Volume 266, number 1,2, 133–141 FEBS 08522 June 1990

Prediction of domain organisation and secondary structure of thyroid peroxidase, a human autoantigen involved in destructive thyroiditis

J. Paul Banga[1], Daruka Mahadevan[1], Geoffrey J. Barton[2], Brian J. Sutton[3], Jose W. Saldanha[2], Eddy Odell[4] and Alan M. McGregor[1]

[1]*Department of Medicine, King's College School of Medicine, Denmark Hill, London SE5 8RX,* [2]*Imperial Cancer Research Fund Laboratories, Biomedical Computing Unit, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX,* [3]*Division of Biomolecular Sciences, Biophysics Section, King's College London, 26–29 Drury Lane, London WC2B 5RL* and [4]*Department of Oral Medicine and Pathology, United Medical and Dental Schools, Guy's Hospital, Floor 28, Guys Tower, London Bridge, London SE1 9RT, UK*

Received 11 April 1990

Organ specific autoimmune diseases are relatively common immunological disorders in man which include thyroid autoimmune disease, insulin-dependent diabetes mellitus and myasthenia gravis. The target autoantigens in some of these diseases have recently been characterised. In thyroid autoimmune disease this includes the key enzyme, thyroid peroxidase (TPO), which is involved in the generation of thyroid hormone. Structural knowledge about autoantigens such as thyroid peroxidase will allow a greater understanding of the interaction between autoantigens and the aberrant immune response, and facilitate the development of strategies for antigen-specific therapeutic manipulation. We report here a prediction of the secondary structure of thyroid peroxidase, together with the results of circular dichroic spectroscopy of a homologous purified enzyme. A combination of 3 secondary structure prediction programs has been used, following multiple sequence alignment, and TPO has been found to consist mainly of α-helical conformation, with little β-sheet present. This structure prediction, together with knowledge of the exon-intron boundaries allows a model for the domain organisation of the TPO molecule to be proposed.

Thyroid peroxidase; Autoantigen; Myeloperoxidase; Circular dichroic spectroscopy; Secondary structure prediction

1. INTRODUCTION

Autoimmune lymphocytic thyroiditis is a relatively common condition affecting 5% of the general population, with a higher incidence in women than men [1]. The disease is characterised by the production of autoantibodies (aAbs) to a number of thyroid cell constituents including thyroglobulin (Tg) and the thyroid microsomal antigen. The thyroid microsomal antigen is likely to be the target autoantigen in destructive thyroid disease leading to thyroid failure [2]. We first showed this antigen to be a 105 kDa protein [3], which has been shown to be identical to thyroid peroxidase (TPO), the key enzyme involved in the generation of thyroid hormones from their 'pro-hormone' Tg [4]. The aAb response to TPO is known to be polyclonal, and several autoantigenic sites have been recognized [5,6]. Amongst the spectrum of aAbs detected are those which, being directed towards the active site of the molecule, inhibit enzymatic activity [5–7].

The complete amino acid sequence of TPO has recently been deduced from cDNA cloning [8]. This revealed the presence of two different forms of mRNA derived by alternative splicing of a single gene product [8]. The two forms of TPO, termed TPO-1 and TPO-2, have been fully sequenced and indicate that 57 residues are deleted in the shorter, TPO-2 form of the molecule [8,9]. The amino acid sequence of TPO exhibits a high degree of sequence similarity (42% identity) to myeloperoxidase (MPO) [10,11]. Furthermore, the structure of the TPO and MPO genes have recently been elucidated [12,13]. A comparison of the two gene sequences reveals that the exon-intron boundaries in both the peroxidases are highly conserved [13]. The close relationship between TPO and MPO is also manifest antigenically with some aAbs to TPO obtained from autoimmune thyroid disease patients cross-reacting with MPO [14]. 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 and allow the development of future strategies for immune intervention to prevent or reverse the course of the disease [15,16].

In order to elucidate secondary or tertiary structure, it is necessary to have access to milligram amounts of purified protein for structural analysis. We are able to purify by HPLC techniques a few hundred microgram amounts of enriched detergent-solubilized, intact TPO (approximate purity 60%) from human thyroid glands

Correspondence address: G.J. Barton, Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

for immunological studies [17], but it is currently not possible to obtain sufficient human thyroid glands for preparation of enough pure TPO for structural studies. Trypsinized TPO from porcine thyroid glands (pTPO) can be prepared in larger quantities, but this lacks approximately 90 residues at the amino- and carboxyl-terminals and also contains a trypsin cleaved site after residue 561 [18]. Such a sample was used for structural studies reported below. However, highly purified preparations of intact, water-soluble human MPO (hMPO) have been obtained for immunological investigations [14]. In view of the high sequence similarity between hMPO and human TPO (hTPO), we have used hMPO for circular dichroic (CD) spectroscopy and extrapolated the results to hTPO. In addition, multiple alignments [19] of the sequences of hTPO, pTPO and hMPO have been performed using the algorithm of Barton and Sternberg [20,21]. By applying 3 independent secondary structure prediction methods to the multiply-aligned peroxidase sequences, it has been possible to predict the secondary structure and the organisation of domains in TPO. By combining the CD spectral information on MPO with the secondary structure prediction of TPO, we show that the structure of TPO is mainly α-helical with little β-sheet, and is organised into distinct domains.

