Thyroperoxidase and Thyroid Autoimmunity. Ed. P. Carayon, J. Ruf. Colloque INSERM/John Libbey Eurotext Ltd. © 1990. Vol. 207, pp. 351-358

Thyroid peroxidase autoantigen : localization of autoantigenic epitopes on recombinant protein and prediction of secondary structure

J.P. Banga, D.L. Ewins, P.S. Barnett, R.W.S. Tomlinson, D. Mahadevan, G.J. Barton, B.J. Sutton, J.W. Saldanha, E. Odell and A.M. McGregor

Department of Medicine, King's College School of Medicine, London; Imperial Cancer Research Fund Laboratories, Biomedical Computing Unit, London; Division of Biomolecular Sciences, King's College, London; United Medical and Dental Schools of Guy's and St Thomas', London Bridge, London, UK

SUMMARY

The cloning and expression of TPO as a recombinant protein has allowed progress on the characterisation of the autoantigenic epitopes recognised by autoantibodies. The main immunogenic region of the molecule recognised by autoantibodies has been localized to the carboxyl terminal. Multiple sequence alignment and secondary structure prediction methods used in conjunction with circular dichroic spectroscopy indicate that TPO structure is principally alpha-helical. Together with a knowledge of the exon-intron boundaries suggests a model for the domain organization of the TPO molecule. The ability in the future to map the autoantigenic epitopes on a three dimensional structural model of the TPO molecule will allow a greater understanding of the aetiological basis of autoimmune disease and allow the design of therapeutic strategies at an antigen specific level to be addressed.

INTRODUCTION

Thyroid peroxidase is the primary enzyme involved in thyroid hormone synthesis and a target autoantigen in destructive thyroid disease leading to thyroid failure in man (Banga et al, 1990a). The autoimmune response to TPO involves both the production of autoantibodies (aAbs) and T cell autoreactivity. A great deal of information is available on the autoantibody response to TPO in autoimmune thyroid disease (AITD). These autoantibodies are known to be polyclonal in nature where several antigenic sites, including the enzyme active site, are known to be recognized (Doble et al, 1988). Additionally, anti-TPO aAbs cross-react with other related peroxidases such as myeloperoxidase (MPO) which shares a high degree of sequence homology (Banga et al, 1989a). However, so far, there is a scarcity of information on the nature of the T cell autoreactivity or the epitopes on the TPO molecule recognized by T cells in AITD (Banga et al, 1990a).

The ability to dissect and study the autoimmune response in vitro requires a source of purified autoantigen. For TPO, a number of procedures have been developed for its purification but the purity and, in particular, the yield of the molecule has been relatively poor to be used for cellular investigations. To circumvent this, we have applied molecular biology techniques to clone and express TPO as a recombinant fusion protein in *E.coli*. In particular, we have taken advantage of polymerase chain reaction (PCR) to amplify various regions of the TPO molecule (averaging 120 amino acids) for expression as recombinant proteins. This has allowed us to identify the main immunogenic region on the TPO molecule recognized by autoantibodies from patients with AITD.

A detailed knowledge of the three dimensional structure and an accurate localization of the autoantigenic epitopes on the TPO molecule would greatly increase our understanding of the autoimmune response to this autoantigen. It would also allow the future design of

therapeutic studies aimed for immune intervention at an antigen specific manner to be addressed (Banga et al, 1989b). Structural knowledge is usually obtained by X-Ray crystallography which requires access to milligram amounts of purified protein. Whilst this is currently not feasible for TPO, even with the availability of recombinant bacterial TPO protein, structure prediction methods developed over the past decade now allow protein structure to be predicted. Combining prediction methods can yield improvements in accuracy as can combining predictions on accurately aligned sequences. Accordingly, we have aligned the sequences of human TPO, porcine TPO and human MPO and applied different prediction methods on the aligned sequences to obtain a consensus structural prediction and the organization of domains of the TPO molecule. In view of the high sequence similarity between human TPO and MPO, we have also used purified MPO for circular dichroic (CD) spectroscopy; since purified preparations of intact TPO were not available, we have also used trypsinized, porcine TPO for CD spectroscopy. By combining the CD spectral information on MPO with the secondary structure prediction of TPO, we show that the structure of TPO is mainly alpha-helical with little beta sheet, and is organized into distinct domains.

