PROTEINS: Structure, Function, and Genetics 28:568–579 (1997)

Flexible Glycine Rich Motif of *Escherichia coli* Deoxyuridine Triphosphate Nucleotidohydrolase Is Important for Functional But Not for Structural Integrity of the Enzyme

Beata G. Vertessy*
Institute of Enzymology, Biological Research Center, Hungarian Academy of Science, Budapest, Hungary

ABSTRACT **Deoxyuridine triphosphate nucleotidohydrolase (dUTPase), a ubiquitous enzyme of DNA metabolism, has been implicated as a novel target of anticancer and antiviral drug design. This task is most efficiently accomplished by X-ray crystallography of the relevant protein–inhibitor complexes. However, the topic of the present investigation, a glycine-rich strictly conserved structural motif of dUTPases, could not be located in the crystal structure of the *Escherichia coli* enzyme, probably due to its increased flexibility. The present work shows that removal of a C-terminal 11-residue fragment, including this motif, by limited trypsinolysis strongly impairs catalytic activity. Kinetic analysis of the intact and digested variants showed that k_{cat} decreases 40-fold, while K_M increases less than twofold upon digestion. The tryptic site was identified by mass spectrometry, amino acid analysis and N-terminal sequencing. The shortened enzyme variant retains the secondary, tertiary, and quaternary (trimeric) structure of the intact species as suggested by UV absorption, fluorescence and circular dichroism spectroscopy, and analytical gel filtration. Moreover, binding affinity of the shortened variant toward the substrate analogue MgdUDP is identical to the one displayed by the intact enzyme. I conclude that the glycine-rich motif is functionally relevant for *E. coli* dUTPase. It may play a role in enzymatic catalysis by contributing to the formation of the catalytically potent enzyme–substrate complex. Proteins 28:568–579, 1997.** © 1997 Wiley-Liss, Inc.

Key words: Gly-rich motif; phosphate binding P-loop; Motif 5 of dUTPases; MgdUDP binding; limited trypsinolysis; circular dichroism spectroscopy; secondary structure determination

INTRODUCTION

The fidelity of DNA replication is of ultimate importance in all living organisms and is accounted for by several regulatory mechanisms (e.g., proofreading, excision repair, mismatch repair), which repair the mistake but do not prevent its occurrence.[1] On the other hand, the ubiquitous enzyme dUTPase, catalyzing the pyrophosphate cleavage of the substrate dUTP into the products dUMP and anorganic pyrophosphate, employs preventive measures: it prevents uracil incorporation into DNA by efficiently decreasing the dUTP/dTTP ratio. The enzyme is essential for viability in organisms as different as yeast[2] and *Escherichia coli*.[3] dUTPase deficiency leads to the appearance of highly uracil-substituted DNA, which, upon excision repair in the presence of a considerable dUTP pool, is excessively fragmented causing cell death.[4] dUTPase also provides dUMP, the precursor of dTTP biosynthesis. Based on the significant role of the enzyme in DNA metabolism (reviewed in Ref. 5), it was proposed as a novel target for anticancer drug design. Accordingly, it has been shown that resistance to the widely used antineoplastic drug fluorodeoxyuridine correlates to an elevation of dUTPase activity,[6] raising the hope that dUTPase inhibitors may improve therapeutic effectiveness. In fact, membrane-permeable substrate-analogous inhibitors of dUTPase were demonstrated to be efficient against human cancer cells in vitro.[7]

The enzyme is developmentally regulated in higher organisms, the highest activity is found in the nondifferentiated, actively proliferating cells.[8,9] Despite the ubiquity of the enzyme, several viruses, including different herpes- and retroviruses, encode their own dUTPase.[10] dUTPase deficiency induced by mutation of viral genes strongly impairs virulence of herpes simplex virus type 1[11] and the retrovirus

Abbreviations: dUTPase, deoxyuridine triphosphate nucleotidohydrolase, EC 3.6.1.23; HuBaR dUTPases, dUTPases of *hu*man, *ba*cterial or *r*etroviral origin; SDS PAGE; sodium dodecyl sulfate polyacrylamide gel electrophoresis; CD, circular dichroism spectroscopy; $K_{d,app}$, apparent dissociation constant; UV, ultraviolet

*Correspondence to: Beata G. Vertessy, Institute of Enzymology, Biological Research Center, Hungarian Academy of Science, Budapest, POB 7, H1518, Hungary.
E-mail: vertessy@enzim.hu
Received 7 February 1997; Accepted 4 March 1997

equine infectious anemia virus[12] in nonproliferating host cells. Consequently, dUTPase is regarded as a promising new target for antiviral drug design as well.

State-of-the-art drug design requires detailed structure–function analysis of relevant enzyme–inhibitor complexes. Structurally, two main dUTPase subgroups can be distinguished[10]: in one group (termed here as herpeslike dUTPases), the monomeric enzyme protein has ~320–380 amino acid residues (e.g., herpes simplex). In the other group (termed here as HuBaR dUTPases for human, bacterial and retroviral sources), the enzyme is a trimer of identical subunits of ~141–160 residues, that is, the approximate half of the herpeslike monomer. The trimeric structure is experimentally proven for the protein from human and rat cells, *E. coli* and equine infectious anemia virus.[13–16] It has been suggested that the herpeslike monomer arose by gene duplication of an ancient dUTPase gene, since the two halves of the ~350-residue polypeptide show significant homology to each other and to the HuBaR dUTPase sequences.[10] The polypeptide chain in both subgroups contains five conserved sequence motifs, numbered 1 through 5 from the N terminus in the HuBaR subgroup, but in the herpeslike subgroup the order of the motifs is altered.[10] In previous studies we showed that an exposed tyrosine residue is functionally relevant in the *E. coli* and the equine infectious anemia virus enzyme, and it is involved in the accommodation of the substrate analogue MgdUDP in the active site of the *E. coli* enzyme.[17,18] The residue was identified as the strongly conserved Tyr-93 of Motif 3 in the *E. coli* enzyme,[18] and the MgdUDP binding site was localized in the shallow cavity seen in the apoenzyme crystal structure.[18,15] The role of Tyr-93 in dUDP binding was recently confirmed in the study presenting the crystal structure of the *E. coli* dUTPase-dUDP complex.[19] An intriguing feature of both this and the previously published apoenzyme structure is that the C-terminal 16 residues, containing the conserved Motif 5, could not be located due to their increased flexibility.[15,19] Residues from all the other four conserved motifs were found to interact with the substrate analogue dUDP, either directly or indirectly, through water molecules.[19] Unfortunately, the invisible C terminus prevented any solid attempt to assign a role to the fifth conserved motif, although reasonable speculations based on sequence comparisons were offered.[19]

In the present work, I probed the structural and functional role of the C-terminal conserved Motif 5, employing tryptic deletion analysis. The shortened enzyme variant is characterized with respect to enzyme kinetic parameters and substrate analogue (MgdUDP) binding affinity. Its structure is investigated by SDS PAGE, amino acid analysis, N-terminal sequencing, mass spectrometry, together with UV absorption, CD and fluorescence spectroscopy, and analytical gel filtration.

MATERIALS AND METHODS

dUTP and low pressure column chromatography materials (Q-Sepharose, Sephacryl) were purchased from Pharmacia, Uppsala, Sweden. dUDP, Tes and Hepes buffer were from Sigma, St. Louis, MO, Phenol Red indicator came from Merck, Darmstadt, Germany. All other materials were of analytical grade purity.