2. MATERIALS AND METHODS

2.1. *Preparation of MPO for CD analysis*

hMPO was isolated from neutrophils purified from blood packs of normal donor blood. The final step in the purification procedure included isoelectric focusing in Sephadex G-25 yielding MPO preparations of >90% purity [14]. Trypsinized pTPO (of purity >90%) was prepared as described [18], and was a gift from Professor A. Taurog.

2.2. *CD spectral analysis*

hMPO was dissolved at 0.1 mg/ml in 0.1 M Tris-HCl, pH 7.4; trypsinised pTPO was resuspended at 0.1 mg/ml in phosphate-buffered saline. CD spectra were recorded on a Jasco J600 CD spectrometer interfaced to an IBM PS/2 computer running commercial software. All spectra were run at room temperature in 0.05 cm 'strain-free' quartz cells and were corrected for solvent absorbance.

2.3. *Alignment of sequences and prediction techniques*

Multiple sequence alignment was performed automatically by the algorithm of Barton and Sternberg [20,21] using the AMPS package [19], whilst conservation values were defined using the protocol of Zvelebil et al. [22]. Turns were predicted using (i) the method of Wilmot and Thornton [23] (Program supplied by Dr C. Wilmot) and (ii) that of Rose [24]. α-Helix and β-strand regions were predicted using the algorithms of Lim [25], Chou and Fasman [26] and Robson [27] as programmed in the Leeds prediction suite (Dr. E. Eliopoulos, Department of Biophysics, University of Leeds, Leeds, UK). The results of the alignment and annotated predictions were formatted using the HOMED programme (Dr P.A. Stockwell, Department of Biochemistry, University of Otago, Otago, New Zealand). B-Cell defined epitopes were predicted using the Hopp and Woods profile [28] with the top 20 hydrophilic peaks considered for further analysis. T-Cell defined epitopes were predicted using the pattern matching technique of Taylor and Rothbard [29].

3. RESULTS AND DISCUSSION

3.1. *Analysis of CD spectroscopy*

The far ultraviolet CD spectrum of hMPO was recorded in aqueous solution and exhibits a moderately broad band with a minimum in the 200–210 nm range (Fig. 1). The values of α-helical content were calculated to be approximately 55–60% for hMPO (Table I), with the remainder of the protein in random coil conformation.

The CD spectrum of a trypsinized preparation of pTPO [18] was measured (not shown) and it also indicated a protein with only α-helical content (approximately 27-33%) and random coil conformation. The lower α-helical content in this preparation compared to hMPO may reflect the fact that the trypsinized pTPO preparation is a fragmented molecule.

3.2. *Analysis of hTPO, pTPO and hMPO sequence data*

3.2.1. Sequence alignment

The multiple alignment of hTPO, pTPO and hMPO is shown in Fig. 2. The high level of similarity between the sequences (>50.0 SD from the mean of 100 randomized sequence scores), strongly suggests that the proteins share a common ancestor, that their overall tertiary folds will be similar and also that the automatically obtained multiple alignment will be correct within common conserved secondary structural regions [20,21].

3.2.2. Combination of secondary structure prediction techniques

Secondary structure prediction methods generally give around 55–65% accuracy for a 3-state prediction of helix, strand and non-helix/non-strand (i.e. coil) structure. However, combined prediction methods can yield improvements in accuracy [30] as can combining predictions on accurately aligned sequences [22]. Accordingly, the 5 prediction methods were performed independently on each aligned sequence, and then the results combined into a consensus.

The fact that TPO and MPO exhibit a high degree of sequence [10,11] and considerable antigenic similarity [14] provides compelling evidence in favour of a similar secondary structure. Since the CD analysis indicates that hMPO is predominantly α-helical, the parameters for all helical proteins were applied in the Chou and Fasman prediction [26], whilst decision constants suitable for proteins of >50% α-helix and <20% β-structure were applied in the Robson prediction [27] ($DC_H = 100$, $DC_E = 50$).

The results of the two turn prediction algorithms, i.e. a total of 6 predictions at each position, were combined by expressing the number of predictions that define each residue as part of a turn as a fraction of the total possible number of predicted turns. Similarly, predic-

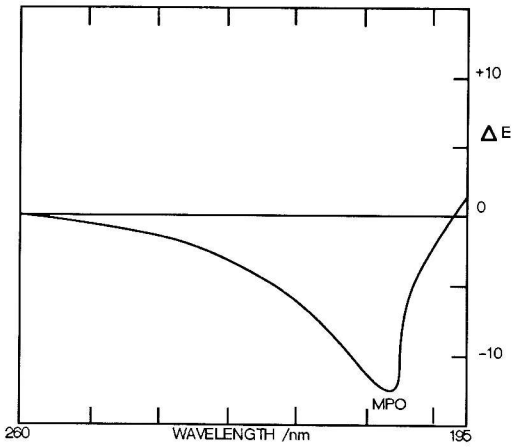


Fig. 1. Circular dichroism (CD) spectrum of hMPO (purity >90%) in the far ultraviolet region in aqueous solution (0.1 mg/ml in 0.1 M Tris-HCl, pH 7.4) in 0.05 cm pathlength. ΔE is the extinction coefficient.

tions of helix, strand and coil were combined for the 9 predictions (Lim [25], Chou and Fasman [26] and those of Robson [27] on each of the 3 sequences) at each position.