METHODS

CD spectral analysis

Trypsinized pTPO was prepared as already described and was a gift from Professor A Taurog (Yokoyama and Taurog, 1988). hMPO was prepared from neutrophils (Banga et al, 1989a). The purity of these peroxidase preparations was assessed by SDS-polyacrylamide gel electrophoresis. For CD analysis, pTPO was resuspended at 0.1mg/ml in phosphate buffered saline whilst hMPO was at 0.1mg/ml in 0.1M Tris-Hcl pH 7.4. CD spectra were run at room temperature in 0.05cm "strain free" quartz 2 cells in a JASCO J600 CD spectrometer.

Multiple sequence alignment and prediction techniques

Sequence alignment was performed by the algorithm of Barton (1990) using the AMPS package. Alpha-helix and beta-strand regions were predicted using the algorithms of Lim (Lim, 1974), Chou and Fasman (Chou & Fasman, 1978) and Robson (Garnier et al, 1978), as programmed in the Leeds Prediction suite. Turns were predicted using the method of Wilmot and Thornton (Wilmot and Thornton, 1988) and that of Rose (Rose, 1987). B cell defined epitopes were predicted using the Hopp and Woods profile (Hopp and Woods, 1987). T cell defined epitopes were predicted using the motifs programme of Rothbard and Taylor (Rothbard and Taylor, 1988).

Expression of TPO as recombinant protein

Different segments of the TPO polypeptide were expressed as fusion proteins in *E.coli* using recombinant biology techniques and polymerase chain reaction as already described (Banga et al, 1989c, 1990b). The recombinant TPO proteins were assessed for their binding to autoantibodies to TPO from patients with thyroid autoimmune disease by immunoblotting techniques.

RESULTS AND DISCUSSION

SDS-polyacrylamide gel electrophoresis of pTPO and hMPO

Preparations of pTPO and hMPO were analysed by SDS-polyacrylamide gel electrophoresis to assess the purity of biochemically purified proteins. The trypsinized pTPO preparations under non-reducing or reducing conditions gave protein staining bands already described (Yokayama and Taurog, 1988) (Fig. 1, Lanes 1 NR and 1R respectively). hMPO preparation under reducing conditions gave a strong and a major protein staining band at 60Kd comprising the heavy polypeptide chain of the enzyme (Banga et al, 1989a). (Fig. 1, Lane 2R). Both the preparations of pTPO and hMPO were > 90% purity and used for CD analysis.

CD analysis of pTPO and hMPO

The far ultra-violet spectra of trypsinized pTPO and hMPO in aqueous solution exhibit a broad band of minimum in the 200-210nm range (Fig. 2). The values of alpha-helical content were calculated to be approximately 27-33% for trypsinized pTPO and 55-60% for hMPO, with the remainder of the protein in both preparations in random coil conformation. The lower alpha-helical content in the trypsinized pTPO preparations is probably due to the fact that it is a fragmented molecule which lacks approximately 90 residues at the NH_2 and COOH terminals and also contains a trypsin cleaved site after residue 561..