Enzyme Preparation and Assay

dUTPase from *E. coli* was purified as previously.[20] The purified preparation appeared as a single band on SDS PAGE gels when investigated by laser densitometry (see below), suggesting at least 98% purity. Protein concentration was measured spectrophotometrically using $A^{0.1\%}$ 1 cm, 280 = 0.52.[18] Molecular mass of the trimeric *E. coli* enzyme is 49 kDa.[15] Throughout the present study, molar enzyme concentrations refer to the monomeric species. Before use, aliquots of the enzyme were dialyzed against respective buffers. Enzyme activity was routinely assayed at enzyme concentrations 25–50 nM in the presence of 40 μM dUTP and 1000 μM $MgCl_2$ (providing saturating excess of Mg^{2+}) in 1 mM Tes/HCl, pH 7.5 containing 150 mM KCl and 40 μM Phenol Red indicator (assay buffer). Proton release during the transformation of dUTP into dUMP and PP_i was followed at 559 nm at 25°C,[18] using Hewlett-Packard 8451A or JASCO-V550 spectrophotometers and 10-mm pathlength thermostatted cuvettes. Initial velocity was determined from the slope of the initial part (first 10 seconds) of the progress curve. This method was used to conveniently follow activity change during trypsinolysis. The enzyme kinetic parameters k_{cat} and K_M were determined from the entire progress curve of dUTP cleavage by using the integrated Michaelis-Menten equation,[21] as described previously[22]:

$$\frac{[P]_t}{t} = V_{max} - \frac{K_M}{t} \ln \frac{[S]_0}{[S]_0 - [P]_t} \qquad (1)$$

where $[P]_t$ is product concentration at reaction time t, V_{max} ($=k_{cat}[E]$, where $[E]$ is enzyme, i.e., dUTPase concentration) and K_M are the apparent maximal velocity and Michaelis constant, respectively, $[S]_0$ is substrate concentration at $t = 0$. $[P]_t$ was calculated from the absorbance change in the indicator assay measured at 559 nm

$$[P]_t = \frac{\Delta A_t}{\Delta A_{max}} [S]_0 \qquad (2)$$

where $\Delta A_t = A_t - A_0$, and $\Delta A_{max} = A_{final} - A_0$, A_t, A_0, and A_{final} are absorbance values at reaction time t, and at the beginning and at the end of the reaction,

respectively. Equation (2) holds true if i) the reaction is practically irreversible, which is the case for the dUTP → dUMP + PP_i pyrophosphorolysis as catalyzed by dUTPase (cf. Ref. 22), and ii) the absorbance signal, due to the differential absorbance of the acidic and basic form of the pH indicator Phenol Red, is essentially linearly related to product formation. The fulfillment of this second criteria was tested in control experiments reacting dUTP at varied concentrations (2–20 μM) with dUTPase. The resulting ΔA_{max} values were linearly related to the substrate concentration, confirming response linearity, in agreement with the observation of Larsson and colleagues.[22] k_{cat} of intact dUTPase was determined to be 5.15 s^{-1}, in accordance with previous investigations.[20,22]

Limited Tryptic Digestion

Limited tryptic digestion of *E. coli* dUTPase was carried out at 30°C, using 2 mg/ml dUTPase concentration and 1:10 trypsin:dUTPase ratio in 10 mM potassium phosphate–150 mM KCl buffer with 1 mM $MgCl_2$, pH 7.5. Aliquots were taken at different time points for activity measurements (initial velocity determinations) and for SDS PAGE. After stopping the digestion by the addition of 1 mM phenylmethanesulfonyl fluoride, the digestion mixture was applied on a Q-Sepharose ion exchange column (66 ml resin volume) equilibrated with 25 mM Hepes buffer, pH 7.0 (Buffer A). The column was washed with 200 ml Buffer A, then a linear gradient of 50 ml Buffer A and 50 ml Buffer A with 0.8 M NaCl was applied. A second linear gradient followed (50 ml Buffer A with 0.8 M NaCl + 50 ml Buffer A with 1.6 M NaCl). Elution was detected at 280 nm, eluted peaks were analyzed by SDS PAGE. A minor peak, corresponding to a protein with a subunit molecular mass of 23 kDa eluted at 0.8 M NaCl and was tentatively identified as trypsin. The main peak (subunit molecular mass estimated as 15.6 kDa) eluted at 1 M NaCl, it was collected, concentrated down to 500 μl volume (using an Amicon ultrafiltration cell, YM 10 membrane from Millipore) and gelfiltrated on a Sephacryl 200 column (internal diameter 2 cm, length 29 cm) equilibrated with 5 mM ammonium bicarbonate. A single peak eluted, which was concentrated as above and submitted for amino acid analysis, Edman degradation, and mass spectrometry, or was used for determination of kinetic parameters using Equations (1) and (2) and analytical gel filtration. For spectroscopic measurements, the digested sample was dialyzed against the respective buffers.

Analytical Gel Filtration

A Sephacryl 200 column was used with an inner diameter of 1.6 cm and a total volume of 162 ml, equilibrated with 10 mM potassium-phosphate–150 mM KCl buffer with 1 mM $MgCl_2$, pH 7.5. Flow rate was 0.93 ml/min. Exclusion volume was determined with Blue Dextran, the column was calibrated with the following proteins: bovine serum albumin, ovalbumin, chymotrypsinogen, and RNase (molecular masses: 67, 43, 25, and 13.7 kDa, respectively). Calibrating proteins or other samples were applied in a total volume of 500 μl, at a concentration of 0.5–1.0 mg/ml.

SDS PAGE

SDS PAGE was performed according to Laemmli[23] by using 16% polyacrylamide gels. Protein bands were visualized by Coomassie Brilliant Blue R-250 staining and quantitative analysis was done by densitometry on a GelDoc densitometer (BioRad).

UV Absorption Spectra

UV absorption spectra of intact and digested dUTPase in 10 mM potassium-phosphate–150 mM KCl buffer with 1 mM $MgCl_2$, pH 7.5 were recorded on a JASCO-V550 UV/VIS spectrophotometer equipped with a data processing system; 10-mm pathlength cuvettes were used, thermostatted at 25°C.

CD Measurements

Near-UV CD spectra (230–350 nm) were recorded on a JASCO 720 spectropolarimeter using 10-mm pathlength cuvettes thermostatted at 25°C. Intact or digested dUTPase from *E. coli* at a concentration of 37 μM was titrated by addition of MgdUDP (0–100 μM) in 10 mM potassium-phosphate–150 mM KCl buffer with 1 mM $MgCl_2$, pH 7.5. Spectra measured immediately after mixing enzyme and MgdUDP were stable for at least 30 minutes. 1 mM $MgCl_2$ provided saturating excess of Mg^{2+}, since results were the same when $MgCl_2$ concentration was increased tenfold in a control experiment. Far-UV CD spectra (down to 184 nm) were recorded on the same instrument in a 1-mm pathlength thermostatted cuvette at 25°C, using the native or the digested protein at 0.2 mg/ml (12.5 μM) in 20 mM potassium-phosphate buffer, pH 7.5. Three scans of every spectrum were averaged. Spectral data processing was done by using the built-in JASCO software of the spectropolarimeter.

Estimation of Secondary Structural Content

Atomic coordinates of the *E. coli* dUTPase apoenzyme, kindly provided by Dr. G. Larsson, Lund University,[15] were processed by the computer program "dssp" (described in Ref. 24) to calculate the content of different secondary structural elements as present in the crystal structure. Far-UV CD spectra were evaluated by several different methods (cf. Ref. 25), using the computer programs kindly distributed by the authors. A personal Iris Silicon Graphics workstation was used for these calculations.

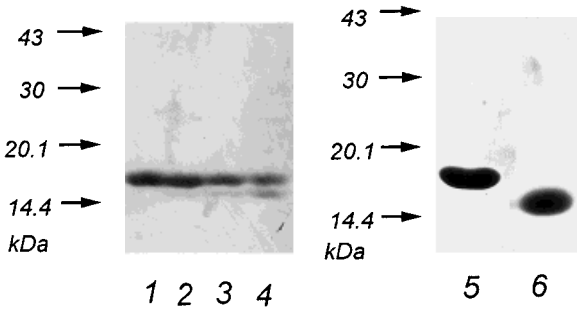

Fig. 1. Electrophoretic analysis of dUTPase trypsinolysis and purification of the digested fragment. **Lane 1:** Intact dUTPase. **Lanes 2, 3, 4:** Digested samples after 5, 10, and 15 minutes of trypsinolysis. **Lane 5:** intact dUTPase. **Lane 6:** Purified digested dUTPase. Calibration kit contained ovalbumin (43 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (20.1 kDa), and α-lactalbumin (14.4 kDa) (cf. respective marks).

