[Frontiers In Bioscience, Scholar, 10, 326-336, March 1, 2018]

# Thyroid hormone-interacting cell and plasma proteins share a common motif

**Salvatore Benvenga[1,2,3], Fabrizio Guarneri[4]**

*[1]Department of Clinical and Experimental Medicine - Endocrinology, University of Messina, via Consolare Valeria - Gazzi, 98125 Messina, Italy, [2]Master Program on Childhood, Adolescent and Women's Endocrine Health, University of Messina, via Consolare Valeria - Gazzi, 98125 Messina, Italy, [3]Interdepartmental Program on Molecular and Clinical Endocrinology and Women's Endocrine Health, AOU Policlinico G. Martino, via Consolare Valeria - Gazzi, 98125 Messina, Italy, [4]Department of Clinical and Experimental Medicine - Dermatology, University of Messina, via Consolare Valeria - Gazzi, 98125 Messina, Italy*

## TABLE OF CONTENTS



## 1. ABSTRACT

We hypothesized that a thyroid hormone (TH)-binding consensus sequence, which is shared by human and animal TH plasma carriers (THPC), might also be shared by cell surface TH transporters (CMTTH) and TH nuclear receptors (THR). We generated the consensus for CMTTH or THR from 8,691 or 624 sequences. In the 49 position-long THPC consensus, eight positions were occupied by very highly conserved (>50% of sequences) and 11 by highly conserved (>33-50% of sequences) groups of residues. Matches between very highly conserved residues of the same group were seven, five or nine when comparing CMTTH vs THPC, THR vs THPC or THR vs CMTTH. Matches between highly conserved residues of the same group were found at one position when comparing CMTTH vs THPC. Noteworthy, the 5-residue TH-binding motif (**Y**,**L/I/M**,X,X,**V/L/I**) originally detected in a few THPC and then confirmed in the 426 THPC (**Y/F/W**,**L/V/I/M**,L/V/M/I,I/v/i/m,**V/L/I/M**) at positions 22-26, was also confirmed in the total 9,741 sequences (**W/F/Y**,**L/I/V/M**,**I/V/M/L**,P,**L/V/I/M**). In conclusion, proteins so genetically and functionally diverse share TH binding because they share a homologous region that remains conserved phylogenetically.

## 2. INTRODUCTION

One of us showed that minor thyroid hormone (TH) plasma carriers of the non-exchangeable (apo B-100) and exchangeable (apo A-I, A-II, A-IV, C-I, C-II, C-III, and apoE) apolipoprotein gene families share local amino acid sequence homology with the major TH plasma carriers (thyroxine-binding globulin (TBG)), prealbumin or transthyretin (TTR) and serum albumin (HSA)) (1). Next, he showed that a 5-residue hydrophobic TH binding motif (**Y**, **L**/**I**/**M**, **X**, **X**,

**V**/**L**/**I**) was evolutionarily conserved in sequences of apolipoproteins known until 1997 (2). Because many other apolipoproteins and nonapolipoprotein TH plasma carriers (THPC) were sequenced subsequently, we wished to verify whether the same phylogenetic conservation of the motif occurred in a much larger number of all such TH binders (3). Importantly, we found that the motif is much larger, and that it is absent in proteins with no known TH interaction property (3). The larger motif is "E/D/Q/N, 0-1X, e/q/d/n, 0-1X, E/D/Q/N, 0-3X, s/t, 0-23X, F/W/Y, 0-1X, e/q/d/n, X, v/l/i/m, X, 0-34X, **Y/F/W**, **L/V/I/M**, **I/v/m/i**, **X**, **V/L/I/M**, 0-1X, e/d/q/n, X, X, X, X, X, I/v/i/m, X, X, I/m/v/i, 0-10X, S/T, q/e/n/d, L/V/I/M, x, n/q/e/d, 0-1X, I/v/i/m", with the 5-residue core sequence typed boldface. Consistently with our previous paper (3), in the above notation, "X" indicates that any amino acid can be present at a given position, uppercase letters for "primary anchors" (that is, residues conserved in more than 50% of the aligned sequences), and lowercase letters for "secondary anchors" (that is, residues conserved in more than 33% and up to 50% of the aligned sequences). Throughout this paper, we will refer to residues conserved in more than 50% of the aligned sequences as being very highly conserved or having the highest degree of conservation. We will refer to residues conserved in more than 33% and up to 50% of the aligned sequences as highly conserved or having the next best degree of conservation. Residues conserved in 33% or less of the aligned sequences will be referred to as modestly conserved.