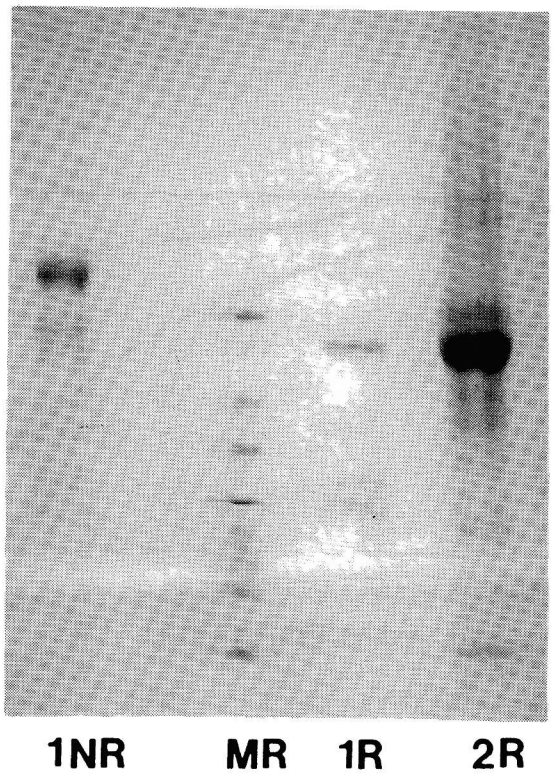

Fig. 1 SDS-polyacrylamide gel electrophoresis of purified preparations of trypsinized, porcine TPO and human MPO which were used for CD spectroscopy. R refers to reducing conditions, NR refers to non-reducing conditions. Sample 1 is trypsinized, porcine TPO whilst sample 2 is MPO. M= molecule weight markers migrating at 66,45,36,29,24,20 and 14.2Kd.

Sequence alignment and secondary structure prediction techniques

Secondary structure prediction methods generally give around 55-65% accuracy for a three state prediction of helix, strand and non-helix/non-strand (i.e. coil) structure. Recently, it has become clear that the accuracy of the various secondary structure prediction methods can be greatly increased by (i) combining the prediction methods and (ii) combining predictions on accurately aligned sequences (Biou et al, 1988; Zvelebil et al, 1987). We have achieved this for TPO by aligning the sequences of hTPO, pTPO and hMPO by the algorithm of Barton and colleagues (Barton and Sternberg, 1987; Barton, 1990) (alignment not shown, but see Banga et al, 1990c). Five prediction methods were then performed independently on each aligned sequence and then the results combined into a consensus.

The overall result of predictions for TPO in the region of homology with MPO (residues 1-741) were alpha-helix 51%, beta sheet 8%, turn 41% and the remainder as random coil (Banga et al, 1990c). This compares favourably with the CD results of MPO which predict alpha-helical conformation of 55-60% and the remainder as random coil.