The consensus prediction shown in Fig. 2 was obtained by taking the defined state to be that with the largest fraction predicted. For example, at aligned position 74, Wilmot and Thornton [23] predict involvement in a turn for hTPO, pTPO and hMPO whilst Rose [24] predicts turn for hTPO and pTPO; a total of 5/6 predictions for turn. At the same position, Lim [25] predicts hMPO as β, Robson [27] predicts hTPO and pTPO as β, and hMPO as α, whilst Chou and Fasman [26] predict hTPO and hMPO as α. Thus, α and β are equally predicted 3/9 at this position, by default, coil is predicted 3/9, whilst turn is predicted 5/6. Accordingly the consensus prediction is assigned as turn for this position. Where a tie occurs, the ambiguity is recorded for all tied combinations of H,B,T predicted at that position (see legend to Fig. 2).

This approach to combining the methods highlights the strongly predicted α, β and turn regions, records ambiguity where it occurs, but leaves regions unassigned where only weakly predicted α or β are indicated. For the purposes of discussion, these unassigned regions will be referred to as 'coil'; however, this is not meant to imply lack of defined structure, but merely that α, β or turn are not strongly predicted in these regions.

Table I

The estimation of the α-helical content of human MPO

Wavelength (nm)	Extinction coefficient (ΔE) ($M^{-1} \cdot cm^{-1}$)	% helix[a]
222	5.4	59.4
225	4.8	55.2

[a] remaining secondary structure is coil

The overall result of this combined prediction for TPO is 31% α-helix, 9% β-sheet and 29% turn, with the remaining 31% being defined as 'coil'. In the region of homology with MPO (residues 1-741) the percentages are α-helix 51%, β-sheet 8%, turn 41% and the remainder as random coil. This compares well with the CD results showing α-helical conformation of 55-60% and the remainder as random coil.

3.2.3. Additional Information used to assist prediction

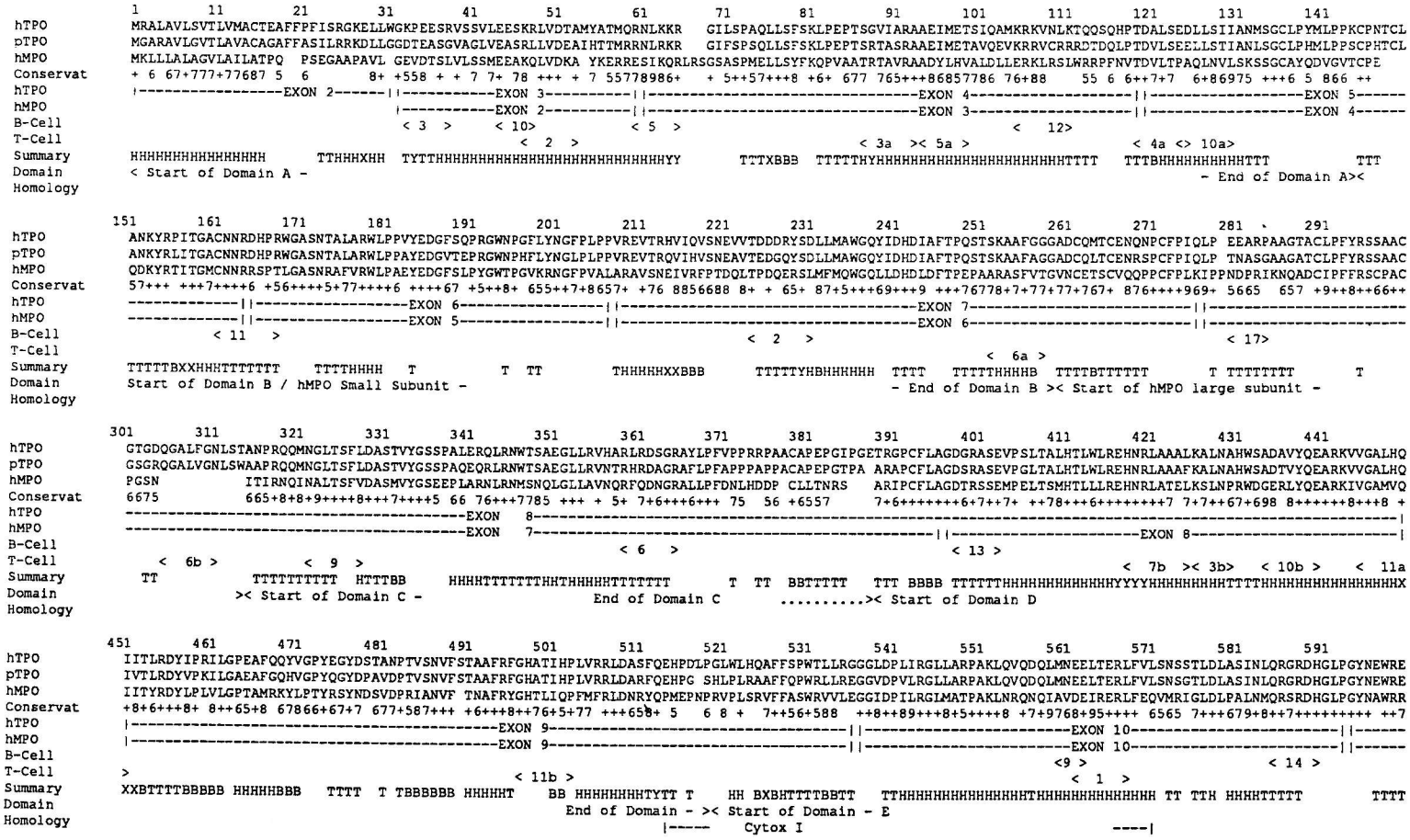
(i) Insertions/deletions. Further structural detail may be inferred from an alignment of two or more sequences by considering the residue conservation at each aligned position. In particular, studies of aligned protein families for which the three-dimensional structure of more than one member is available [31] show that insertions and deletions normally occur in surface loop regions between regular secondary structures. Analysis of the peroxidase alignment (Fig. 2) shows only 4 insertions between the human and porcine TPO (insertion of E at 388 and L at 518 of hTPO and insertion of GK at 896 and LPG at 903 of pTPO; see Table II). All 4 insertions are adjacent to Gly, Pro and charged amino acids which are frequently associated with turns or surface loop regions. The insertions are also adjacent to predicted turn and coil regions. These observations lend strong support to the prediction of loops at these positions. When the analysis is extended to hMPO a further 10 insertions are observed (Table II). Nine of these fall within turn/coil regions. The exception is the insertion at position 54 of h/pTPO which occurs in the middle of a strongly predicted helix.