Fluorescence Spectroscopy

Emission spectra were recorded by using a JASCO FP 777 spectrofluorimeter. Intact or digested dUTPase at 0.1 mg/ml in 20 mM potassium-phosphate buffer, pH 7.5 was measured in a 5-mm pathlength thermostatted cuvette at 25°C. Excitation wavelength was 280 nm, emission was scanned between 290 and 400 nm. Slit widths for monochromators were set at 1.5 or 3 nm for excitation and emission, respectively.

Amino Acid Analysis

Amino acid analysis was carried out on digested dUTPase at the Analysis and Synthesis Laboratory of the Agricultural Biotechnological Research Center of Gödöllö, Hungary. Samples were hydrolyzed in 6 M HCl at 110°C for 24 hours, followed by neutralization and derivatization, using phenylisothiocyanate. Derivatized amino acids were separated on a reversed phase HPLC column.

Edman Degradation

Edman degradation of trypsinolysed *E. coli* dUTPase was carried out at the Analysis and Synthesis Laboratory of the Agricultural Biotechnological Research Center of Gödöllö, Hungary using the pulsed-liquid phase sequencer ABI 471A of Applied Biosystems, Inc. USA. Samples were dried on glass fiber filter, and the instrument was operated according to the instructions of the manufacturer.

Electrospray Mass Spectrometry

Electrospray mass spectrometry of intact and digested *E. coli* dUTPase was carried out on a Finnigan MAT TSQ 7000 triple-stage quadrupol instrument at the University of József Attila, Szeged, Hungary, according to the operational instructions.

Table I. Change in the Amino Acid Composition of dUTPase on Limited Trypsinolysis*

Residue	Predicted for the intact sample	Digested sample	Predicted with R-141 cut
Asn + Asp	16	16.5	16
Thr	7	7.3	7
Ser	8	7.4	7
Glu + Gln	13	11.8	11
Pro	9	8.3	9
Gly	**18**	**13.3**	**13**
Ala	12	11.8	12
Val	11	10.5	11
Met	6	5.4	6
Ile	10	9.6	10
Leu	16	15.5	16
Tyr	2	2.0	2
Phe	6	5.5	5
His	4	3.1	3
Lys	5	4.5	5
Arg	7	6.9	6

*Tryptophan and cysteine were not determined. The composition predicted from the protein sequence is included for comparison. Data for Gly are highlighted.

RESULTS

Trypsinolysis and Determination of the Cleavage Site

Trypsinolysis of dUTPase was followed by analyzing aliquots from the digestion reaction mixture by SDS PAGE (Fig. 1). Reflecting tryptic cleavage, a new band appeared upon digestion at an apparent molecular mass of ~1 kDa less than the molecular mass of 16 kDa corresponding to the intact species. No lower bands could be detected even after 120 minutes of digestion when trypsin was inactivated by the addition of freshly dissolved phenylmethanesulfonyl fluoride to 1 mM final concentration, and the digestion mixture was purified by ion-exchange and gel filtration chromatography. The purified shortened species appeared as a single band on the gel, suggesting at least 98% purity (Fig. 1).

Amino acid analysis of the digested enzyme showed that Gly content was significantly reduced while Met and Lys content changed only slightly upon trypsinolysis as compared to the respective expected data of the intact enzyme (Table I). This finding suggested that the tryptic cut occurred at the Gly-rich C terminus of the dUTPase polypeptide. N-terminal sequencing of the digested protein confirmed this suggestion: the determined sequence **MetMetLysLysIle** is identical to the N terminus of intact *E. coli* dUTPase.[15]

E. coli dUTPase consists of 152 amino acid residues, the last three possible tryptic sites in the sequence are the peptide bonds: Arg 116–Ile 117, Arg 141–Gly 142 and Arg 151–Glu 152, implying losses in molecular mass of 3921, 1071, and 129 Da, respectively, as calculated from the amino acid se-

quence. Experimentally, approximately 1 kDa loss in subunit molecular mass upon trypsinolysis has been estimated from SDS PAGE, which points to cleavage at Arg 141. In order to get independent proof for the location of the tryptic site, electrospray mass spectrometry was also used to measure the molecular mass of dUTPase (Fig. 2). Measured values were 16,296 and 15,223 Da for the intact and the digested samples, respectively, that is, the loss in molecular mass is 1073 Da, in good agreement with the presumed tryptic cut at Arg 141. In the mass spectrum of the digested sample, no peak corresponding to the molecular mass of the intact enzyme could be identified (not shown). There are some shoulders and minor peaks on the mass spectrum of the digested variant, the corresponding molecular mass values were determined to be: 15,087, 15,213, 15,234, 15,298, and 15,536 Da. Efforts to account for these values by assuming further possible tryptic sites, in addition to the Arg 141–Gly 142 peptide bond, failed. Although there are several possible tryptic sites at the N terminus, cleavage at these sites would result in fragments with molecular masses deviating at least 23–120 Da from the values associated with the mass spectrum shoulders, while the maximal error of the instrument was determined to be 9 Da (cf. Fig. 2, control experiment with myoglobin). Moreover, the N terminus was shown to be unharmed upon limited digestion (cf. results from Edman degradation discussed above). The trypic fragments resulting from cleavage of the peptide bonds Arg 116–Ile 117 or Arg 151–Glu 152 at the C terminus would possess molecular masses of 12,366 and 16,158 Da, respectively, far outside of the range of the presented mass spectrum, where no peaks could be identified. Therefore, the mass spectrum shoulders cannot represent tryptic fragments in addition to the 141 residue-digested product, rather may be due to instrumental factors. Taken together, the results of the analytical methods in this section prove that the shortened species lost the C-terminal 11 residues of the intact dUTPase (containing the conserved Motif 5 of dUTPases), with no change at the N terminus.

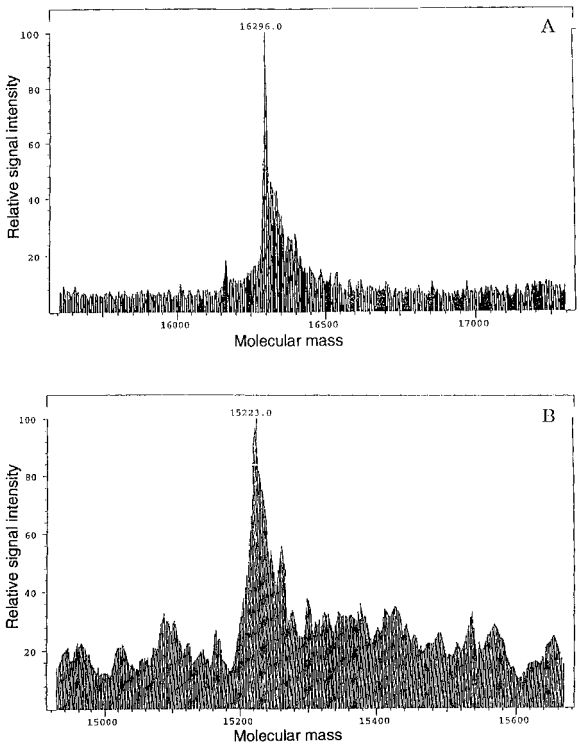

Fig. 2. Mass spectrometric analysis of the limited tryptic digestion of dUTPase. **A:** Intact samples. **B:** Digested samples. Analysis of the data to yield molecular mass values was done by using the software of the instrument. The molecular mass determined for the intact sample is 9 Da more than the value calculated from the amino acid sequence. This deviation was found to be due to the instrument, since it also appeared when myoglobin was measured (not shown).