Now, we wished to test the hypothesis that the said large motif was present into two types of physiologically important proteins that bind thyroid hormones but that are localized in cells (4): the TH cell membrane transporters (CMTTH) and the TH nuclear receptors (THR).

## 3. MATERIALS AND METHODS

To maintain consistency with our previous work (3), we probed the motif "E/Q/D/N, p, e/q/d, 0-1X, e/d/q, 0-3X, s/t, 0-14X, W/Y/F, d/e/n/q, 0-1X, a/g, I/v/i, I/m/v/i, 0-34X, **Y**/**F**/**W**, **L**/**V**/**I**/**M**, X, X, **V**/**L**/**I**/**M**, q/e/d/n, e/q/d/n, X, I/v/m, v/i/l/m, X, X, v/i/l/m, 4-10X, S/T, X, L/I/M/V, 2-3X, I/v/i/m" on CMTTH (OATPs, MCT8, MCT10, LAT1, LAT2, NCTP, MDR1) and THR (THR alpha1, beta1, beta2). THR alpha2 was omitted *a priori*, in that it contains a 122-residue C-terminus that replaces a region in THR alpha 1 ligand binding domain, thus rendering THR alpha 2 unable to bind TH (4).

In line with previous criteria (1, 3), the motif is made by groups of amino acids conserved, at each position, in more than 33% of the aligned sequences; at each position, amino acids are listed in decreasing order of frequency. Details on the notation used were given in the Introduction of the present paper. The amino acid sequences of CMTTH and THR from *Homo sapiens* and animal species were retrieved from the Entrez Protein database, omitting replicated entries and fragments. Search for the aforementioned motif in these sequences was performed with the MotiFinder software (5), as in the previous paper (3).

Groups of amino acids were: A/G (Ala/Gly), D/E/N/Q (Asp/Glu/Asn/Gln), F/Y/W (Phe/Tyr/Trp), H/K/R (His/Lys/Arg), I/L/M/V (Ile/Leu/Met/Val), S/T (Ser/Thr) (1-3).

In addition to conservation of the motif *per se*, we were also interested in detecting some consistency. Thus, regardless of the being plasmatic or cellular, we sought to ascertain whether conservation of the motif within proteins having preferential interaction with T4 was greater compared with proteins having preferential interaction with T3, and *vice versa*. The binding affinity of the different cell proteins for T4 or T3 was not always clearly defined in literature. Accordingly, our comparison was limited to the cell proteins which are experimentally and unambiguously classified in literature (4, 6-11) as trasporting/binding T4 more than T3 (OATP1C1) or trasporting/binding T3 more than T4 (MCT8, MCT10, THR alpha1, beta1, beta2). Differences between proportions were analyzed by the chi-square test or Fisher's exact test, as appropriate.

## 4. RESULTS

Consensus motifs for each group of TH binding proteins considered (apolipoproteins, plasmatic non-apolipoproteins, all plasmatic proteins, membrane transporters, nuclear receptors, all cell proteins) and overall consensus for the grand total of 9,741 sequences (from 450 species) are shown in Figure 1. Residues having a very high degree of conservation appear boldface white against a black background in all figures; their corresponding position is also typed boldface in the text. Residues highly conserved appear boldface black against a gray background. Residues modestly conserved appear in regular print against no background. The consensus motifs of the different categories of proteins are summarized in Table 1. To independently verify the statistical significance of our findings, the protein-protein BLAST (Basic Local Alignment Search Tool) software, available at https://blast.ncbi.nlm.nih.gov, was used to compare the motif-containing sequences of human proteins with those of their animal counterparts. The E values reported by the software were all highly significant, as their order of magnitude ranged between $10^{-15}$ and $10^{-62}$.