(ii) Intron/exon boundaries. Recently the gene structures for hTPO and hMPO have been described [12,13]. The hTPO gene comprises 17 exons whilst the hMPO gene consists of 12 exons. When the exon-intron boundaries of these two genes are compared, exons 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 [13]. The use of sequence alignment and exon-intron boundaries between related proteins has previously been used to predict successfully the domain organization of β,γ-crystallins [32] and immunoglobulins [33].

The position of the exon boundaries is illustrated on Fig. 2 and summarised in Table II. Between positions 1 and 741, ten out of the eleven boundaries fall within one residue of a strongly predicted turn, or within a coil region.

In the C-terminal region of h/pTPO (742-943) the secondary structure of the EGF-like region (795-847) may be defined with confidence from the known three-dimensional structure of EGF as determined by NMR [34] (see below). The 13-14 boundary coincides with the random coil N-terminus of EGF, whilst the 14-15 boundary lies in a very short β-strand segment near the C-terminus of EGF. The remaining 3 boundaries (out-

1


hTPO
pTPO
hMPO
Conservat
B-Cell
T-Cell
Summary
Domain
Homology
Start of Domain A
End of Domain A
Start of Domain B / hMPO Small Subunit
End of Domain B
Start of hMPO large subunit
Start of Domain C
End of Domain C
Start of Domain D
Cytox I