Characterization of the Motif 5-less Variant

Kinetic analysis and MgdUDP binding

The main panel of Figure 3 shows that loss of Motif 5 seriously perturbs enzymatic function of the protein: its activity showed an exponential decay due to trypsinolysis, in accordance with the single tryptic cut. After 120 minutes of digestion, the residual activity of the Motif 5-less variant is <5% of the activity of the intact enzyme. The presence of MgdUDP had no effect on the rate of tryptic cleavage, that is, the Arg 141–Gly 142 peptide bond is not protected by MgdUDP (Fig. 3). To determine the binding affinity of the digested inactive protein toward MgdUDP, I used the recently developed CD technique,[18] based on a specific near-UV CD signal induced by binding of MgdUDP to the enzyme. As shown in Figure 3, inset, the apparent dissociation constant of the MgdUDP–dUTPase complex is unaffected by trypsinolysis of the enzyme.

Detailed kinetic analysis of the enzyme was performed by measuring the entire progress curve of MgdUTP cleavage (Fig. 4). Assuming Michaelis-Menten behavior,[22] data were treated according to the linearized Michaelis-Menten equation using Equations (1) and (2) as described in the Methods section (Fig. 4, insets). The determined k_{cat} and K_M values were 5.15 s^{-1}, 0.28 μM and 0.14 s^{-1}, 0.49 μM for the intact and digested species, respectively. There is a dramatic 40-fold decrease in k_{cat}, while K_M is increased less than twofold as a result of digestion.

Spectroscopic investigations

Protein conformational changes can often be detected by sensitive spectroscopic techniques. Accordingly, the intact and the Motif 5-less variants of *E. coli* dUTPase have been analyzed by UV absorption, fluorescence, and near- and far-UV CD spectroscopies (Fig. 5). The UV absorption spectrum of the protein shows characteristic double maxima at 278 and 282 nm, together with a shoulder at 295 nm, in accordance with previous results (Fig. 5A: cf. Ref.

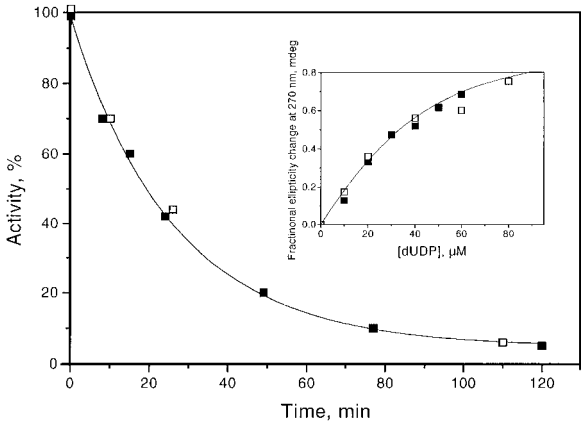

Fig. 3. Inactivation of dUTPase on tryptic digestion. Main Panel. 2 mg/ml (125 μM) *E. coli* dUTPase was digested with 0.2 mg/ml trypsin in the absence (filled symbols) or presence (open symbols) of 200 μM MgdUDP. Data were fitted with the equation $A = A_{final} + A_{initial} \exp(-kt)$ where A = activity at time t, $A_{initial}$ = activity at $t = 0$, A_{final} = activity at $t = \infty$, k = first-order rate constant. Fitting parameters were: $A_{initial}$ = 100%, A_{final} = 4.7%, $k = 0.038\ min^{-1}$. A control experiment showed that in the absence of trypsin, dUTPase retains 100% activity even after 5 hours incubation at 30°C. **Inset.** Binding of MgdUDP to intact (filled symbols) or digested (open symbols) dUTPase. Enzyme at 37 μM was titrated with dUDP in the presence of saturating excess of Mg^{2+}. Differential fractional ellipticity increase at 270 nm, induced by the formation of the MgdUDP–dUTPase complex was followed as previously.[17] Data were fitted by the equation $\Delta\Theta/\Delta[\Theta] = (a - (a/\backslash 2 - 4c_1c_2)/\backslash 0.58/2)$, where $\Delta\Theta$ is the differential ellipticity determined by substracting the ellipticity of the dUDP and dUTPase solution, measured separately, from the ellipticity measured in their mixture, $\Delta[\Theta]$ is the differential molar ellipticity of the MgdUDP–dUTPase complex, $a = c_1 + c_2 + K_{d,app}$, c_1 and c_2 are enzyme and dUDP concentrations, respectively (solid lines). $\Delta[\Theta]$ was 0.068 mdeg/μM for both data sets, $K_{d,app}$ (the apparent dissociation constant) was 15 μM for both datasets.

20). The maximum of the fluorescence emission spectrum is situated at 333 nm (Fig. 5B), as it is usually found for the fluorescence emission of tryptophane residues buried in an apolar proteinaceous microenvironment. This is in accordance with the crystal structure data showing the single Trp 102 of *E. coli* dUTPase in a buried position.[15] Together with fluorescence, near-UV CD spectroscopy has been suggested as one of the most reliable methods to reveal a sensitive fingerprint of tertiary protein structure. There are several characteristic peaks in the near-UV CD of *E. coli* dUTPase, the most intensive peak being situated at 290 nm (Fig. 5C). Far-UV CD spectroscopy is a relevant method to estimate the content of different secondary structural elements in a protein, the spectra recorded are shown in Fig. 5D. The intact and digested samples were found to elicit quite similar spectra in all cases, with the sole exception of the far-UV CD measurements, where the negative peak situated at around 208 nm in the intact species is blue-shifted to 202 nm, and the shape of the spectrum is also somewhat altered upon trypsinolysis. This change was analyzed in the context of secondary structural composition of dUTPase, as estimated from the far-UV CD spectra,[25] and results were compared with the secondary structural composition present in the crystal structure.[15] Table II shows the values estimated from the far UV spectra of the intact and the digested species using different methods. The methods used differ in both their techniques of data treatment (singular value decomposition, least-squares, ridge regression, neural network) and in their databases (sets of different proteins or basis curves for the main secondary conformational states). Advantages and disadvantages of these methods are discussed in detail by Greenfield.[25] *E. coli* dUTPase is a mainly β-sheet protein,[15] consequently, for the purpose of the present investigation all of those main CD evaluation methods were probed which are suggested to report β-pleated sheet conformation fairly well. Although there are considerable differences in the absolute values of α-helical, β sheet, turn and other conformations estimated by the different methods, the changes in these values induced by removal of the C-terminal 11-residue peptide fragment are similar, independently of the specific approach used (Table II). Namely, most methods show that loss of the C terminus induces small, but significant increase in β-pleated sheet content, while the fractional content of other (random) conformation is slightly decreased. The α-helical content is predicted to significantly decrease by the nonconstrained least-squares analysis (MLR), but this change is not confirmed by the other, more reliable methods. When compared to the values calculated from the crystal structure of the intact enzyme, the variable selection (VARSLC[26,27]) and the neural network (K2D[28]) method give closest matches, although the latter does not provide an estimate for the turn conformation. These two methods are also referred to as being between the most useful techniques in the recent review of Greenfield.[25] In conclusion, the spectroscopic evaluation of tertiary protein conformation have not revealed any major conformational change upon trypsinolysis of dUTPase. The Motif 5-less variant is suggested to retain the mostly β-pleated sheet conformation of the intact protein, with a minor increase in the fraction of ordered conformation.

Oligomeric Status of dUTPase

Oligomeric status of dUTPase was determined by analytical gel filtration (Fig. 6). Both the intact and the digested forms eluted at the same position corresponding to the trimeric protein (around 46 kDa), that is, tryptic cleavage did not perturb subunit interactions.