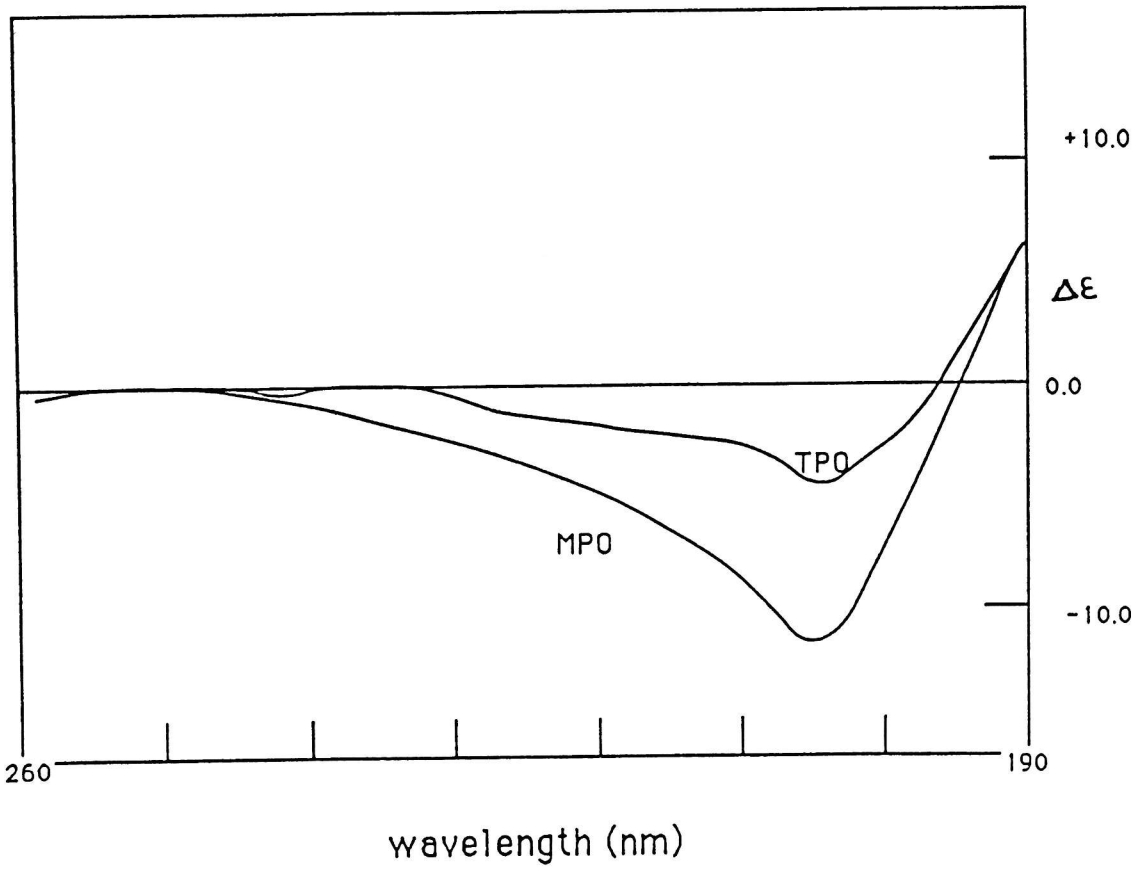

Fig. 2 Circular dichroic (CD) spectroscopy of trypsinized, porcine TPO and human MPO in the far ultraviolet region in aqueous solution. The difference in the two spectra are likely to reside in the fact that the trypsinized, porcine TPO is a fragmented molecule which is also truncated at the amino and carboxyl terminals.

The alignment of two or more sequences also allows further structural information to be obtained. The three proteins share a large number of identical amino acids (hTPO vs pTPO = 70% identity; hTPO vs hMPO as aligned by the algorithm of Barton (1990) = 50%) (for the alignment, see Banga et al, 1990c). Consideration of insertions or deletions in the aligned sequences shows areas in surface loop regions (Bashford et al, 1987). Four insertions between the hTPO and pTPO are apparent at positions E at 388 and L at 518 of hTPO and insertion of GK at 896 and LPG at 903 of pTPO and loops are predicted in these positions. When this analysis is extended to hMPO, a further 10 insertions are observed. The alignment of sequences also allows the conservation of cysteine residues to be ascertained. Sixteen cysteine residues are conserved between hTPO and pTPO. In the region of sequence homology with hMPO (residues 1-741), fourteen of these cysteine residues are conserved with hMPO (Banga et al, 1990c). The extensive conservation of cysteine residues strongly supports the contention

that TPO and MPO are structurally homologous throughout the region of TPO that corresponds to MPO.

Recently the gene structures for hTPO and hMPO have been described where the hTPO gene comprises 17 exons and the hMPO gene consists of 12 exons (Kimura et al, 1989). When the exon-intron boundaries of these two genes are compared, exon 3-11 of TPO and exons 2-11 of MPO coincide exactly except that exon 8 of TPO corresponds to exons 7 and 8 of MPO (Kimura et al, 1989). The use of sequence alignment and exon-intron boundaries between related proteins has previously been used to successfully predict the domain organization of other proteins (Inna et al, 1983; Williams, 1984). Using this approach, the domain organization of TPO is predicted into nine distinct domains (Banga et al, 1990c).

The TPO region 799-847 has been shown to have a high sequence similarity to EGF (Libert et al, 1987). Additionally, there is absolute conservation of the six cysteine residues that are disulphide bonded in the EGF molecule - these cysteine residues in hTPO are at residue positions 799, 805, 819, 828 and 843 (Banga et al, 1990c). This allows the structure of TPO in this region to be predicted confidently by analogy with the NMR structure of EGF (Cooke et al, 1987). The structure of EGF consists of two B-strands and turns held together by a cluster of three disulphide bonds.

TABLE 1

Predicted B cell (antibody) and T cell epitopes of TPO

	ANTIBODY EPITOPE		T CELL EPITOPE	
	Residues	Sequence	Residues	Sequence
1.	684-697	DAQRRELEKHSLSR	562-568	EELTERL
2.	223-230	VTDDDRYS	47-53	KRLVDTA
3.	33-38	KPEESR	87-93	SGVIARA
4.	760-767	HCEESGRR	426-431	KALNAH
5.	60-65	RNLKKR	119-125	TDALSED
6.	359-365	RLRDSGR	94-99	AEIMET

Taken together, multiple sequence alignment and a combination of secondary structure prediction methods indicate that the TPO structure is principally alpha-helical. Together with a knowledge of the exon-intron boundaries suggests that the TPO molecule may fold into nine separate domains (Banga et al, 1990c).