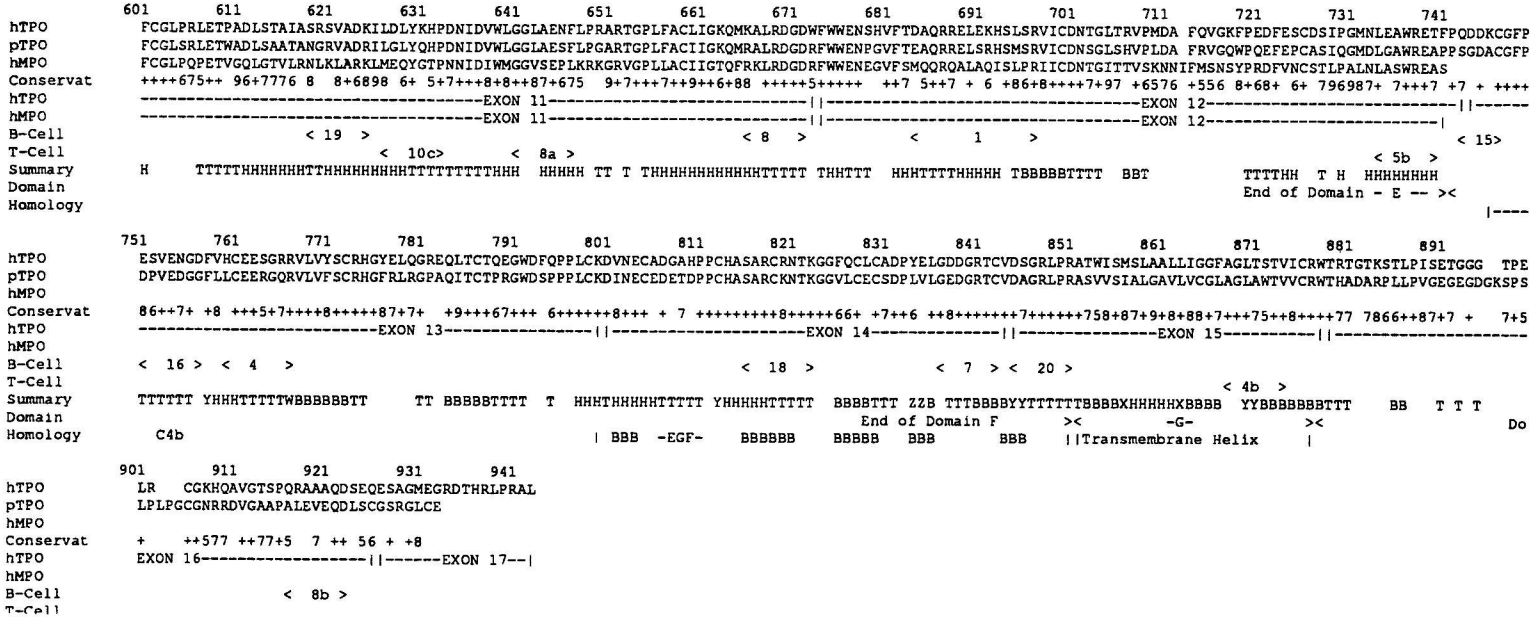


Fig. 2. Summary of structural features. Lines 1, 2 and 3 show the multiply aligned sequences of hTPO, pTPO and hMPO, respectively. Line 4 shows conservation values calculated according to the method of Zvelebil et al. [22] then multiplied by 10 to give a scale 0–10. Conservation values below 5 are not shown, whilst '+' indicates total identity at the position (also see Fig. 3). Lines 5 and 6 show the exon structure of hTPO and hMPO, respectively. Line 7 shows the location of predicted B-cell defined epitopes for hTPO by the method of Hopp and Woods [28], and the predicted epitopes are numbered according to their relative hydrophilicity with <1> being the most hydrophilic region. Line 8 shows the location of predicted T cell defined epitopes according to the patterns of Rothbard and Taylor [29]; numbering refers to the strength of the pattern match. 1 represents the highest score with equally scoring matches labelled a, b, c, etc. Line 9 shows the consensus secondary structure prediction (see text). 'H', 'B' and 'T' indicate strongly predicted helix, strand and turn regions. Where two or more states are equally predicted the ambiguity is recorded as follows: H/T = *Y*, H/B = *X*, B/T = *Z*, H/B/T = *W*. Line 10 shows the predicted domain structure for h/pTPO (see text) and line 11 (where appropriate) indicates the regions of sequence similarity shared with cytochrome *c* oxidase polypeptide I (Cytox I), C4b and EGF. The location of the β-strands in the EGF-like domain (indicated B) are those determined by NMR [34].

Table II

Insertion in protein(s)	Amino acids	Alignment position	Comments
h/pTPO	FF		
	FF	19–20	Adjacent Pro; turn predicted
h/pTPO	W		
	G	31	Adjacent Gly; turn predicted
h/pTPO	M		
	I	54	Predicted helix
hMPO	LRS	66–68	Adjacent Gly; turn predicted
h/pTPO	TCL		
	TCL	148–150	Turn predicted
hMPO	P	279	Adjacent Pro; turn predicted
h/pTPO	QGALFGNLST		
	QGALVGNLSW	305–314	Turn predicted
h/pTPO	A		
	A	378	Adjacent Pro; turn predicted
h/pTPO	PGE		
	PA	386–388	Turn predicted
h/pTPO	S		
	S	490	Coil predicted
hTPO	L	518	Gly/Pro region and predicted turn
hMPO	I	713	Coil predicted
pTPO	GK	896–897	Gly/Pro region and predicted turn
pTPO	LPG	903–904	Gly/Pro region and predicted turn

side the EGF homology region) all lie in predicted turn, or coil regions.

(iii) Conservation of Cys residues. TPO and MPO are both known to contain disulphide bridges [14,18]. Of the 16 Cys residues conserved between hTPO and pTPO in the region 1–741 only those at residue 15 and residue 149 are not also conserved in hMPO. The conserved Cys residues are not uniformly distributed through the protein sequence. Three occur within a 26 amino acid stretch (136–161), five in a 39 residue stretch (262–300) and two within 15 amino acids (379–393) of each other. The remaining 4 conserved Cys positions are distributed through the region from 602 to 741, at 602, 659, 700 and 726.

This extensive conservation of cysteine residues strongly supports the contention that TPO and MPO are structurally homologous throughout the region of TPO that corresponds to the two subunits of MPO.

(iv) General residue conservation. Although the 3 proteins share a large number of identical amino acids (hTPO versus pTPO = 70% identity; hTPO versus hMPO as aligned in Fig. 2 = 50% identity), by considering a smoothed profile of conservation values (Fig. 3), regions where identities and conservative substitutions predominate may be identified. These suggest the residues most likely to be involved in structural and/or functional roles common to all three molecules. Conversely, regions of low conservation can indicate variable regions that may correspond to surface loops, or domains that do not exhibit similar functions in all three proteins. With one exception, at position 115 where a turn is also strongly predicted, the most significant minima shown on Fig. 3 correspond to the positions of insertions in one or more proteins. The regions bounded by these minima are discussed in more detail below.

3.2.4. Homology with EGF, complement C4b and cytochrome *c* oxidase

Database scans have revealed sequence similarity between TPO and EGF [10] in the region 799–847 (Fig. 2). The high sequence similarity between the two molecules in this region allows the structure of TPO to be predicted confidently by analogy with the NMR structure of EGF [34]. In particular this is supported by the absolute conservation of the six Cys residues that are disulphide-bonded in the EGF molecule. The structure of EGF consists of two β-strands and turns held together by a cluster of three disulphide bonds [34].

