Thyroglobulin type-I-like domains in invariant chain fusion proteins mediate resistance to cathepsin L digestion

Christoph Hitzel[a,1], Holger Kanzler[a,1,2], Angelika König[a], Markus P. Kummer[b], Klaudia Brix[b], Volker Herzog[b], Norbert Koch[a,*]

[a] *Division of Immunobiology, University of Bonn, Römerstraße 164, D-53117 Bonn, Germany*
[b] *Institute of Cell Biology, University of Bonn, Bonn, Germany*

Received 22 August 2000; revised 11 October 2000; accepted 23 October 2000

First published online 2 November 2000

Edited by Pierre Jolles

Abstract The MHCII associated invariant chain isoform Ii41 shows homology to a repeat in thyroglobulin (TgR). We show that the Ii31 isoform, which lacks the TgR-like domain, is sensitive to cathepsin L treatment whereas Ii41 displays substantial resistance. The TgR-like sequence of Ii41 was exchanged for thyroglobulin type-IA and -IB repeats, that contain six or four cysteine residues. Resistance to cathepsin L digestion was maintained upon substitution of the Ii41 TgR for homologous sequences from TgR type-IA. Mutation of a conserved cysteine in the TgR domain of an Ii fusion protein strongly reduced resistance to cathepsin L digestion. **© 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.**

Key words: MHC; Invariant chain; Cathepsin inhibitor; Thyroglobulin

1. Introduction

Thyroglobulin, the precursor of thyroid hormones, contains a domain with ten repeats that include six or four cysteine residues (TgR type-IA or -IB) [1–3]. In the MHCII associated invariant chain (Ii41), a sequence homologous to the TgR type-IA is encoded by an alternatively spliced exon of the Ii gene [4]. First evidence for a functional role of this TgR-like domain of Ii was obtained when a fragment of Ii41 was co-isolated with cathepsin L [5]. Inhibition of enzyme catalyzed substrate conversion revealed that the Ii41 fragment blocks the function of cathepsin L [6]. The crystal structure of the Ii41 TgR-like fragment shows a three loop arrangement stabilized by three disulfide bonds with the wedge shaped Ii41 fragment anchored in the active center of cathepsin L [7]. The importance of this finding is apparent [8] as the regulated degradation of Ii governs the uptake of antigenic peptides by MHCII dimers, a key step in initiating an immune response. Ii is a substrate for cathepsin L under physiological conditions, as was shown using cathepsin L deficient mice [9].

TgR-like domains are found in an increasing number of functionally unrelated proteins from different species [10]. Some of the proteins contain several TgR-like domains. Efforts have been made to identify the function of these domains. Partial proteolysis has been used to separate TgR-like domains in equistatin from sea anemone [11]. One of the fragments was an inhibitor of cysteine proteases, while another fragment blocks cathepsin D.

To determine whether TgR-like domains other than the Ii41 inhibit the function of cathepsin L, the Ii41 TgR domain was exchanged for TgR type-IA and -IB elements from bovine thyroglobulin or a homologous sequence from human epithelial glycoprotein (EGP). We show here that Ii TgR recombinant molecules exhibit a remarkable resistance to cathepsin L digestion.

2. Materials and methods

2.1. Transfection, labeling, immunoprecipitation and SDS–PAGE

5×10^5 COS cells were transfected with 1 μg plasmid DNA and 10 μg cationic lipid (DOSPER, Boehringer-Mannheim, Germany). Fusion of the lipid complex with cell membranes was accomplished by addition of 10% FCS. Cells were labeled for 30 min with 30 μCi [^{35}S]methionine (Amersham) and lysed with 1% NP40 [12]. Ii was immunoprecipitated by a mixture of mAb In1 (rat anti-mouse Ii), mAb MAR18.5 (anti-rat Ig), and 10 μl protein A Sepharose CL4B. Following overnight incubation, precipitates were washed four times and analyzed on SDS–polyacrylamide gels. Gels were dried and radioactivity was visualized by exposure to X-ray films.