DISCUSSION

Recently, a high-resolution three-dimensional structure became available for the trimeric dUTPase from *E. coli*.[15,19] This enzyme belongs to the same dUTPase subgroup, termed here HuBaR dUTPases,

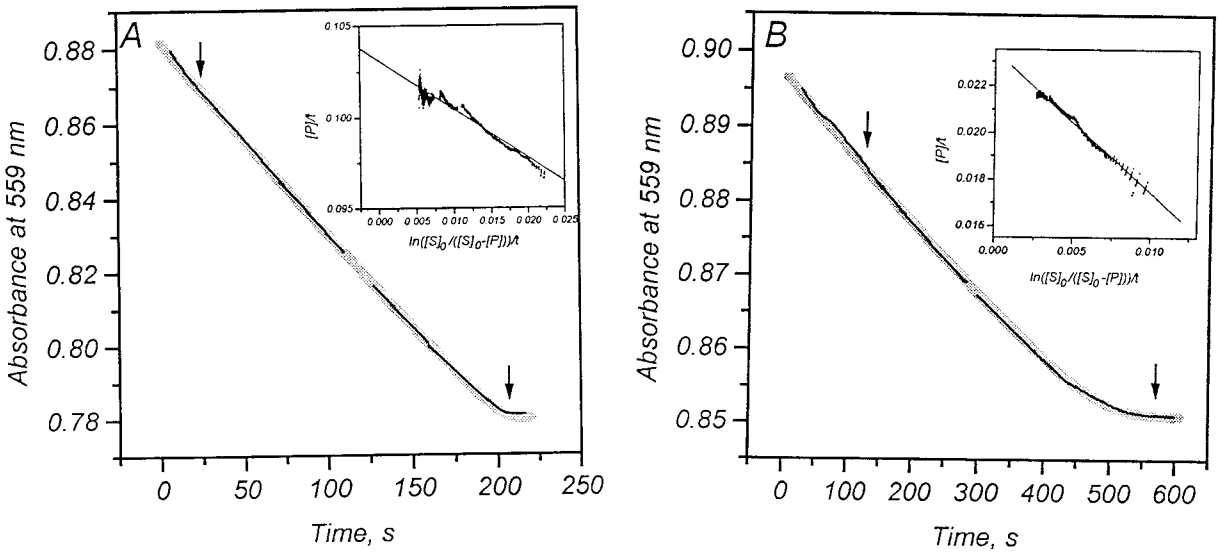

Fig. 4. Determination of enzyme kinetic parameters of intact (**A**) and digested (**B**) dUTPase. Reaction was started by the addition of 20 nM intact or 100 nM digested enzyme to the assay solution containing 20 μM (A) or 10 μM (B) dUTP. Insets show the data between the arrows transformed according to Equations (1) and (2) as described in the Materials and Methods section. Measured values are presented by scattered graphs, gray shading on main panels, or solid lines on insets display the results of least-squares fitting. V_{max} ($=k_{cat}$ [dUTPase]) and K_M determined from the Inset fittings were used to develop the theoretical progress curves (gray) on the main panels. The errors of determinations were 2% for V_{max} and 3% for K_M values, respectively.

as the human and several retroviral enzymes. The oligomeric structure of several other species in the HuBaR group is also known to be trimeric (e.g., human, rat, equine infectious anemia virus), and the homology between these sequences is about 35%.[13,14,16] These observations might allow a similar fold for the proteins in the HuBaR subgroup, resulting in an exposed C-terminal region of the corresponding polypeptides. To test this hypothesis, the tryptic deletion method presented here seems adequate, since all dUTPases contain the strictly conserved Arg, as a possible site for tryptic cleavage, at the beginning of Motif 5. There are numerous basic residues in *E. coli* dUTPase, but only two of them (Arg 141 and Arg 151) are found among the C-terminal 16 residues, suggested to behave rather flexible in the otherwise quite compact jellyroll fold.[15] Moreover, the last peptide bond Arg 151–Gln 152 is not expected to be a good substrate for the endopeptidase trypsin, suggesting that a selective tryptic cleavage may occur at the Arg 141–Gly 142 peptide bond. The present work confirmed this hypothesis and enabled me to construct a Motif 5-less variant of *E. coli* dUTPase. The enzymatic function of this variant, catalyzing the cleavage of the α–β phosphate bond of dUTP, is seriously harmed: k_{cat} is decreased to 2.4% of the intact species (Figs. 3 and 4). There is also a small, less than twofold increase in K_M. Assuming rapid equilibrium kinetics and equating K_M with the apparent dissociation constant of the MgdUTP–dUTPase complex (cf. Ref. 22), the contribution of the C-terminal 11-residue fragment to the substrate-binding free energy is calculated to be a mere 4%. Binding affinity for the substrate analogue MgdUDP is completely retained in the Motif 5-less protein. dUDP also possesses the α–β phosphate bond, nevertheless, it is a nonhydrolyzable inhibitor of the enzyme.[22] Based on the present results it seems that Motif 5, although apparently does not participate in accommodating the scissile bond, still has some vital role either in positioning dUTP for productive enzyme–substrate complex formation or in enzymatic catalysis.

The crystal structure of an N-terminal shortened, but catalytically active, genetically constructed variant of human dUTPase was published after submission of the present investigation for publication.[29] This structure confirms the suggested similar fold of the *E. coli* and the human protein and the proposed flexible behavior of Motif 5 renders the C terminus invisible in the case of the ligand-free human enzyme, as well.[29] However, in the dUDP-human dUTPase complex, in contrast to the dUDP–*E. coli* dUTPase complex,[19] the C terminus becomes partially visible (the last five residues are still delocalized). This structure reveals that the conserved Phe of Motif 5 (the residue at position 10 in Fig. 7) stacks over the uracil ring of the substrate analogue dUDP bound to the human enzyme. A similar movement of the C terminus induced by dUTP binding to *E. coli* dUTPase, but absent upon dUDP binding to the bacterial enzyme,[19] could initiate the catalytic reaction by fostering correct positioning of key amino acid residues. There is a functional difference be-