### 4.1. Plasma carriers

Consensus motifs for the 220 apolipoproteins (from 69 species), for the 206 nonapolipoproteins (from 54 species) and the resulting total of 426 plasmatic TH

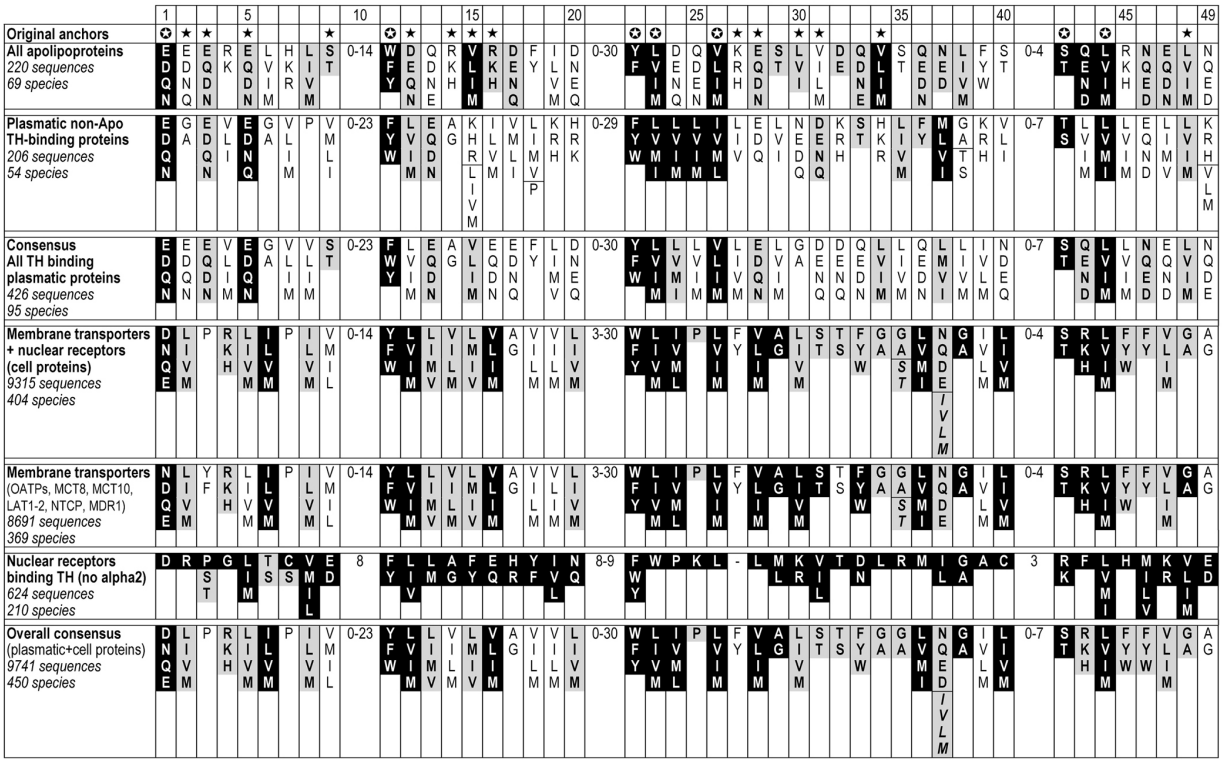


**Figure 1.** Consensus sequences for TH-binding plasmatic and cell proteins. Based on frequency in a given position, groups of similar amino acids are written in white boldface on black background (frequency >50%) or in black boldface on gray background (frequency 34% to 50%). When the latter case occurs for two groups in the same position but their frequencies are different, amino acids belonging to the less frequent group are written in italics. In each position of the overall consensus, amino acids belonging to the same group are shown from top to bottom in decreasing order of frequency.

carriers (from 95 species) are from ref. 3. Conservation of the TH binding motif in these 426 sequences is very high at eight positions (that is, columns 1, 5, 11, 22, 23, 26, 42 and 44). Highly conserved positions in the same 426 sequences are 11 (columns 3, 9, 13, 15, 24, 28, 34, 37, 43, 46 and 48).

