Solid phase synthesis of TyrT, a thymine–tyrosine conjugate with poly(A) RNA-binding ability

Giovanni N. Roviello,*a Valentina Roviello,b Ida Autieroa and Michele Savianoc

The present work deals with the synthesis and characterization of a novel nucleoamino acid derivative based on a L-tyrosine moiety to which a thymine nucleobase was anchored by means of an amide bond to the N-alpha group. This derivative, denominatored by us TyrT, belongs to the family of the nucleobase–amino acid conjugates that show a wide range of biological activities, frequently associated with their ability to interact with nucleic acids. In this respect, the interaction of TyrT with poly(A), a proposed RNA target for anticancer strategies, was studied by circular dichroism (CD) which suggested its ability to bind this RNA. Moreover, the modification of the morphology of a sample of TyrT in the presence of poly(A) was visualised by scanning electron microscopy (SEM) which was in agreement with the evidence that the thyminyl L-tyrosine interacts with poly(A). Finally, computational analyses have been performed to hypothesize the binding mode from a structural point of view, suggesting that the binding is mainly kept via hydrophobic contacts, reproducing a stacking-like interaction between the thymine ring of TyrT and the two successive adenine rings of a poly(A) model.

Introduction

Several examples of biologically-active compounds based on amino acid units carrying nucleobases are known in the literature, as in the case of the willardiine, an uracil-carrying L-alanine derivative which acts as a partial agonist of the AMPA receptor, involved in several neurological disorders.1 Another example of a nucleobase–amino acid conjugate is the aminoacyl nucleoside N6-threonylcarbamoyl adenosine, found in transfer RNAs, involved in the protein synthesis process in which it has a fundamental importance in maintaining the decoding accuracy.2 Besides the natural nucleoamino acids and aminoacyl nucleosides, also artificial nucleobase–amino acid conjugates were described, some of which were found able to bind the HIV trans-activation response (TAR) RNA element.3

L-Tyrosine plays a key role in the protein–nucleic acid binding, as in the case of the DNA recognition by the herpes simplex virus DNA binding proteins,4 a process requiring the interaction of the nucleic acid with L-tyrosine residues present on the protein surface. Interestingly, L-tyrosine residues are also involved in RNA-binding processes as in case of (i) free L-tyrosine/RNA aptamers binding,5 or of (ii) the residue Tyr236 which is required for RNA-binding activity of the subunit of the Cleavage Factor CFI125 in Entamoeba histolytica.6 Tyrine–adenine–tyrosine aromatic stacking pairing was found of crucial importance in the N-glycosidase activity of gelonin, a ribosome inactivating protein from the plant Gelonium multiflorum, suggesting the key role of this interaction in the RNA binding by this protein.7 Besides the aromatic non-covalent nucleobase–tyrosine interactions, the formation of covalent thymine–tyrosine adducts was demonstrated suggesting that this kind of derivatives are at the origin of DNA–protein cross-linking occurring under DNA exposure to ionizing radiation.8

Poly(A) tail present at 3’ terminus of eukaryotic mRNA was recognized as a potential target for anticancer biomedical approaches9 and thus, an increasing attention is currently devoted to various molecular systems (including small organic molecules, modified peptides etc.) able to recognize poly(A) in view of their employment as possible anticancer drugs.10–14

Starting from the information that L-tyrosine residues in proteins can drive the binding to nucleic acids and also that free L-tyrosine can interact with RNA, we aimed to investigate the properties of a L-tyrosine chemically modified by a thymine nucleobase with respect to its RNA binding. This adduct (Fig. 1) is associable to the nucleobase–amino acid conjugates that were found able to bind TAR RNA by Joly et al.,3 with the major difference consisting in the different nature of the amino acid employed (basic versus aromatic of the present study).

Results and discussion

Synthesis of TyrT

The nucleoamino acid derivative was assembled on a Rink Amide MBHA resin (0.65 mmol g⁻¹, 77 mg, 50 μmol), which was

---

*Istituto di Biosstrutture e Bioimmagini – CNR (UOS Napoli Centro), 80134 Napoli, Italy. E-mail: giovanni.roviello@cnr.it; Fax: +39-81-2534574; Tel: +39-81-2534585

1 Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale (DICMaPI), Università di Napoli “Federico II”, 80125 Napoli, Italy