The EGF-like domain is preceded by a region (744–798) that has similarity to C4b [10] and is followed by a predicted transmembrane helix (852–876). These 3 regions correspond to exons 13, 14 and 15, respectively. hMPO and h/pTPO also share a region of strong similarity with cytochrome *c* oxidase polypeptide I (514–571). This does not correspond to a single exon, but spans the boundary between exons 9 and 10 and includes only a part of each.

3.3. *hMPO subunit organisation*

hMPO is synthesised as a large precursor polypeptide which, during intracellular transport, is processed to yield mature heavy and light chain subunits of approximately 65 and 14 kDa, respectively [35]. The tetrameric complex of functional hMPO is formed by cleavage of the precursor hMPO at 144–145 (Thr-Cys) and 259–260

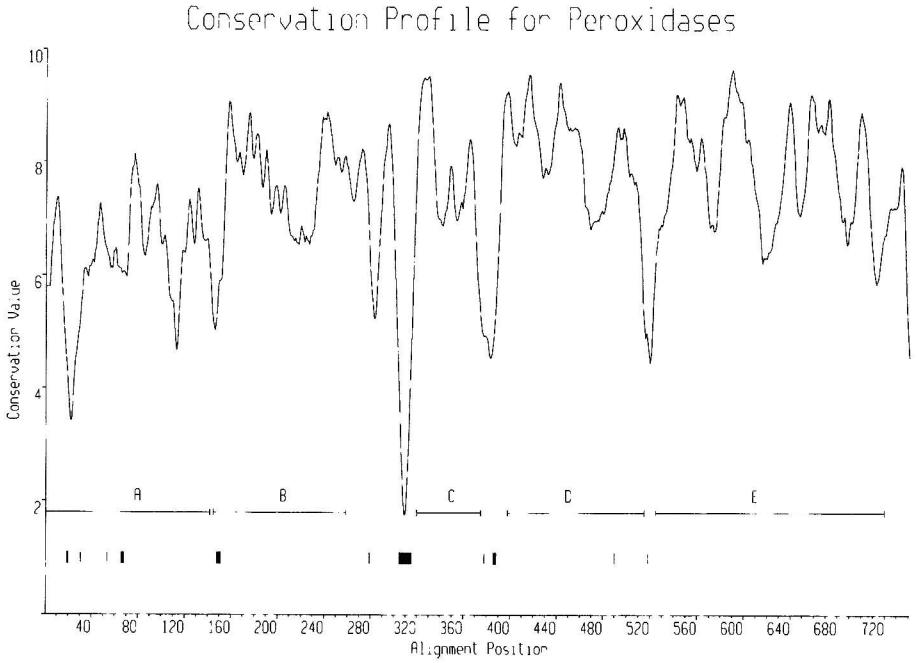


Fig. 3. Conservation profile for region common to TPO and MPO. The conservation profile was calculated by smoothing the conservation values shown in Fig. 2. using a moving average with a window of 11 amino acids, followed by a further smoothing process using a window of 3 amino acids. The putative domain positions are illustrated (A–E) together with the location of all insertions in the alignment (vertical bars). High values indicate conservation; the minima indicate regions of greatest sequence variation.

(Gly-Val) to yield two heavy (260–741) and two light (145–259) chains. Both cleavage points fall within strongly predicted turn regions as might be expected for a loop or coil accessible to proteolysis.

3.4. *Prediction of domain locations for the peroxidases*

Proteins of the length of the region common to all three peroxidases (1–741) are often divided into several structural domains. The subunit structure of hMPO clearly suggests at least two distinct domains that by analogy may exist in TPO. Supporting evidence for this suggestion is the presence of an insertion in TPO close to the start of the hMPO small subunit and a glycine-rich turn region at the end (Fig. 2). Inspection of the conservation profile (Fig. 3) shows the hMPO small subunit region to be one of the longest continuous regions of conservative substitutions. Additionally, both subunits start with stretches of sequence for which little regular secondary structure (α-helix and β-strand) is predicted. This may indicate that in TPO these are flexible domain linking regions.

The N-terminal region that precedes the small subunit of hMPO (1–144) consists of 6 predicted helices. These residues 1–144 may form an independently folding domain in h/pTPO. Fig. 3 shows this putative domain (Domain A) to exhibit lower overall conservation over its length than is seen over the rest of the molecule. This observation is consistent with the knowledge that this region is cleaved off in the mature hMPO and so, presumably is not subject to the same pressure to conserve precise detailed tertiary structure as the rest of the molecule.

The hMPO small subunit spans 114 (145–259) residues in the alignment with only 5 short helices strongly predicted. However, the subunit corresponds to a region of continuous high conservation (Fig. 3) as would be expected for a domain sharing common properties of function or structure. We have labelled the hMPO small subunit region, Domain B.