### 4.2. Cell membrane transporters of thyroid hormones (CMTTH) and comparison with plasma carriers (THPC)

Figure 1 shows that conservation of the TH binding motif in the 8,691 CMTTH sequences (from 369 species) is very high at 21 positions, a number almost 3-fold greater than the eight positions in the THPC (see above, heading Plasma carriers). Of these 21 positions, seven (columns no. 1, 11, 22, 23, 26, 42 and 44) are also those very highly conserved in the 426 THPC and that share the same group of amino acids. Noteworthy, the hierarchy of representation of conservative residues within each group of residues varies, except for the two C-terminal of the seven positions. Indeed, at column 42, Ser occurs more frequently than Thr in both THPC and CMTTH, as does Leu>Val>Ile>Met at column 44. The highly conserved positions in the 8,691 CMTTH sequences are 14, only one of which (column no. 15) matches the consensus of the 426 THPC. Noteworthy, at position 14, Val is the most frequent residue in these THPC but the rarest in the CMTTH.

We wished to ascertain which THPC class (*viz*, apolipoprotein or the nonapolipoprotein) was more similar to the CMTTH. To this end, we considered the total number of matches in the two top degrees of conservation (*viz*, very highly conserved or highly conserved residues in CMTTH had to match either very highly conserved or highly conserved residues in THPC, and *vice versa*) (Figure 1). We also considered the number of intervening residues at the buffer positions (columns 10, 21 and 41). As shown in Figure 1, there are a total of 16 such matches with apolipoproteins (columns 1, 3, 5, 9, 11, 15, 22, 23, 26, 28, 34, 42, 43, 44, 46, 48) and 13 with the nonapolipoproteins (columns 1, 3, 5, 11, 13, 22, 23, 24, 26, 37, 42, 44, 48). The number of buffer residues of CMTTH is more similar to that of apolipoproteins (0-14 *vs* 0-14 at column 10, 3-30 *vs* 0-30 at column 21, and 0-4 *vs* 0-4 at column 41) than that of nonapolipoproteins (0-14 *vs* 0-23, 3-30 *vs* 0-29, and 0-4 *vs* 0-7, respectively) (Figure 1).

### 4.3. Thyroid hormone nuclear receptors (THR) and comparison with plasma carriers (THPC)

In the 624 sequences (from 210 species) of the three THR, all positions but one (column 6) are

**Table 1.** Matches of very highly (more than 50% of sequences) and highly (more than 33% and up to 50%) conserved amino acid groups in the consensus of thyroid hormone plasma carriers (apolipoproteins and nonapolipoproteins), cell surface thyroid hormone transporters and thyroid hormone nuclear receptors