2.2. Cathepsin L digestion of Ii

Cathepsin L (Calbiochem) was adjusted to 0.05 μU/ml to achieve degradation of wild-type Ii41 within 24 h. Ii immunoprecipitates immobilized with protein A Sepharose were suspended in 15 μl of 20 mU cathepsin L/ml in 0.04 M citrate buffer, pH 5. The cysteine protease was activated by 5 mM L-cysteine and Ii was digested at 37°C. The digestion was stopped by SDS sample buffer (pH 7.8) and 5 min boiling. Protein fragments were separated by SDS–PAGE.

2.3. Mutagenesis of the Ii41 TgR domain

Sequences of the TgR-like domain of murine Ii were substituted with TgR domain I.1 (type-IA) or I.9 (type-IB) encoded by the bovine thyroglobulin gene [13] or the TgR domain encoded by the human epithelial glycoprotein (EGP) gene using published methods [14].

Sequences of domain I.1 [13] were amplified using primer p41.TgI.1/5 (5′-GTCTACCCGGGTGCGTTCCGTCCCCAGTGCGCCGAGGATGGC-3′) in combination with p41.TgI.1/3 (5′-GTCTTCCATGTCCAGTGGCTCCAGACAGGCCGCAGGCCGCCC-3′). Sequences binding to Ii41 cDNA are underlined.

*Corresponding author. Fax: (49)-228-734555.
E-mail: norbert.koch@uni-bonn.de

[1] Contributed equally to this work.
[2] Present address: DNAX Research Institute, California, USA.

Abbreviations: MHCII, major histocompatibility complex class II molecules; Ii, MHCII associated invariant chain; Tg, thyroglobulin; TgR, thyroglobulin repeat; FCS, fetal calf serum; EGP, epithelial glycoprotein

0014-5793/00/$20.00 © 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.
PII: S0014-5793(00)02189-X

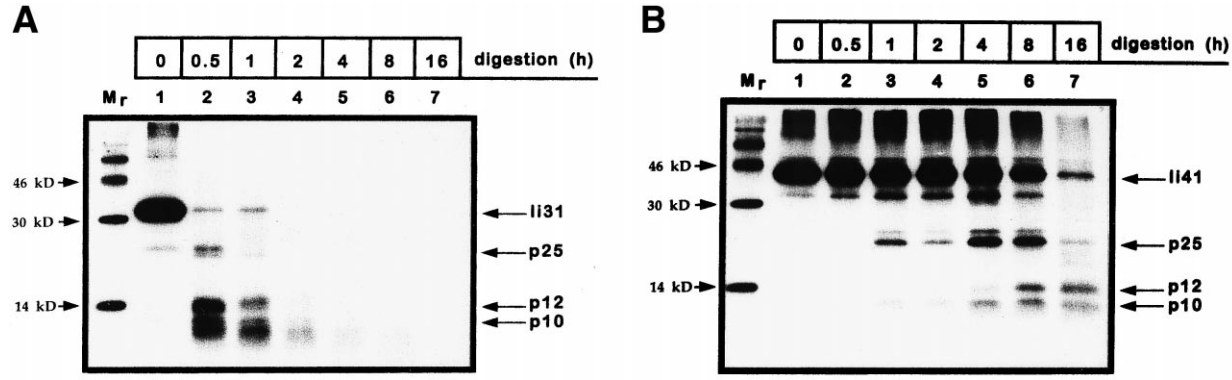

Fig. 1. Digestion of Ii31 and of Ii41 with cathepsin L. COS cells were transiently transfected with cDNA encoding Ii31 (A) or Ii41 (B). Cells were labeled with [^{35}S]methionine and Ii was immunoprecipitated with mAb In1. Immobilized Ii immunoprecipitates were digested with 20 mU/ml of cathepsin L. The digestion was stopped after 0.5, 1, 2, 4, 8, and 16 h by SDS sample buffer and boiling. Ii and its degradation products are separated by SDS–PAGE. Mobilities of proteins are indicated at the right and molecular weight markers on the left.