Prediction of antibody and T cell defined epitopes

The TPO molecule is the target autoantigen in lymphocytic thyroiditis (Hashimoto's disease). We have used the Hopp and Woods hydrophilicity profiles to predict the amino acid residues in TPO which may be recognized by antibodies (Table 1). Since T cell help is required for antibody production, amino acid sequences that may serve as T cell epitopes are also described, using the motifs programme of Rothbard and Taylor (1988) (Table 1). The top six hydrophilic peaks for antibody epitopes and the top six T cell epitopes are predicted (Table 1).

Generation of recombinant preparations of TPO and mapping of autoantigenic sites

Different segments of the TPO molecule have been expressed as recombinant protein in *E.coli* as fusion proteins with glutathione-S-transferase (Banga et al, 1989c, 1990b). The recombinant overlapping TPO fusion proteins encompass the following amino acid residues 1-160, 145-250, 457-589, 577-677, 657-767 and 737-845. A larger fragment encompassing the latter three recombinants was also produced (residues 577-845). Using Western blotting and autoantibodies to TPO from patients with thyroid autoimmune disease, we recently showed that these sera recognize recombinant TPO proteins containing residues 657-767 and 145-250 of the fusion molecule. Interestingly, the latter recombinant TPO preparations harbour the two most hydrophilic peaks of position 684-697 and 223-230 respectively which have been predicted to contain the sequential antigenic determinants (Table 1) (Banga et al, 1989c, 1990b). Nine out of ten AITD sera contained autoantibodies to TPO which react with recombinant TPO fusion proteins derived from the COOH terminal of the molecule. These studies, together with another report (Ludgate et al, 1989), suggests that the main immunogenic region of the TPO polypeptide is localized towards the carboxyl region of the molecule.

ACKNOWLEDGEMENTS

Trypsinized, porcine TPO was kindly given by Professor A Taurog. We thank Drs J Fox and C Rawlings for advice and Dr A Drake and Mr J Hoardley for help with CD spectroscopy. Many thanks to Mrs J De Groote for organizing the manuscript preparation. Supported by grants from The Wellcome Trust and the Medical Research Council.

REFERENCES

Banga, J.P., Barnett, P.S., and McGregor, A.M. (1990a): Immunological and molecular characteristics of the thyroid peroxidase autoantigen. Autoimmunity (in press)

Banga, J.P., Barnett, P.S., Ewins, D.L., Page, M.J., and McGregor, A.M. (1990b): Mapping of autoantigenic epitopes on recombinant thyroid peroxidase fragments using the polymerase chain reaction. Autoimmunity (in press)

Banga, J.P., Mahadevan, D., Barton, G.J., Sutton, B.J., Saldanha, J.W., Odell, E., and McGregor, A.M. (1990c): Prediction of domain organization and secondary structure of thyroid peroxidase, a human autoantigen involved in destructive thyroiditis. FEBS Letts (in press)

Banga, J.P., Tomlinson, R.W.S., Doble, N., Odell, E., and McGregor, A.M. (1989a): Thyroid microsomal/thyroid peroxidase and autoantibodies show discrete patterns of cross-reactivity to myeloperoxidase, lactoperoxidase and horseradish peroxidase. Immunol. 67: 197-204

Banga, J.P., Barnett, P.S., Mahadevan, D., and McGregor, A.M. (1989b): Immune recognition of antigen and its relevance to autoimmune disease: recent advances at the molecular level. Eur. J. Clin. Invest. 19: 107-116

Banga, J.P., Barnett, P.S., Ewins, D.L., Page, M.J., and McGregor, A.M. (1989c): Autoantigenic epitopes on thyroid microsome/thyroid peroxidase antigen recognised by autoantibodies reside in the amino terminal, the carboxyl region and the central core of the molecule. Autoimmun. 5: 142-144

Barton, G.J. (1990): Protein multiple sequence alignment and flexible pattern matching. Meths. in Enzymol. 183: 403-428

Barton, G.J., and Sternberg, M.J.E. (1987): A strategy for the rapid alignment of protein sequences. J. Mol. Biol. 198: 327-337

Bashford, D., Chothia, C., and Lesk, A.M. (1987): Determinants of a protein fold: unique features of the globulin amino acid sequences. J. Mol. Biol. 196: 199-216