2 Istituto Di Cristallolografia – CNR, 70126 Bari, Italy
previously Fmoc-deprotected by treatment with 40% piperidine in DMF (30 min), following the synthetic route of Fig. 2. More in detail, a solution of Fmoc-L-Tyr(Bu)-OH (250 μmol, 5 equiv.), HATU (250 μmol, 5 equiv.) and DIEA (500 μmol, 10 equiv.) in 0.8 mL of DMF was added to the resin-NH₂ and the mixture was stirred over 30 min. This procedure was repeated two times and, subsequently, Fmoc group was removed with 40% piperidine in DMF (20 min). The tyrosine-functionalised resin was reacted, then, with TCH2COOH (250 μmol, 5 equiv.), HATU (250 μmol, 5 equiv.) and DIEA (500 μmol, 10 equiv.) dissolved in 0.8 mL of DMF. After 30 min the same procedure was repeated once more. Finally, TyrT was cleaved from the solid support by treatment with TFA/H₂O/TIS (95/2.5/2.5, v/v) over 2 h and, a solvent precipitation from cold diethyl ether, centrifugation and lyophilisation.

The sample obtained by the above-described procedure was analyzed by HPLC, NMR and ESIMS that confirmed the identity of the desired product which was recovered pure with an overall yield of 79%; LC-ESIMS (Fig. 3) \( m/z = 346.76 \) (found), 347.36 (expected for \( [C_{16}H_{18}N_4O_5 + H]^+ \)); 693.31 (found), 693.71 (expected for \( 2 \times [C_{16}H_{18}N_4O_3 + H]^+ \)); NMR \( \delta_H \) (400 MHz, CD,OD) 7.18 (1H, s, Cα-H), 7.06 (2H, d, J = 8.4, CH Ar-Phe), 6.94 (2H, d, J = 5.6, 8.8, CHα), 4.39 (1H, d, J = 16.4, NCH3H), 4.31 (1H, d, J = 16.4, NCH3H), 3.09 (1H, dd, J = 5.2, 14, CHβH), 2.80 (1H, dd, J = 9.2, 14, CHβH), 1.84 (3H, s, CH3); \( \delta_C \) (100 MHz, CD,OD) 176.3, 169.7, 167.2, 153.4, 147.7, 143.7, 131.6, 129.3, 116.5, 111.5, 56.5, 51.2, 38.4, 12.5.

Prior to its utilisation in the subsequent experiments, TyrT solid was added with H2O/ethanol = 1/1. Interestingly, following the synthetic procedure of the present work TyrT was obtained as amide-derivative (-CONH₂) and it should be regarded as a nucleoamino amide. In this regard, it is worth to outline the importance of amino amides, molecules of remarkable biomedical significance as seen in case of numerous anti-inflammatory, antibacterial, analgesic and anaesthetic drugs.¹⁵

**CD poly(A) RNA-binding studies**

First the CD profile of TyrT was analysed at pH 7.5 and 10 °C. Fig. 4 shows the CD spectrum of TyrT ( ), which presents two positive bands (at 199 and 225 nm) and no other relevant signal around 260 nm where DNA ( ) and RNA ( ) have their significant spectroscopical contribution.

CD binding experiments with TyrT and DNA (dA12) in 10 mM phosphate buffer solution at pH 7.5 an 10 °C did not show any difference in spectra of DNA and TyrT/DNA (T/A = 2/1) mixture, suggesting that the thymine derivative is not able to bind the DNA molecule. On the other hand, the same study performed with poly(A) RNA led to a certain variation of the CD spectrum upon complexation with TyrT. This finding can be explained in terms of a molecular interaction between the nucleobase–tyrosine conjugate and poly(A), which causes perturbations of the structure of poly(A) RNA (Fig. 5).

**Microscopy analysis**

In order to perform SEM observations, TyrT was crystallized from a H2O/ethanol (70/30, v/v) solution by slow evaporation of a drop of solution deposited on the SEM sample holder. The TyrT/poly(A) complex obtained after CD studies was also analyzed, whereas a second sample of TyrT/poly(A) complex was realized by adding TyrT to poly(A) in the proper stoichiometric ratio directly on the sample holder. Fig. 6 shows two SEM images of pure TyrT at different magnifications (top left and right), TyrT/poly(A) obtained after CD experiments (bottom left) and freshly-prepared on the SEM sample holder (bottom right).

SEM shows pure TyrT as tufts of lamellar needle-like crystals (Fig. 6, top left): these are well formed, with lamellae being characterized by a very regular surface clearly observable at a magnification of 24 000× (Fig. 6, top right). These findings suggest a high level of purity of TyrT and indicate strong...
**Fig. 3** LC-ESIMS (positive ions) of TyrT (RP C18 column, HPLC method: 2 to 30% of B in A over 10 minutes).

**Fig. 4** CD profile relative to a 10 μM solution of TyrT (red) in comparison with the spectra of DNA (blue) and RNA (green) both at a 20 μM concentration in 10 mM phosphate buffer, pH 7.5, at 10 °C.