Bovine TgR domain I.9 was amplified using primer p41.TgI.9/5 (5′-GTCTACCCGGGTGCGTTCCGTCCAACCTGCTTAGAGACAGGA-3′) in combination with p41.TgI.9/3 (5′-GTCTTCCATGTCCAGTGGCTCCGGGCACTGGGCACTGCTGTT-3′). Sequences of the TgR domain of EGP [15] were amplified using primers p41.EGP/5 (5′-GGGTGCGTTCCGTCCTGACTGCGATGAG-3′) and p41.EGP/3 (5′-TAGGTCTTCCATGTCCAGTGGCTCAGAGCAGGTTATTTCAGT-3′). The reaction mix contained 10 mM Tris/HCl, pH 8.4, 50 mM KCl, 0.1 mg/ml gelatin, 200 μM dNTPs, 100 pM phosphorylated primers, $MgCl_2$ (1.5 mM) for bovine TgR I.1 or EGP and TgR, 4.5 mM for bovine TgR I.9, 50 ng of plasmid containing the target DNA and 2.5 U Taq DNA polymerase (Boehringer-Mannheim). The enzyme was added after initial denaturation. Products representing the bovine TgR I.1 and I.9 were phosphorylated using polynucleotide kinase. PCR products were mixed (50:1) with plasmid pUC18.Ii41 containing the Ii41 cDNA. Annealing was performed by heating the samples to 100°C for 5 min and immediate cooling on ice. Four μl of 100 mM Tris/HCl, pH 7.5, 5 mM dNTPs, 10 mM ATP, 20 mM DTT, 2 μl H_2O, 1 μl T4 DNA polymerase (Amersham/USB; 3.5 U/μl) and 3 μl T4 DNA ligase (Amersham/USB; 1 U/μl) were added. The samples were incubated at 4°C for 5 min, at room temperature (RT) for 5 min and at 37°C for 120 min. Aliquots were used to transform *Escherichia coli* strain BMH 71-18 mutS by electroporation. 200 ng of plasmids purified from overnight bacteria cultures were digested with restriction enzyme *Sac*II and aliquots were used in electro-transformation of *E. coli* strain DH5α. DNA fragments encoding Ii41 with a substituted TgR-like domain were subcloned into expression vector pcEXV3 and sequence substitutions were confirmed by nucleotide sequencing.

3. Results

3.1. Differential sensitivity of Ii31 and Ii41 to cathepsin L digestion

To investigate degradation of the two Ii isoforms by cathepsin L in vitro, we expressed Ii31 or Ii41 cDNA in COS cells. After metabolic labeling with [^{35}S]methionine, cells were lysed and Ii was immunoprecipitated with mAb In1 and subsequently digested with cathepsin L. Ii31 is rapidly degraded to fragments in the range of 25 kDa and 10–12 kDa (Fig. 1A). In contrast, Ii41 exhibits a substantial resistance to degradation by cathepsin L (Fig. 1B) with complete digest taking longer than 16 h. Since the TgR domain of Ii41 responsible for inhibition of cathepsin L [6] was also found in other proteins, we wanted to determine whether homologous TgR domains exhibit an inhibitory function on Ii degradation by cathepsin L. Recombinant Ii chains were constructed where aa 213–255 of the Ii41 sequence were replaced by homologous sequences of TgR (type-IA or -IB repeats) or of EGP (Fig. 2). In addition, a mutant Ii chain, Ii EGPC/F was generated. The mutant shows an exchange of Cys-215 for Phe of the EGP-TgR domain, which disrupts the disulfide bond in the first loop.

3.2. Structural requirements of the TgR for inhibition of cathepsin L

To test for inhibition of cathepsin L digestion, recombinant Ii chains were immunoprecipitated from transfected COS cells and digested (Fig. 3). Proteolysis of Ii TgR I.1 (Fig. 3A) and