Biou, V., Gibrat, J.F., Lexin, J.M., Robson, B., and Garnier, J. (1988): Secondary structure prediction combination of three different methods. Prot. Eng. 2: 185-191

Chou, P.Y., and Fasman, G.D. (1978). Prediction of a secondary structure of proteins from their amino acid sequences. Adv. Enzymol. 47: 145-148

Cooke, R.M., Wilkinson, A.J., Baron, M., Pastore, A., Tappin, M.J., Campbell, I.D., Gregory, H., and Sheard, D. (1987): The solution structure of human epidermal growth factor. Nature 327: 339-341

Doble, N.D., Banga, J.P., Pope, R., Lalor, R., Kilduff, P., and McGregor, A.M. (1988): Autoantibodies to the thyroid microsomal/thyroid peroxidase antigen are polyclonal and directed to several distinct sites. Immunol. 64: 23-29

Garnier, J., Osguthorpe, D.L. and Robson, B. (1987): Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120: 97-120

Hopp, T.P., and Woods, K.R. (1981): Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. 78: 3824-3828

Inna, G., Piatigorsky, J., Norman, B., Slingsby, C., and Blundell, T. (1983): Gene and protein structure of a B-crystallin polypeptide in murine lens: relationship of exons and structural motifs. Nature 302: 310-315

Kimura, S., Hong, Y.S., Kotani, T., Ohtaki, S., and Kikkawa, F. (1989): Structure of the human thyroid peroxidase gene: comparison and relationship to the human thyroid peroxidase gene. Biochem. 28: 4481-4489

Libert, F., Ruel, J., Ludgate, M., Swillens, S., Alexander, N., Vassart, G., and Dinsart, C. (1987): Thyroperoxidase, an autoantigen with a mosaic structure made of nuclear and mitochondrial gene modules. EMBO J. 6: 4193-4196

Lim, V.I. (1974): Algorithms for prediction of -helical and B-structural regions in globular proteins. J. Mol. Biol. 88: 873-894

Ludgate, M., Mariotti, S., Libert, F., Dinsart, C., Piccolo, P., Santini, F., Ruf, J., Pinchera, A., and Vassart, G. (1989): Antibodies to human thyroid peroxidase in autoimmune thyroid disease: studies with a cloned recombinant complementary deoxyribonucleic acid epitope. J Clin. Endocrinol. Metab. 68: 1091-1096

Rose, G.D. (1978): Prediction of chain turns in globular proteins on a hydrophobic basis. Nature 272: 586-591

Rothbard, J.B., and Taylor, W.R. (1988): A sequence pattern common to T cell epitopes. EMBO J 7: 93-100

Williams, A.F. (1984): The immunoglobulin super family takes shape. Nature 308: 12

Wilmot, C.M., and Thornton, J.M. (1988): Analysis and prediction of the different types of B-turns in protein. J. Mol. Biol. 203: 221-232

Yokoyama, N., and Taurog, A. (1988): Porcine thyroid peroxidase: relationship between the native enzyme and an active, highly purified tryptic fragment. Mol. Endocrinol. 2: 838-849

Zvelebil, M.J.J., Barton, G.J., Taylor, W.J., and Sternberg, M.J.E. (1987): Prediction of protein secondary structure and active sites using the alignment of homologous sequences. J. Mol. Biol. 195: 957-961

Résumé

Le clonage de la TPO et son expression par recombinaison génétique a permis de progresser dans la caractérisation des épitopes auto-antigéniques reconnus par les auto-anticorps. La principale région immunogénique reconnue par les auto-anticorps a été localisée sur la partie C-terminale de la molécule. Des méthodes d'alignements de séquences et de prediction de structures secondaires combinées à la spectroscopie par dichroisme circulaire indiquent que la structure de la TPO est principalement en hélice alpha La connaissance supplémentaire des enchaînements des exons et des introns suggère un modele d'organisation en domaines de la molécules de TPO La possibilité future de positionner les épitopes auto-antigeniques sur un modèle structural tri-dimensionnel de la TPO permettra une meilleure compréhension de l'étiologie des maladies auto-immunes et permettra d'envisager des strategies thérapeutiques specifiques d'un antigene donne