| Protein group | THPC 8 - 11 | Apolipoproteins 10 - 18 | Nonapolipoproteins 11 - 8 | CMTTH 21 – 14 [1] | THR 44 - 3 [2] |
|---|---|---|---|---|---|
| THPC 8 - 11 | --- | **7** (1, 11, 22, 23, 26, 42, 44) | **8** (1, 5, 11, 22, 23, 26, 42, 44) | **7** (1, 11, 22, 23, 26, 42, 44) | **5** (1, 11, 22, 26, 44) |
| | --- | **5** (3, 9, 28, 46, 48) | **3** (3, 13, 48) | **1** (15) | **0** |
| | --- | **4** (5, 15, 34, 43) | **2** (24, 37) | **1** (24) | **3** (34, 37, 48) |
| Apolipoproteins 10 - 18 | **7** (1, 11, 22, 23, 26, 42, 44) | --- | **7** (1, 11, 22, 23, 26, 42, 44) | **7** (1, 11, 22, 23, 26, 42, 44) | **6** (1, 11, 22, 26, 34, 44) |
| | **5** (3, 9, 28, 46, 48) | --- | **2** (3, 48) | **2** (8, 37) | **0** |
| | **4** (5, 15, 34, 43) | --- | **1** (5) | **2** (15, 30) | **3** (8, 33, 48) |
| Nonapolipoproteins 11 - 8 | **8** (1, 5, 11, 22, 23, 26, 42, 44) | **7** (1, 11, 22, 23, 26, 42, 44) | --- | **8** (1, 11, 22, 23, 24, 26, 42, 44) | **6** (1, 11, 22, 26, 37, 44) |
| | **3** (3, 13, 48) | **2** (3, 48) | --- | **0** | **0** |
| | **2** (24, 37) | **1** (5) | --- | **1** (12) | **2** (12, 48) |
| CMTTH 21 - 14 [1] | **7** (1, 11, 22, 23, 26, 42, 44) | **7** (1, 11, 22, 23, 26, 42, 44) | **8** (1, 11, 22, 23, 24, 26, 42, 44) | --- | **9** (1, 11, 12, 22, 26, 28, 36, 38, 44) |
| | **1** (15) | **2** (8, 37) | **0** | --- | **0** |
| | **1** (24) | **2** (15, 30) | **1** (12) | --- | **2** (8, 13) |
| THR 44 - 3 [2] | **5** (1, 11, 22, 26, 44) | **6** (1, 11, 22, 26, 34, 44) | **6** (1, 11, 22, 26, 37, 44) | **9** (1, 11, 12, 22, 26, 28, 36, 38, 44) | --- |
| | **0** | **0** | **0** | **0** | --- |
| | **3** (34, 37, 48) | **3** (8, 33, 48) | **2** (12, 48) | **2** (8, 13) | --- |

[1]: column 35 contains two highly conserved amino acid groups [2]: columns 3 and 7 contain a very highly and a highly conserved amino acid group, CMTTH: Cell surface thyroid hormone transporters. THPC: thyroid hormone plasma carriers. THR: Thyroid hormone nuclear receptors. THPC consists of apolipoproteins and nonapolipoproteins. For each combination, number and (in brackets) position of the matches within the 49 position scheme (see Figure 1) are given. Matches between very highly conserved amino acid groups are written in white in black cells, matches between highly conserved amino acid groups are written in black in gray cells, matches of a group which is very highly conserved in one series of proteins and highly conserved in the other are written in black in white cells. In the row and column headings, the number of very highly or highly conserved amino acid groups in each group of proteins (see also Figure 1) are typed in white boldface on black background and black boldface on gray background, respectively.

very highly conserved; in addition, columns 3 and 7 have also a highly conserved group (Figure 1). Of the 43 ungapped positions, those shared with the THPC are five (no. 1, 11, 22, 26 and 44), while those shared with the CMTTH are nine (no. 1, 11, 12, 22, 26, 28, 36, 38 and 44). Noteworthy, THR feature (i) at position 1, only Asp (which is the second very highly conserved residue at that position in both THPC and CMTTH); (ii) a differential usage of the aromatic group Phe/Trp/Tyr compared to both THPC and CMTTH at either column 11 and 22; (iii) only Leu (which is the very highly conserved residue of the Leu/Ile/Met/Val group in CMTTH and the second very highly conserved residue of the same group in the THPC) at column 26; (iv) the usage of Leu and Val as the first and second very highly used residue of the Leu/Ile/Met/Val group at position 44 in THR, THPC and CMTTH. Finally, amino acids that are, at given positions, those very highly conserved in THR are downgraded to highly conserved in THPC (columns 34, 37, 48) or CMTTH (columns 8 and 13) (Figure 1). At other positions, the very highly conserved residues of THR are

downgraded to modestly conserved in THPC (columns 8, 12, 14, 16, 18, 19, 20, 29, 33 and 49) or CMTTH (columns 5, 19 and 32).