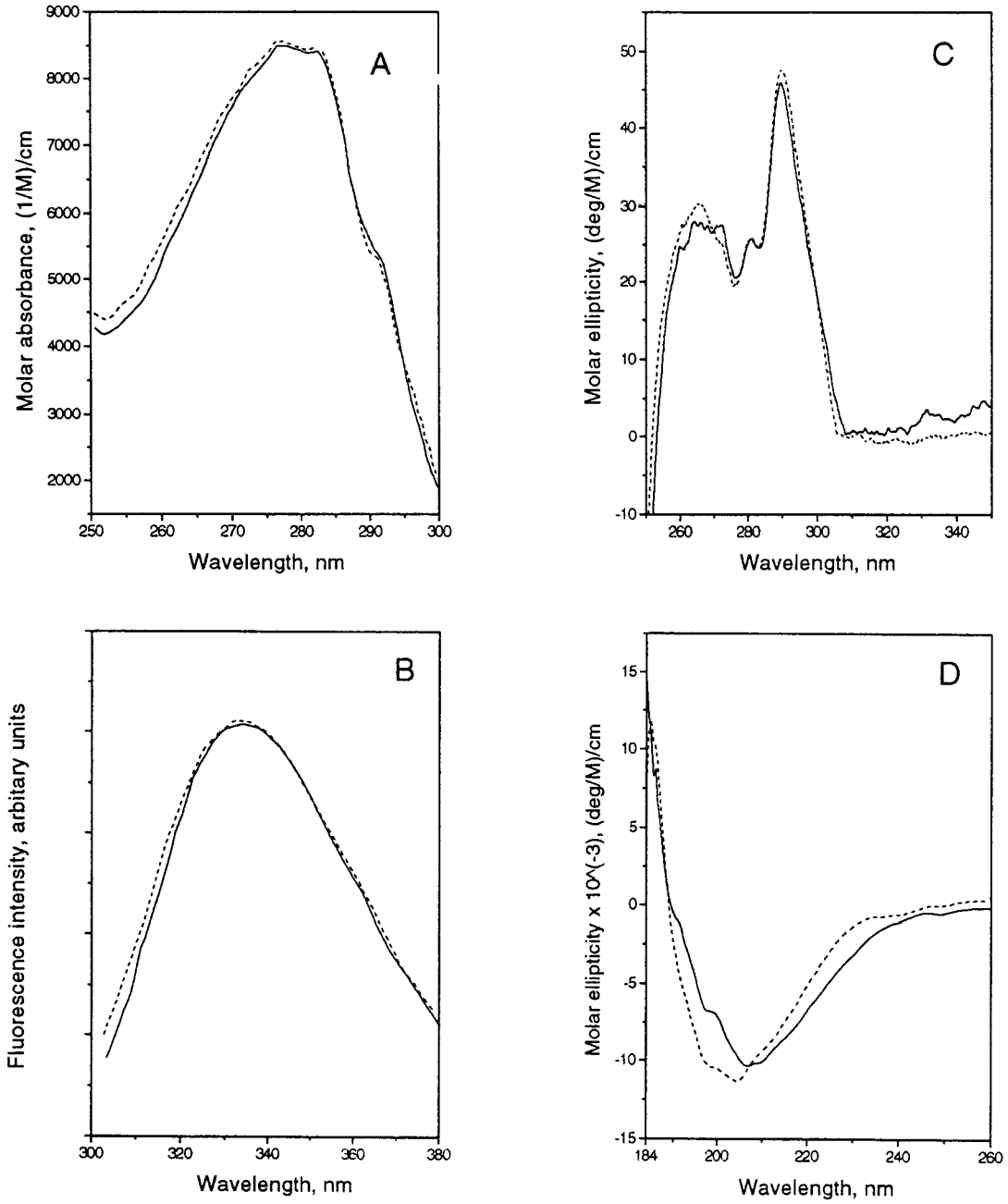

Fig. 5. Investigation of protein conformation of intact (solid line) and digested (dashed line) dUTPase by spectroscopic techniques. **A:** UV absorption. **B:** Fluorescence emission. **C:** Near-UV circular dichroism. **D:** Far-UV circular dichroism.

tween the bacterial and the human enzyme: Motif 5 does not participate in dUDP (crystal structure data, Ref. 19) or MgdUDP (present study, Fig. 3) binding, whereas it moves to lock the substrate analogue by phenylalanine-uracil stacking in the human structure.[29] This difference may suggest that the closing of the active site pocket requires the presence of the complete α–β–γ phosphate chain of the substrate dUTP in the bacterial enzyme, and upon cleavage of the scissile α–β phosphate bond, the active site pops open (the C-terminal tail returns to the unordered conformation) to release the products dUMP and PP_i. The mechanism of product release must be different in the human enzyme, since the phenylalanine-uracil stacking (i.e., the closure of the active site) is still present in the dUMP-dUTPase complex.[29] This difference may be exploited in the design of species-specific inhibitors of the enzyme.

Table II. Estimation of Content of the Different Secondary Structural Conformations in Intact and Digested dUTPase*

	Secondary structural elements							
	α Helix		β Sheet		Turn		Other	
Method	Intact	Digested	Intact	Digested	Intact	Digested	Intact	Digested
VARSLC†	8 ± 1	8 ± 1	32 ± 2	34 ± 2	17 ± 1	16 ± 1	38 ± 1	37 ± 2
CONTIN‡	0 ± 0.1	0 ± 0.1	58 ± 1	61 ± 1	34 ± 1	34 ± 1	8 ± 1	5 ± 1
G&F§	0	0	39	39	ND	ND	68	69
MLR¶	5	2	47	55	38	38	10	5
K2D**	11	9	42	47	ND	ND	47	44
X-ray††	5	ND	41	ND	11	ND	43	ND

*Far-UV CD spectra presented in Figure 4D were processed by the main methods reported to be reasonably adequate for proteins containing mainly β sheets.[25] Analyses were conducted according to the authors' instructions, using CD data collected between 184 and 260 nm in the case of the VARSLC program or 200–240 nm for all the other evaluations. VARSLC, singular value decomposition with variable selection[26,27]; CONTIN, ridge regression analysis of Provencher and Glöckner[34]; G&F, constrained least-squares analysis of Greenfield and Fasman[35]; MLR, nonconstrained least-squares analysis[25]; K2D, neural network method of Andrade et al.[28]
†Database used was the set of 33 proteins as in Ref. 217.
‡Database used was the set of 16 proteins originally suggested by Provencher and Glöckner.[34]
§Database consisted of the spectra of poly-L-lysine in the α-helical, β-pleated sheet and random conformations.[35]
¶Database contained the spectra for α helix, β sheet, β turn, and random coil extracted from 15 proteins by multilinear regression.[36]
**Database used was that of Andrade et al.[28]
††The conformation of intact dUTPase was calculated from the atomic coordinates,[15] using the method of Kabsch and Sanders.[24]

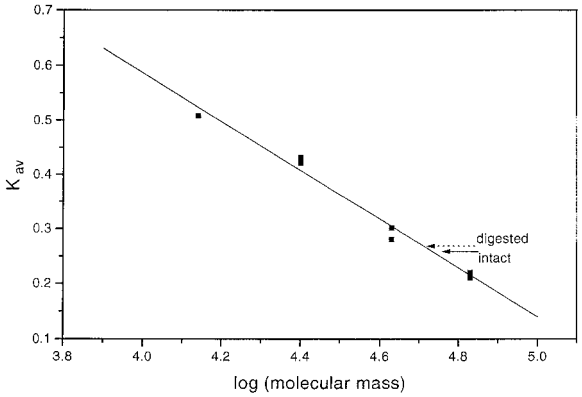

Fig. 6. Analysis of the oligomeric structure of intact and digested dUTPase by gel filtration. Calibrating proteins (full squares), intact dUTPase (solid arrow), or digested dUTPase (dotted arrow) were applied separately. Data on the calibration curve were fitted to a straight line: $K_{av} = (V_e - V_O)/(V_t - V_0) = 2.37\ (\pm 0.13) - 0.47\ (\pm 0.03)$ log (molecular mass), where V_e = elution volume, V_O = exclusion volume determined by Blue Dextran = 70 ml, V_t = total volume of the column = 162 ml. Values are given as mean (± error). K_{av} values for intact or digested dUTPase were 0.28 (±0.025, two independent chromatographic runs) and 0.29 (±0.023, two independent chromatographic runs). Consequently, the estimated molecular masses are 47 and 45 kDa, for the intact and the digested proteins, respectively, the difference between these two values being within the error of the determination.

The Motif 5-less variant of the enzyme seems to retain the secondary, tertiary, and quaternary conformation of the intact species (Figs. 5 and 6). The secondary structural composition calculated from the crystal structure is in reasonable agreement with the data estimated from the far-UV CD spectrum of the intact enzyme (Table II). Both datasets confirm high occurrence of β elements, together with a low α-helical content. Although there are differences between the calculated and the estimated values, these may be explained by the usual finding that secondary structural analysis of CD spectra provides best estimates for α-helical content.[25] The change in the shape of the far-UV CD spectrum on digestion can be interpreted by the aid of various evaluation techniques as a small, but significant, increase in the fractional content of ordered (β-pleated sheet) structural elements. This finding is in accordance with the expectation. In the crystal structure the C-terminal 16 residues were found not to possess any stable secondary structure.[15,19] Removal of the 11-residue fragment of this unordered region may be expected to increase the fractional content of the ordered structures in the remaining larger fragment. Nevertheless, since the cutoff region amounts to only 7% of the intact polypeptide chain, no major effect would be reasonable.

It was suggested that Motif 5 of dUTPases is somewhat similar to the Walker A phosphate binding P-loop motif[30] found in mononucleotide binding proteins.[19] This similarity is seriously harmed for two main reasons.