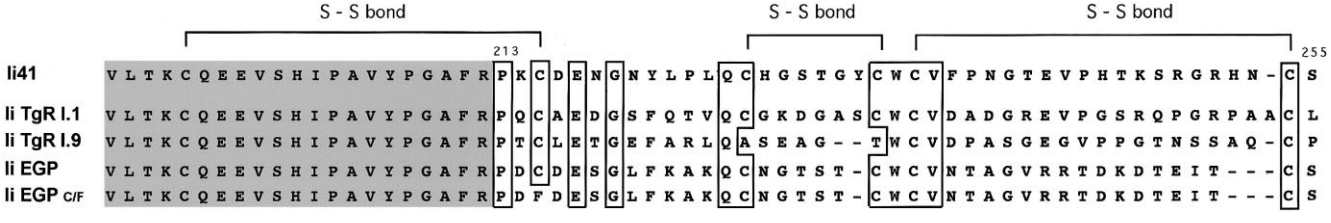

Fig. 2. Partial sequence of the Ii41 domain and Ii TgR fusion proteins. The sequence aa 213–255 of Ii41 was exchanged for sequences from bovine thyroglobulin type-IA or -IB repeats (TgR I.1 or TgR I.9), by EGP, or by the mutant EGPC/F. The type-IA TgR repeat consists of three loops that are stabilized by disulfide bonds. To minimize structural changes to the Ii fusion protein, the sequence of the first loop that shows most variable length in the TgR domains was not exchanged. In TgR I.9 (type-IB repeat) which contains only four cysteines, the second disulfide bond cannot form. EGPC/F has Cys-215 exchanged by Phe. Locations of disulfide bonds are indicated. Regions with a high degree of conservation are boxed. The Ii41 sequence is shaded.

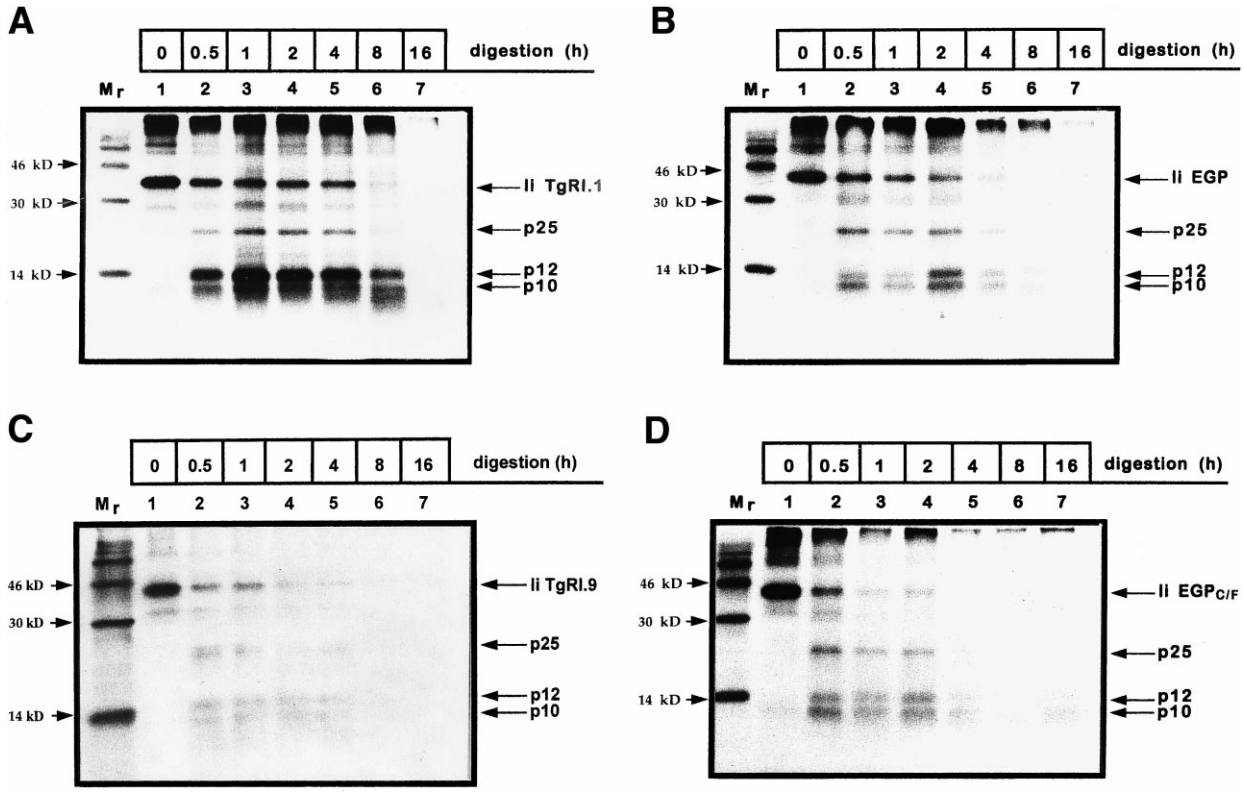

Fig. 3. Digestion of recombinant Ii by cathepsin L. Ii TgR I.1 (A), Ii EGP (B), Ii TgR I.9 (C), and Ii EGPC/F (D) were transiently expressed in COS cells, labeled with [^{35}S]methionine and immunoprecipitated with mAb In1. Immunoprecipitates were digested with 20 mU/ml cathepsin L. Digestion was stopped by addition of SDS sample buffer. Samples were separated by SDS–PAGE. Molecular weight markers are shown on the left.

