Proc. Natl. Acad. Sci. USA *88*, 2118–2121.

Gajewski, E., Rao, G., Nackerdien, Z., and Dizdaroglu, M. (1990). Modification of DNA bases in mammalian chromatin by radiation-generated free radicals. Biochemistry *29*, 7876–7882.

Gommers-Ampt, J., Lutgerink, J., and Borst, P. (1991). A novel DNA nucleotide in *Trypanosoma brucei* only present in the mammalian phase of the life-cycle. Nucl. Acids Res. *19*, 1745–1751.

Gommers-Ampt, J. H., Teixeira, A. J. R., van de Werken, G., van Dijk, W. J., and Borst, P. (1993). The identification of hydroxymethyluracil in DNA of *Trypanosoma brucei*. Nucl. Acids Res. *9*, 2039–2043.

Gupta, R. C., Reddy, M. V., and Randerath, K. (1982). ^{32}P-postlabeling analysis of non-radioactive aromatic carcinogen–DNA adducts. Carcinogenesis *3*, 1081–1092.

Kornberg, A., and Baker, T. A. (1992). DNA Replication (San Francisco: W. H. Freeman).

Kornberg, S. R., Zimmerman, S. B., and Kornberg, A. (1961). Glucosylation of deoxyribonucleic acid by enzymes from bacteriophage-infected *Escherichia coli*. J. Biol. Chem. *236*, 1487–1493.

Lee, M. G.-S., and Van der Ploeg, L. H. T. (1990). Homologous recombination and stable transfection in the parasitic protozoan *Trypanosoma brucei*. Science *250*, 1583–1587.

Lehman, I. R., and Pratt, E. A. (1960). On the structure of the glucosylated hydroxymethylcytosine nucleotides of coliphages T_2, T_4 and T_6. J. Biol. Chem. *235*, 3254–3259.

Lichtenstein, J., and Cohen, S. S. (1960). Nucleotides derived from enzymatic digests of nucleic acids of T_2, T_4 and T_6 bacteriophages. J. Biol. Chem. *235*, 1134–1141.

Lyon, M. F. (1993). Epigenetic inheritance in mammals. Trends Genet. *9*, 123–128.

Nishida, Y., Ohrui, H., and Meguro, H. (1984). ^{1}H-NMR studies of (6R)- and (6S)-deuterated D-hexoses: assignment of the preferred rotamers about C5–C6 bond of D-glucose and D-galactose derivatives in solu-

tions. Tetrahed. Lett. *25*, 1575–1578.

Pays, E., Delauw, M. F., Laurent, M., and Steinert, M. (1984). Possible DNA modification in GC dinucleotides of *Trypanosoma brucei* telomeric sequences: relationship with antigen gene transcription. Nucl. Acids Res. *12*, 5235–5247.

Perkins, S. J., Johnson, L. N., Philips, D. C., and Dwek, R. A. (1977). High-resolution ^{1}H- and ^{13}C-NMR spectra of D-glucopyranose, 2-acetamido-2-deoxy-D-glucopyranose, and related compounds in aqueous media. Carbohydr. Res. *59*, 19–34.

Rae, P. M. M., and Steele, R. E. (1978). Modified bases in the DNAs of unicellular eukaryotes: an examination of distributions and possible roles, with emphasis on hydroxymethyluracil in dinoflagellates. Biosystems *10*, 37–53.

Schanbacher, F. L., and Ebner, K. E. (1970). Galactosyltransferase acceptor specificity of the lactose synthetase A protein. J. Biol. Chem. *245*, 5057–5061.

Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J., and Udseth, H. R. (1990). New developments in biochemical mass spectrometry: electrospray ionization. Anal. Chem. *62*, 882–899.

ten Asbroek, A. L. M. A., Ouelette, M., and Borst, P. (1991). Targeted insertion of the neomycin phosphotransferase gene into the tubulin gene cluster of *Trypanosoma brucei*. Nature *348*, 174–175.

ten Asbroek, A. L. M. A., Mol, C. A. A. M., Kieft, R., and Borst, P. (1993). Stable transformation of *Trypanosoma brucei*. Mol. Biochem. Parasitol. *59*, 133–142.

Van der Ploeg, L. H. T. (1987). Control of variant surface antigen switching in trypanosomes. Cell *51*, 159–161.

Vliegenthart, J. F. G., Dorland, L., and van Halbeek, H. (1983). High-resolution ^{1}H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. Adv. Carbohydr. Chem. Biochem. *41*, 209–374.

Wood, D. J., Hruska, F. E., and Ogilvie, K. K. (1974). Proton magnetic resonance studies of 2'-deoxythymidine, its 3'- and 5'-monophosphates and 2'-deoxythymidylyl-(3',5')-2'-deoxythymidine in aqueous solution. Can. J. Chem. *52*, 3353–3366.

Wyatt, G. R., and Cohen, S. S. (1953). The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. Biochem. J. *55*, 774–782.