1. As shown in Figure 7, Motif 5 lacks the strictly conserved Lys between Gly and Ser/Thr of the Walker A motif, where Lys was shown to accommodate the β and/or γ phosphates of the respective substrates in several crystal structures and it is proposed to stabilize the transition state during catalysis (reviewed in Ref. 31).
2. Walker A motifs usually constitute a loop connecting a β strand to a following α helix nearby the N terminus of the corresponding polypeptides (32), while Motif 5 of dUTPases is situated at the very

Motif 5 sequences, Pos. #		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Non-viral sources	Candida	**D**	N	**T**	E	R	G	E	G	G	F	G	**S**	**T**	G		
	E. coli	**D**	A	**T**	D	R	G	E	G	G	F	G	H	S	G		
	Yeast	E	E	S	A	R	G	A	G	G	F	G	**S**	**T**	G		
	Human	**D**	D	**T**	E	R	G	S	G	G	F	G	**S**	**T**	G		
	Tomato	**D**	S	**T**	V	R	G	S	G	G	F	G	**S**	**T**	G		
Herpes viruses	HSV-1	P	P	S	E	R	**G**	T	G	G	**F**	G	S	**T**	**G**		
	BHV	P	P	S	A	R	**G**	P	R	G	**F**	G	S	**T**	**G**		
	VZV	T	P	S	E	R	**G**	T	R	G	**F**	G	S	**T**	D		
	EBV	A	L	E	G	R	Q	G	R	G	**F**	G	S	S	**G**		
	CCV	E	R	T	R	R	**G**	T	G	G	**F**	G	S	S	**G**		
	HSVSA	S	Q	M	S	R	**G**	D	A	G	L	G	S	S	**G**		
	EHV	P	S	S	L	R	A	D	G	G	**F**	G	S	T	**G**		
Pox viruses	Variola			T	D	R	G	**D**	Q	G	F	G	S	T	G		
	Vaccinia			T	N	R	G	**D**	Q	G	F	G	S	T	G		
	SFV			T	E	R	G	N	S	G	F	G	S	S	G		
	ORF			T	D	R	G	**D**	S	G	F	G	S	T	G		
ASFV				T	P	R	G	E	G	R	F	G	S	T	G		
Lenti retro-viruses	Ovine			T	E	R	**G**	N	Q	G	**F**	G	S	**T**	**G**		
	Visna			T	E	R	**G**	E	Q	G	Y	G	S	**T**	**G**		
	JSRV			I	N	R	Q	D	K	G	**F**	G	S	S	D		
	FIV			S	E	R	**G**	D	N	G	Y	G	S	**T**	**G**		
	CAEV			T	E	R	**G**	E	K	G	**F**	G	S	**T**	**G**		
	EIAV			S	Q	R	**G**	D	K	G	**F**	G	S	**T**	**G**		
Onco retro-viruses	SRV					R	G	Q	G	S	**F**	G	S	**S**	**D**	I	**Y**
	MPMV					R	G	Q	G	S	**F**	G	S	**S**	**D**	I	**Y**
	SMRV					R	G	A	S	A	P	G	S	**S**	**D**	V	**Y**
	MMTV					R	G	S	E	G	**F**	G	S	T	S	H	V

Motif 5 consensi, Pos. #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Non-viral	**D** E	x	**T** S	x	R	G	x	G	G	F	G	**S** H	**T** S	G		
Herpes					R	G	x	x	G	**F** L	G	S	**T** S	**G** D		
Pox			T	x	R	G	**D** N	S Q	G	F	G	S	**T** S	G		
Lenti					R	**G** Q	x	x	G	**F** Y	G	S	**T** S	**G** D		
Onco					R	**G** A	x	x	x	**F** P	G	S	**S** T	**D** S	x	**Y** V
Motif 5 Common					R	G	x	x	G	F	G	**S** H	**T** S	G		
Other P-loops																
Walker A ref.(30)						G	x	x	x	x	G	K	S T			
Walker B[a] ref.(30)					R K	x	x	x	G	x	x	h	x	h	D E	
Dinucl.[b] ref.(31)							h	h	G	x	G	x	x	G		
PP_i loop[c] ref.(33)			h	h	h	h	h	S	G	G	x	D	S T			
Protein kinase ref.(31)					h	G	x	G	x	F Y	G	x	V			

Fig. 7.

C terminus of the polypeptide chain in both the HuBaR and the herpeslike subgroup. Moreover, at least in the case of the *E. coli* enzyme, Motif 5 does not form any turn (loop) structure being flexibly unlocalized in the crystal structure. Figure 7 lists several other Gly-rich P-loop motifs, which again usually constitute real loop structures connecting β strand to α helix (Walker A loop, dinucleotide binding P loop, pyrophosphatase P loop) or β strand to another β strand (protein kinase P loop),[31–33] that is, they differ from Motif 5 in this respect. On the other hand as far as the sequence data are concerned, these motifs all agree with the dUTPase Motif 5 in having at least two (sometimes three) common residues situated at the same position. It is worthwhile to note that three conserved residues of Motif 5, not present in the Walker A loop, the invariable Arg at the beginning (5), the Phe position (10) in the middle and the Gly position (14) at the end are also found in some of the listed P loops, in the Walker B motif, the protein kinase loop, and the dinucleotide P loop, respectively. These considerations prompt toward two conclusions i) Motif 5 of dUTPases shows restricted homology to several different P loop motifs while ii) the dUTPase Gly-rich motif is definitely distinct from any of these P loops. The role of this motif is presently established as relevant for enzymatic function for the *E. coli* dUTPase, and it seems reasonable to argue for a similar importance of Motif 5 in other dUTPases.[29]

Fig. 7. Comparison of Motif 5 of dUTPases with P-loop motifs found in other proteins. Amino acid sequences were taken from the sequence databases SwissProt and GenBank. Alignment of Motif 5 sequences was performed according to Ref. 10. One-letter code is used for amino acid residues, x stands for any amino acid, h stands for a hydrophobic residue. In positions where only two residues may occur, boldface indicates higher incidence, where applicable. Underlined residues are invariable. Motif 5 consensuses are based on at least 70% occurrence. In case of several available clones, only one has been used to avoid almost identical sequences. Pos. #, position number; Candida, *Candida albicans* GenBank X77925; *E. coli,* SwissProt P06968; Yeast, *Saccharomyces cerevisiae,* SwissProt P33317; Human, *Homo sapiens* SwissProt P33316; Tomato, *Lycopersicon esculentum* SwissProt P32518; HSV-1, herpes simplex virus type 1 SwissProt P10234; BHV, bovine herpes virus; VZV, varicella zoster virus SwissProt P09254; EBV, Epstein-Barr virus SwissProt P03195; CCV, channel catfish virus SwissProt P28893; HSVSA, herpes virus saimiri SwissProt Q01034; EHV, equine herpesvirus type 1 (Strain AB4P) SwissProt P28892; Variola, variola virus SwissProt P33826; Vaccinia, vaccinia virus (strain WR) SwissProt P17374; SFV, Shope fibroma virus SwissProt P32308; ORF, orf virus SwissProt P14597; African swine fever virus GenBank X71982; Ovine, ovine lentivirus SwissProt P16901; Visna, visna lentivirus (strain KV1772) SwissProt P35956; JSRV, Jaagsiekte sheep retrovirus SwissProt P31625; FIV, feline immunodeficiency virus (Isolate San Diego) SwissProt P19028; CAEV, caprine encephalitis virus; EIAV, equine infectious anemia virus; SRV, simian type D retrovirus; MPMV, Mason-Pfizer monkey virus; SMRV, squirrel monkey retrovirus; MMTV, mouse mammary tumor virus. [a]The more general consensus of the Walker B motif is (R, K) $-x_{1-4}Gx_{2-4}$hxh(D, E) (ref. (31)). [b]Dinucl. = P-loop found in dinucleotide binding proteins usually possessing the Rossmann fold.[31] [c]PP_i loop = P-loop found in ATP pyrophosphatases.[33]

CONCLUSION

Our long-term aim is to reveal detailed structure–function relationships of dUTPase, ultimately contributing to the rational design of dUTPase inhibitors for use in pharmacological trials. The vital importance of dUTPase in DNA metabolism is recently being appreciated as high resolution details of the structure/function relationships of the enzyme become published.[5,15,19,29] Hereby I presented a simple tryptic method to construct a Motif 5-less variant of *E. coli* dUTPase, perturbed in catalytic function but not in higher order protein structure. This variant still contains the strictly conserved Arg 141 residue, immediately preceding the Gly-rich P-looplike motif. The shortened variant is expected to be helpful in determining the contribution of the Gly-rich motif to the interaction of the enzyme with different inhibitors (investigations are in progress in our laboratory). The functional role of Motif 5 for *E. coli* dUTPase, invisible flexible in the crystal structure, is presently established.