of Ii EGP (Fig. 3B) revealed that both polypeptides exhibit a reduced, but still substantial resistance to cathepsin L treatment when compared to wild-type Ii41 (Fig. 1A). The TgR type-IB variant expressed in Ii TgR I.9 (Fig. 3C) shows rapid proteolysis by cathepsin L treatment that resembles the rate of digestion of Ii31 (Fig. 1B), suggesting that the absence of the disulfide bond in the second loop of the TgR type-IB repeat results in a reduced sensitivity to the protease. Similarly, exchange of Cys-215 in Ii EGPC/F (Fig. 2) that prevents the first disulfide bond largely abolishes the resistance to cathepsin L digestion (Fig. 3D). This recombinant Ii chain is almost as sensitive to degradation by cathepsin L as wild-type Ii31 (Figs. 1B and 3D).

For a semiquantitative comparison the decay constants of recombinant Ii chain degradation by cathepsin L were determined. The chimeric TgR I.1 and EGP show decay constants of $k = 1 \times 10^{-4}$ s^{-1} and $k = 2 \times 10^{-4}$ s^{-1} compared to a k of about 5×10^{-5} s^{-1} for Ii41. This result may suggest that inhibitory constants of TgR I.1 and EGP for cathepsin L activity are markedly lower than that of Ii41.

4. Discussion

Ii is expressed in two forms, Ii31 and Ii41, that differ in a 64 aa sequence containing a TgR-like domain [4]. An immunological function of the TgR-like domain in Ii41 has not yet been elucidated. Antigen presenting cells that express either Ii31 or Ii41 process and present antigen equally [16]. This result was confirmed in recombinant mice, expressing either Ii31 or Ii41, which show a similar phenotype [17]. The role of the TgR-like domain in Ii41 could be to modulate degradation of antigenic peptides or to provide a mechanism to regulate degradation of Ii by cathepsin L. The different susceptibility of Ii31 and Ii41 to degradation by cathepsin L, as shown in this report (Fig. 1), points to a mechanism of modulating the release of Ii from MHCII molecules. This finding might provide an explanation why Ii41 has a prolonged half life compared to Ii31 [18].

Both Ii31 and Ii41 form trimers and are present in the same complex [19]. Although in most cells Ii31 is expressed more abundantly than Ii41, in some cell types, like DC, macrophages and Langerhans cells, the amount of Ii41 approaches 30–40% of total Ii expression [18]. Using transfected cells, it was shown that the presence of Ii41 in Ii trimers hampers degradation of Ii31 [20]. Upon release of Ii, MHCII dimers become susceptible for binding of antigenic peptides [21]. Hence the inhibition of cathepsin L by Ii41 may delay loading of MHCII dimers with peptides.

The main characteristics of the TgR domain are a Cys-Trp-Cys-Val motif and two or four additional cysteines that stabilize the structure by two to three disulfide bonds. Although the aa residues that maintain the tertiary structure of the TgR domain, as revealed by the crystal structure [7], are highly conserved among divers proteins, the total identity of the sequences is less than 26%. Thus, it is remarkable that exchange of the TgR-like domain of Ii41 by homologous sequences from bovine TgR and human EGP, still resulted in retention of the domain's inhibitory function (Fig. 3).

It is as yet unknown whether the type-IA TgR repeat exerts a cathepsin L inhibitory function in situ. Cathepsin L has

been detected on the lumenal plasma membrane of thyrocytes [22]. TgR might, therefore, be able to regulate cathepsin L activity in the follicular lumen and to inhibit precocious proteolysis of thyroglobulin.

The discovery of the TgR-like domain in a number of proteins suggests widespread importance. TgR homologous sequences have also been found in a cysteine proteinase inhibitor from *Actinia equina* [23] and a sequence from *Caenorabditis elegans* [24] (gb/AAB92030.2/). The presence of the TgR-like domain in invertebrates suggests that a protease inhibitory function predates the immune system. Possibly, the protease inhibitory function of the TgR-like domain of Ii41 embodies an early evolutionary defense mechanism, e.g. to parasites such as *Leishmania* or *Trypanosoma cruzi*. These organisms require cysteine proteases for survival and invasion [25].