The region corresponding to the hMPO large subunit consists of 481 aligned positions (260–741) which is atypically large for a single structural protein unit. The secondary structure predictions, conservation analysis and homology information suggest that the hMPO large subunit region may be subdivided into up to 3 distinct sub-regions (see Fig. 3). There are two long segments of continuous high conservation. The segment 398–520 includes the highly conserved region, suggested by Kimura and Ikeda-Saito [11] to contain the proximal histidine at 411 (indicated D on Fig.3). This is followed by a second long segment which continues to position 715 (indicated E on Fig. 3). This region includes the stretch of sequence homologous to cytochrome *c* oxidase polypeptide I [10]. It also includes the longest stretch of sequence that is completely conserved in all 3 sequences which occurs at the exon 10/11 boundary (positions 588–599) and includes a conserved His at 590, strongly suggesting that this His is involved in the function of all 3 proteins. In TPO-2 this conserved His-590 is spliced out in exon 10 and it is possible that

Table III

Exon/intron boundaries in hTPO and hMPO

hTPO	hMPO	
2-3	1-2	Coil/turn region and deletion in hMPO
3-4	2-3	End of helix-turn region - close to insertion in hMPO
4-5	3-4	Strongly predicted turn
5-6	4-5	Strongly predicted turn
6-7	5-6	Strongly predicted turn/coil
7-8	6-7	Coil - insertion in hMPO
-	7-8	Strongly predicted turn
8-9	8-9	Ambiguous helix/β prediction
9-10	9-10	Strong turn/coil prediction
10-11	10-11	Strong turn/coil prediction
11-12	11-12	Strong turn/coil prediction
12-13	-	Coil
13-14	-	Coil - (from homology to EGF)
14-15	-	β - (from homology to EGF)
15-16	-	Turn
16-17	-	Turn

its removal may lead to an enzymatically inactive molecule.

A further highly conserved segment precedes these domains D and E, spanning positions 315–380 and which we have labelled as Domain C (Fig. 3). Between the end of Domain B and the beginning of C there is a region of 44 residues (260–314) which shows no regular secondary structure, includes the longest insertion between TPO and MPO, and is unusually rich in Cys residues. At the end of Domain C there is a further unusual stretch of sequence (369–386) in which 9 out of the 18 residues are proline. We suggest that this segment, at the end of which are found deletions in pTPO and hMPO relative to hTPO, marks the boundary with Domain D. However, evidence for the boundary between Domains D and E cannot be considered conclusive. The breakdown in homology in this region around residue 518 which gives rise to the substantial minimum in the conservation profile in Fig. 3 could be due to an extensive surface loop feature.

The C-terminal region of TPO from residue 742 includes the segments that bear homology with the C4b and EGF (Fig. 2). These segments which correspond to exons 13 and 14, may well correspond to independently folded domains. They are followed by the predicted transmembrane helix and a short intracellular region for which little or no regular secondary structure is predicted.

3.5. *Prediction of antibody and T cell defined epitopes*

The TPO molecule is known to be the target autoantigen recognized by aAbs from patients with autoimmune thyroid disease. We have used the Hopp and Woods hydrophilicity profiles [27] to predict sequences of amino acid residues in the TPO molecule which may be recognized by antibodies. Since antibody recognition by B-lymphocytes invariably requires T cell help, we have also predicted sequences of amino acid residues that may serve as putative recognition sites for T-lymphocyte populations using the patterns described by Rothbard and Taylor [29]. (Fig. 2). Distinct T-lymphocyte populations specific for target autoantigens have been implicated in the pathogenesis of experimentally induced and human autoimmune diseases [15,16,36]. Furthermore, the ability to predict T cell recognition sites for human autoimmune diseases is important in defining the specificity of cloned human T cells using synthetic polypeptides [37]. The B cell defined epitopes were predicted as the top 20 hydrophilic peaks (Fig. 1). The two most hydrophilic peaks are found at positions 684–697 and 223–230. Interestingly, using recombinant TPO preparations which encompass different segments of the TPO molecule, our recent data have shown that polyclonal antisera to TPO and patient's autoantibodies to TPO recognize recombinant TPO preparations containing residues 657–767 and 145–250 [38]. These two recombinant preparations harbour the two most hydrophilic peaks mentioned above. In contrast to B cell epitopes, no empirical data are available on the T cell epitopes of the TPO molecule [39]. However, the amino acid residues comprising the T cell sites can now be empirically determined using synthetic peptides for stimulation of thyroid gland infiltrating or peripheral blood lymphocytes from thyroid disease patients undergoing active disease.

4. CONCLUSIONS

In summary, multiple sequence alignment and secondary structure prediction methods used in conjunction with CD spectroscopy of a homologous enzyme indicate that TPO structure is principally α-helical. These data taken together with a knowledge of the exon-intron boundaries suggest a model for the domain

organization of the hTPO molecule. This domain model will allow a systematic approach to the design of oligonucleotides for PCR amplification of discrete regions of cDNA to generate recombinant TPO preparations for immunological [38] and crystallographic studies.

Acknowledgments: We thank Drs J. Fox and C. Rawlings for advice and support, Dr A. Drake and Mr J. Hoardley for help with the CD spectroscopy. Many thanks to Dr J. Rothbard for access to the motifs programme for predicting T cell epitopes. We thank Professor A. Taurog for the gift of purified, trypsinized preparation of porcine TPO. We are grateful to Mrs. J. De Groote for the expert preparation of the manuscript.