**Fig. 5** Sum (green) and complex (blue) CD spectra relative to TyrT (8 nmol in T) + poly(A) (16 nmol in A) in 10 mM phosphate buffer (pH = 7.5). Before complexation the two ligand solutions are contained in two separated reservoirs (each having a volume of 0.8 mL), while after the complexation the overall volume of the resulting solution was 1.6 mL.
intermolecular attractions within the sample, also according to the high thermal stability found for the crystals up to 200 °C in air, as evidenced by use of an optical polarizing microscope equipped with a hot stage, revealing that thermal decomposition occurs only at temperatures higher than 250 °C. On the other hand, the two samples of TyrT/poly(A) complex show a morphology very different from that relative to pure TyrT, as it can be deduced by comparing the images on the top and bottom left of Fig. 6, both obtained at similar magnifications (∼1000×), as well as those present on the top and bottom right of the same figure at higher magnifications which show that the lamellar crystals of TyrT disappeared in presence of poly(A).

Computational analyses
Several computational methods were used to study the interaction between TyrT with our model of poly(A) RNA to analyze the binding mode from a structural point of view, at atomistic levels. The molecular docking did not allow a sufficient receptor flexibility. Since poly(A) is a very flexible molecule, MD simulation was deemed necessary to assess how the ligand accommodates in interaction with poly(A) and to evaluate which are the main interactions keeping the binding. During 1000 ps of simulation without any restraints, the flexibility of both molecules allows a better arrangement of the binding leaving the ligand to fit in an energetically favorable pose as indicated by the MMPBSA calculation.16,17 Indeed, the relative binding energy of the complex was found to be −17.5 (Generalized Born) and −4.6 (Poisson Boltzman) kcal mol⁻¹. In the proposed model, it is clearly indicated that TyrT fits into a region formed by two successive adenine rings, which slightly open to accommodate the thymine-ring portion of the ligand, the binding is mainly kept via hydrophobic contacts, reproducing a stacking-like interaction between the thymine ring of TyrT and the two successive adenine rings (see Fig. 7).

Serum stability study
Enzymatic stability of the thymine-based derivative was studied by incubating TyrT (150 μM) in 98% fresh human serum at 37 °C and analysing by RP-HPLC samples withdrawn from the reaction mixture at various times (0, 1, 2, 3, 4, 5, 24 and 168 h) after treatment with 7 M urea at 95 °C over 2 min (Fig. 8). As evident in Fig. 8, TyrT is stable even after 168 hour incubation.

Experimental section
Abbreviations
CD (Circular Dichroism); DIEA (N,N-diisopropylethylamine); DMF (dimethylformamide); Fmoc (9-fluorenylethoxycarbonyl);
bu

10
15

168 h

HATU  (O-(7-azabenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate); PNA (peptide nucleic acids); TCH$_2$COOH (thymin-1-yl acetic acid); TFA (trifluoroacetic acid); TIS (trislopropyl silane); Tyr (l-tyrosine).

**Fig. 8** HPLC profiles on a C-18 analytical column of aliquots relative to 0 and 168 hours withdrawn from incubation of the TyrT in 98% fresh human serum at 37 °C (detection at 260 nm). RP-HPLC chromatography was performed using a linear gradient of 2% (for 5 min) to 80% B in A over 20 min: (flow = 1 mL min$^{-1}$; A = 0.1% TFA in water; B = 0.1% TFA in acetonitrile).

**Chemicals**
Fmoc-\(\text{L-}{\text{Tyr}}(\text{Bu})\)-OH was Inbios, while HATU was Novabiochem. Ethanol was purchased from Fluka; piperidine, TCH$_2$COOH, TIS, diethyl ether and human serum were Sigma Aldrich. DMF, DIEA, TFA and HPLC chromatography solvents were purchased from Romil. Poly ra was Biomers.