Acknowledgements: We thank Dr. S. Kellermann for critical reading of the manuscript. This work was supported by the Sonderforschungsbereich 284 Teilprojekte B1 and B6 and by Fonds der chemischen Industrie.

References

[1] Mercken, L., Simons, M.J., Swillens, S., Massaer, M. and Vassart, G. (1985) Nature 316, 647–651.
[2] Malthiery, Y. and Lissitzky, S. (1985) Eur. J. Biochem. 147, 53–58.
[3] Molina, F., Bouanani, M., Pau, B. and Granier, C. (1996) Eur. J. Biochem. 240, 125–133.
[4] Koch, N., Lauer, W., Habicht, J. and Dobberstein, B. (1987) EMBO J. 6, 1677–1683.
[5] Ogrinc, T., Dolenc, I., Ritonja, A. and Turk, V. (1993) FEBS Lett. 3, 555–559.
[6] Bevec, T., Stoka, V., Pungercic, G., Dolenc, I. and Turk, V. (1996) J. Exp. Med. 183, 1331–1338.
[7] Guncar, G., Pungercic, G., Klemencic, I., Turk, V. and Turk, D. (1999) EMBO J. 18, 793–803.
[8] Cresswell, P. (1994) Annu. Rev. Immunol. 12, 259–293.
[9] Nakagawa, T., Roth, W., Wong, P., Nelson, A., Farr, A., Deussing, J., Villanangos, J.A., Ploegh, H., Peters, C. and Rudensky, A.Y. (1998) Science 280, 450–453.
[10] Lenarcic, B. and Bevec, T. (1998) Biol. Chem. 379, 105–111.
[11] Lenarcic, B. and Turk, V. (1999) J. Biol. Chem. 274, 563–566.
[12] Hitzel, C., van Endert, P. and Koch, N. (1995) J. Immunol. 154, 1048–1056.
[13] Mercken, L., Simons, M.-J., De Martinoff, G., Swillens, S. and Vassart, G. (1985) Eur. J. Biochem. 147, 59–64.
[14] Deng, W.P. and Nickoloff, J.A. (1992) Anal. Biochem. 200, 81–88.
[15] Simon, B., Podolsky, D.K., Moldenhauer, G., Isselbacher, K.J., Gattoni-Celli, S. and Brand, S.J. (1990) Proc. Natl. Acad. Sci. USA 87, 2755–2759.
[16] Stockinger, B., Pessara, U., Lin, R.H., Habicht, J., Grez, M. and Koch, N. (1989) Cell 56, 683–689.
[17] Shachar, I., Elliott, E.A., Chasnoff, B., Grewal, I.S. and Flavell, R.A. (1995) Immunity 3, 373–383.
[18] Kämpgen, E., Koch, N., Koch, F., Stöger, P., Heufler, C., Schuler, G. and Romani, N. (1991) Proc. Natl. Acad. Sci. USA 88, 3014–3018.
[19] Gedde-Dahl, M., Freisewinkel, I., Staschewski, M., Schenck, K., Koch, N. and Bakke, O. (1997) J. Biol. Chem. 272, 8287.
[20] Fineschi, B., Arneson, L.S., Naujokas, M. and Miller, J. (1995) Proc. Natl. Acad. Sci. USA 92, 10257–10261.
[21] Rabinowitz, J.D., Vrljic, M., Kasson, P.M., Liang, M.N., Busch, R., Boniface, J.J., Davies, M.M. and McConnel, H.M. (1998) Immunity 9, 699–709.
[22] Brix, K., Lemansky, P. and Herzog, V. (1996) Endocrinology 137, 1963–1974.
[23] Lenarcic, B., Ritonja, A., Strukelj, B., Turk, B. and Turk, V. (1997) J. Biol. Chem. 272, 13899–13903.
[24] The *C. elegans* Sequenzing Consortium (1998) Science 282, 2012–2018.
[25] Bevec, T., Stoka, V., Pungercic, G., Cazzulo, J.J. and Turk, V. (1997) FEBS Lett. 401, 259–261.