ACKNOWLEDGMENTS

I gratefully thank Rebecca Persson and Prof. Per Olof Nyman, University of Lund, Sweden, for making *E. coli* dUTPase available to me. Thanks are due to Dr. Gunilla Larsson for providing the atomic coordinates of the *E. coli* dUTPase apoenzyme; Profs. N. Greenfield and the original authors of CD evaluating computer programs (Refs. 26–28, 34–37), Profs. W. Kabsch and C. Sander for software distribution; Dr. András Patthy, Analysis and Synthesis Laboratory, Agricultural Biotechnological Research Center of Gödöllö, Hungary, for amino acid analysis and N-terminal sequencing; and Zoltán Kele, University of József Attila, Szeged, Hungary, for mass spectrometry.

REFERENCES

1. Kornberg, A., Barker, T.A. "DNA Replication," 2nd ed. New York: Freeman, 1991.
2. Gadsden, M.H., McIntosh, E.M., Game, J.C., Wilson, P.J., Haynes, R.H. dUTP pyrophosphatase is an essential enzyme in *Saccharomyces cerevisiae.* EMBO J. 12:4425–4431, 1993.
3. El-Hajj, H.H., Zhang, H., Weiss, B. Lethality of a *dut* (deoxyuridine triphosphatase) mutation in *Escherichia coli.* J. Bacteriol. 170:1069–1075, 1988.
4. Ingraham, H.A., Dickey, L., Goulian, M. DNA fragmentation and cytotoxicity from increased cellular deoxyuridylate. Biochemistry 25:3235–3230, 1986.
5. Pearl, L.H., Savva, R. The problem with pyrimidines. Nature Struct. Biol. 3:485–487, 1996.
6. Canman, C.E., Radany, E.H., Parsels, L.A., Davis, M.A., Lawrence, T.S., Maybaum, J. Induction of resistance to fluorodeoxyuridine cytotoxicity and DNA damage in human tumor cells by expression of *Escherichia coli* deoxyuridine triphosphatase. Cancer Res. 54:2296–2298, 1994.
7. Zalud, P., Wachs, W.O., Nyman, P.O., Zeppezauer, M. In "Purine and Pyrimidine Metabolism in Man," Vol. VIII:

Sahota, A., Taylor, M. (eds.). New York: Plenum, pp. 135–138.
8. Nation, M.D., Guzder, S.N., Giroir, L.E., Deutsch, W.A. Control of *Drosophila* deoxyuridine triphosphatase. Biochem. J. 259:593–596, 1989.
9. Strahler, J.R., Zhu, X.-X., Hora, N., Wang, Y.K., Andrews, P.C., Roseman, N.A., Neel, J.V., Turka, L., Hanash, S.M. Maturation stage and proliferation-dependent expression of dUTPase in human T cells. Proc. Natl. Acad. Sci. U.S.A. 90:4991–4995, 1993.
10. McGeoch, D.J. Protein sequence comparisons show that the "pseudoprotease" encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase enzyme family. Nucleic Acids Res. 18:4105–4110, 1990.
11. Pyles, R.B., Sawtell, N.M., Thompson, R.L. Herpes simplex virus type 1 dUTPase mutants are attenuated for neurovirulence, neuroinvasiveness and reactivation from latency. J. Virol. 66:6706–6713, 1992.
12. Threadgill, D.S., Steagall, W.K., Flaherty, M.T., Fuller, F.J., Perry, S.T., Rushlow, K.E., LeGrice, S.F.J., Payne, S.L. Characterization of equine infectious anemia virus dUTPase: Growth properties of a dUTPase deficient mutant. J. Virol. 67:2592–2600, 1993.
13. Climie, S., Lutz, T., Radul, J., Sumner-Smith, M., Vanderberg, E., McIntosh, E. Expression of trimeric human dUTP pyrophosphatase in *Escherichia coli* and purification of the enzyme. Protein Expr. Purif. 5:252–258, 1994.
14. Hokari, S., Sakagishi, Y. Purification and characterization of deoxyuridine triphosphate nucleotidohydrolase from anemic rat spleen: a trimer composition of the enzyme protein. Arch. Biochem. Biophys. 253:350–356, 1987.
15. Cedergren-Zeppezauer, E.S., Larsson, G., Nyman, P.O. Dauter, Z., Wilson, K.S. Crystal structure of a dUTPase. Nature 355:740–742, 1992.
16. Bergman, A.-C., Björnberg, O., Nord, J., Rosengren, A.M., Nyman, P.O. dUTPase from retrovirus equine infectious anemia virus: High-level expression in *Escherichia coli* and purification. Protein Expr. Purif. 6:379–387, 1995.
17. Vertessy, B.G., Zalud, P., Nyman, P.O., Zeppezauer, M. Identification of tyrosine as a functional residue in the active site of *Escherichia coli* dUTPase. Biochim. Biophys. Acta 1205:146–150, 1994.
18. Vertessy, B.G., Persson, R., Rosengren, A.M., Zeppezauer, M., Nyman, P.O. Specific derivatization of the active site tyrosine in dUTPase perturbs ligand binding to the active site. Biochem. Biophys. Res. Commun. 219:294–300, 1996.
19. Larsson, G., Svensson, L.A., Nyman, P.O. Crystal structure of the *Escerichia coli* dUTPase in complex with a substrate analogue (dUDP). Nature Struct. Biol. 3:532–538, 1996.
20. Hoffmann, I., Widström, J., Zeppezauer, M., Nyman, P.O. Overproduction and large-scale preparation of deoxyuridine triphosphate nucleotidohydrolase from *Escherichi coli*. Eur. J. Biochem. 164:45–51, 1987.
21. Keleti, T. Basic enzyme kinetics. Akadémiai Kiadó, Budapest, 1986.
22. Larsson, G., Nyman, P.O., Kvassman, J. Kinetic characterization of dUTPase from *Escherichia coli*. J. Biol. Chem. 271:24010–24016, 1996.
23. Cleveland, D.W., Fischer, S.G., Kirschner, M.W., Laemmli, U.K. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. 252:1102–1106, 1977.
24. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
25. Greenfield, N. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. Anal. Biochem. 235:1–10, 1996.
26. Compton, L.A., Johnson, W.C. Jr. Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. Anal. Biochem. 155:155–167, 1986.
27. Manavalan, P., Johnson, W.C. Jr. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. Anal. Biochem. 167:76–85, 1987.
28. Andrade, M.A., Chacón, P., Merolo, J.J., Morán, F. Protein Eng. 6:383–390, 1993.
29. Mol, C.D., Harris, J.M., McIntosh, E.M., Tainer, J.A. Human dUTP pyrophosphatase: Uracil recognition by a β hairpin and active sites formed by three separate subunits. Structure 4:1077–1092, 1996.
30. Walker, J.E., Saraste, M., Runswick, M.J., Gay, N.J. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1:945–951, 1982.
31. Bossemeyer, D. The glycine-rich sequence of protein kinases: A multifunctional element. Trends Biol. Sci. 19:201–205, 1994.
32. Traut, T.W. The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. Eur. J. Biochem. 222:9–19, 1994.
33. Bork, P., Koonin, E.V. A P-looplike motif in a widespread ATP pyrophosphatase domain: Implications for the evolution of sequence motifs and enzyme activity. Proteins 20:347–355, 1994.
34. Provencher, S.W., Glöckner, J. Estimation of protein secondary structure from circular dichroism. Biochemistry 20:33–37, 1981.
35. Greenfield, N., Fasman, G.D. Biochemistry 8:4108–4116, 1969.
36. Yang, J.T., Wu, C.S., Martinez, H.M. Calculation of protein conformation from circular dichroism. Methods Enzymol. 130:208–269, 1974.