REFERENCES

[1] Weetman, A.P. and McGregor, A.M. (1984) Endocrine Rev. 5, 309-361.
[2] Khoury, E.L., Hammond, L., Bottazzo, G.F. and Doniach, D. (1981) Clin. Exp. Immunol. 45, 316-328.
[3] Banga, J.P., Pryce, G., Hammond, L. and Roitt, I.M. (1985) Mol. Immunol. 22, 629-642.
[4] Czarnocka, B., Ruff, J., Ferrand, M., Carayan, P. and Lissitzky, S. (1985) FEBS Lett. 190, 147-151.
[5] Doble, N., Banga, J.P., Pope, R., Lalor, E., Kilduff, P. and McGregor, A.M. (1988) Immunol. 64, 23-29.
[6] Kohno, Y., Himaya, Y., Shimojo, N., Niimi, H., Nakajima, H. and Hosoya, T. (1986) Clin. Exp. Immunol. 64, 534.
[7] Yokoyama, N., Taurog, A. and Klee, G.G. (1989) J. Clin. Endocrinol. Metab. 68, 766-773.
[8] Kimura, S., Kotani, T., McBride, O.M., Umeki, K., Hirai, K., Nakayama, T. and Ohtaki, S. (1987) Proc. Natl. Acad. Sci. USA 84, 5555-5559.
[9] Barnett, P.S., Banga, J.P., Watkins, J., Huang, G.C., Gluckman, D., Page, M.J. and McGregor, A.M. (1990) Nucleic Acids Res. 18, 670.
[10] Libert, F., Ruel, J., Ludgate, M., Swillens, S., Alexander, N., Vassart, G. and Dinsart, C. (1987) EMBO J. 6, 4193-4196.
[11] Kimura, S. and Ikeda-Saito, M. (1988) Proteins: Structure, Function and Genetics 3, 113-120.
[12] Morishita, K., Tsuchiya, M., Asano, S., Kaziro, Y. and Nagata, S. (1987) J. Biol. Chem. 262, 15208-15215.
[13] Kimura, S., Hong, Y., Kotani, T., Ohtaki, S. and Kikkawa, F. (1989) Biochemistry 28, 4481-4489.
[14] Banga, J.P., Tomlinson, R.W.S., Doble, N., Odell, E. and McGregor, A.M. (1989) Immunology 67, 197-204.
[15] Banga, J.P., Barnett, P.S., Mahadevan, D. and McGregor, A.M. (1989) Eur. J. Clin. Invest. 19, 107-116.
[16] Wraith, D.C., McDevitt, H.O., Steinman, L. and Acha-Orbea, H. (1989) Cell. 57, 709-715.
[17] Tomlinson, R.W.S., Banga, J.P. and McGregor, A.M. (1989) J. Endocrinol. 121, 117 (Abstr.).
[18] Yokoyama, N. and Taurog, A. (1988) Mol. Endocrinol. 2, 838-844.
[19] Barton, G.J. (1990) Methods Enzymol. 18, 403-428.
[20] Barton, G.J. and Sternberg, M.J.E. (1987) J. Mol. Biol. 198, 327-337.
[21] Barton, G.J. and Sternberg, M.J.E. (1987) Protein Eng. 1, 89-94.
[22] Zvelebil, M.J.J., Barton, G.J., Taylor, W.R. and Sternberg, M.J.E. (1987) J. Mol. Biol. 195, 957-961.
[23] Wilmot, C.M. and Thornton, J.M. (1988) J. Mol. Biol. 203, 221-232.
[24] Rose, G.D. (1987) Nature 272, 586-591.
[25] Lim, V.I. (1974) J. Mol. Biol. 88, 873-878.
[26] Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymol. 47, 145-148.
[27] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97-120.
[28] Hopp, T.P. and Woods, K.R. (1987) Proc. Natl. Acad. Sci. USA 78, 3824-3828.
[29] Rothbard, J.B. and Taylor, W.R. (1988) EMBO J. 7, 93-100.
[30] Riou, V., Gibrat, J.F., Levin, J.M., Robson, B. and Garnier (1988) Protein Eng. 2, 185-191.
[31] Bashford, D., Chothia, C. and Lesk, A.M. (1987) J. Mol. Biol. 196, 199-216.
[32] Inana, G., Piatigorsky, J., Norman, B,, Shingsby, C. and Blundell, T.L. (1983) Nature 302, 310-315.
[33] Williams, A.F. (1984) Nature 308, 12-14.
[34] Cooke, R.M., Wilkinson, A.J., Baron, M., Pastore, A., Tippin, M.J., Campbell, I.D., Gregory, H. and Sherad, B. (1987) Nature 327, 339-341.
[35] Morishita, K., Kubota, N., Asano, S., Kaziro, Y. and Nagata, S. (1987) J. Biol. Chem. 262, 3844-3851.
[36] Heber-Katz, E. and Acha-Orbea, H. (1989) Immunol. Today 10, 164-169.
[37] Harcourt, G.C., Sommer, N., Rothbard, J., Wilcox, N.A. and Newsom-Davis, J. (1988) J. Clin. Invest. 82, 1894-1900.
[38] Banga, J.P., Barnett, P.S., Ewins, D.C., Page, M.J. and McGregor, A.M. (1990) Autoimmunity in press.
[39] Banga, J.P., Barnett, P.S. and McGregor, A.M. (1990) Autoimmunity in press.