**Apparatus**
\(^1\)H NMR and \(^{13}\)C NMR spectra were recorded at 25 °C on Varian unity 400 MHz spectrometer. Chemical shifts (\(\delta\)) are given in parts per million (ppm). Proton chemical shifts were referenced to residual CH$_3$OD (\(\delta = 3.30\), quin) signals. \(^{13}\)C NMR chemical shifts were referenced to the solvent (CD$_3$OD: \(\delta = 49.3\), sept). Samples were analysed and characterized by LC-MS on an MSQ mass spectrometer (ThermoElectron, Milan, Italy) equipped with an ESI source operating at 3 kV needle voltage and 320 °C, and with a complete Surveyor HPLC system, comprising an MS pump, an autosampler, and a PDA detector, by using a Phenomenex Jupiter (4u, 90A) Proteo (4 \(\mu\)m, 4.6 \(\times\) 150 mm) column. Gradient elution was performed (monitoring at 260 nm) by building up a gradient starting with buffer A (0.05% TFA in water) and applying buffer B (0.05% TFA in acetonitrile) with a flow rate of 0.8 mL min$^{-1}$. Analytical chromatograms were obtained on a Hewlett Packard/Agilent 1200 series HPLC, equipped with a diode array detector, by using a Phenomenex Jupiter C18 300 Å (5 \(\mu\)m, 4.6 \(\times\) 250 mm) column. Gradient elution was performed at 25 °C (monitoring at 260 nm) by building up a gradient starting with buffer A' (0.1% TFA in water) and applying buffer B' (0.1% TFA in acetonitrile) with a flow rate of 1 mL min$^{-1}$. Samples were lyophilized in a FD4 Freeze Dryer (Heto Lab Equipment) for 16 hours. Circular dichroism (CD) studies were conducted in analogy to other previous reports. CD spectra were obtained at 10 °C on a Jasco J-715 spectropolarimeter with a Tandem Hellma quartz cell, while, ultraviolet (UV) spectra were recorded on a UV-Vis Jasco model V-550 spectrophotometer equipped with a Peltier ETC-505T temperature controller, using a Hellma quartz cell with a light path of 1 cm.

After slow solvent evaporation, samples were coated with a conductive layer of Au–Pd and examined in a Nova NanoSem 450 FEI SEM at 15 kV in high vacuum mode. Phase behaviour was analysed using an optical microscope (Zeiss Axioscop) equipped with a Mettler Toledo FP82HT Hot Stage.

**Computational studies**
A starting model of single strand RNA poly(A) structure, composed by 20 adenine residues, was built using the server: http://structure.usc.edu/make-na/server.html, this model was then used as target structure for docking studies with the TyrT molecule. Molecular docking was performed using AutoDock Vina, AutoDock Tools was utilized to prepare the input pdbqt file for both poly(A) and TyrT molecules and to set the size and center of the grid box. The predicted binding affinity (kcal mol$^{-1}$), which indicates how strongly a ligand binds to the receptor, is calculated based on the scoring function used in AutoDock Vina and is used to select the best pose of the ligand. The final binding pose with the best AutoDock Vina score was selected for further optimization through molecular dynamics (MD). In details, 1000 ps of MD simulation at 283 K in explicit waters was performed and analyzed using the GROMACS simulation package, with Parmbsc0 force field, applying a computational protocol successfully employed to the study of standard with modified oligonucleotides. Since the ligand is not recognized by the force field, the PRODRG (http://davapc1.bioch.dundee.ac.uk/prodrg/) server was used to create an automated topology of it, which was manually modified taking into account the atom type of the Amber99sb force field library. To obtain the representative pose of the interaction, the MD trajectory was clustered using Gromacs package, a representative pose of the whole trajectory was selected from the most populated cluster on the basis of the best relative binding free energy, computed through MMPBSA.

**Conclusions**
The nucleobase–tyrosine conjugate object of the present study was obtained by a solid phase synthetic strategy and, after purification and MS/NMR characterization, it was employed in poly(A) RNA binding studies performed via CD spectroscopy in analogy to our previous studies, that revealed its ability to bind the RNA target. The interaction of TyrT with poly(A) was studied also by molecular docking that suggested the importance of the T nucleobase in the molecular recognition of poly(A) in the
overall interaction process. Moreover, SEM microscopy evidenced that the interaction with poly(A) dramatically altered the morphology of samples of TyrT treated with the RNA. Finally, a high enzymatic stability of TyrT was found by human serum HPLC assay. Overall, all these findings encourage us to devote our future efforts on molecular systems of the type presented in the present study as potential poly(A) binders with potential importance in biomedicine.

Acknowledgements

We thank Prof. Antonio Roviello for his precious suggestions, and Mr Leopoldo Zona for his invaluable technical assistance. We also thank Consiglio Nazionale delle Ricerche (CNR) for the research grant received under the bilateral CNR, Italy-SRNSF, Georgia research Program (2014–2015) and Regione Campania for the research grant received for the research project “Nuovi sistemi nucleopeptidici per applicazioni diagnostiche” under the research program Legge Regionale n. 5 (annualità 2007).

References

4 W. T. Ruyechan and J. W. Olson, Surface lysine and tyrosine residues are required for interaction of the major herpes simplex virus type 1 DNA-binding protein with single-stranded DNA, J. Virol., 1992, 66, 6273–6279.