As done for CMTTH (see above), we wished to ascertain if THR were more similar to the apolipoprotein or the nonapolipoportein THPC. There are a total of nine matches with apolipoproteins (columns 1, 8, 11, 22, 26, 33, 34, 44, 48), and eight with nonapolipoproteins (columns 1, 11, 12, 22, 26, 37, 44, 48). The number of buffer residues at columns 10, 21 and 41 is slightly more in favor of apolipoproteins than nonapolipoproteins (8 *vs* 0-14 *vs* 0-23, 8-9 *vs* 0-30 *vs* 0-29, and 3 *vs* 0-4 *vs* 0-7).

### 4.4. Comparison between thyroid hormone membrane transporters (CMTTH) and nuclear receptors (THR)

While CMTTH or THR share with THPC seven or five positions where the very highly conserved residues belong to the same group, CMTTH and THR share nine (columns 1, 11, 12, 22, 26, 28, 36, 38, 44) (Figure 1). At another two positions, a highly conserved group of residues in CMTTH is upgraded to very highly conserved group in THR (columns 8, 13). Finally, the number of intervening residues is more similar at the N-terminal (0-14 *vs* 8) and C-terminal (0-4 *vs* 3) buffer positions than in the central buffer position (3-30 *vs* 8-9).

### 4.5. Overall comparison of the types of proteins and usage of amino acids at given positions

In the THPC, there is an overrepresentation of the D/E/N/Q group, especially in the apolipoproteins (19/46 columns or 41.3.%) (Figure 1). This compares with a frequency of 8/46 (17.4%, P=0.01) in the nonapolipoproteins, 2/46 (4.3%, P<0.0001) in the CMTTH and 6/46 (13.0%, P=0.002) in the THR. Noteworthy, of the said 19 columns, those N-terminal to the first buffer column have E (Glu) as the most used residue and N (Asn) the least used, while between columns 12 and 36 the most used one becomes D (Asp) almost always, and from column 37 it becomes N (Asn). The D/E/N/Q group occupies, though at a scanty degree of conservation, the variable third and fourth position (columns 24, 25) of the core TH-binding motif Y/F/W, L/I/M/V, X, X, V/L/I/M in apolipoproteins. In contrast, columns 24 and 25 in the nonapolipoprotein plasma carriers are occupied by the L/I/M/V group, and consistently in the hierarchical order of frequency L/V/M/I at either position. This group of residues (but in the order I/V/M/L) occupies column 24 in CMTTH.

In CMTTH and the nonapolipoprotein carriers, there is an overrepresentation of the L/I/M/V group (24/46 columns in CMTTH (52.2%), and 23/46 columns in nonapolipoproteins (50.0%)). To a lesser degree, this overrepresentation also occurs in the nuclear receptors (15/46 (32.6%), P=0.06 *vs* 24/46, P=0.09 *vs* 23/46). In CMTTH, the hierarchy of usage is L/I/V/M at six positions (column 2, 5, 20, 23, 30 and 40) and V/L/I/M at another three positions (columns 16, 28 and 47). In the nonapolipoprotein plasma carriers, the hierarchy of usage is L/V/I/M at six positions (columns 12, 24, 25, 43, 45 and 48), L/I/V/M at two positions (columns 15 and 35), and L/V/M/I at another two (columns 22 and 44). In THR, noteworthy is the frequent usage of Met, especially between columns 29 and 46.

The apolipoproteins and the nonapolipopoproteins have a similar number of positions where the conserved residues are the positively charged H/K/R (6 and 7 positions, respectively), but never at identical positions. R (Arg) is used more frequently than K (Lys) in the apolipoproteins, just the opposite of the nonapolipoproteins. The THR also have such residues conserved at 8 positions, only one (H at column 45) matching at least another type of proteins (R/K/H of apolipoproteins). THR are remarkable for having R/K, in lieu of S/T, at column 42. CMTTH have only two positions (no. 4 and 43) where the positively charged residues are conserved, and consistently in the order R>K>H. On the other hand, CMTTH are noticeable for the relative abundance of the A/G group between columns 29 and 49.

Concerning amino acid usage at the invariable positions in the 5-residue core thyroid hormone binding motif, (1, 2 and 5, that is columns 22, 23 and 26), there are evident differences. Column 22 is occupied by Y (Tyr) more frequently than F (Phe) in apolipoproteins, but the hierarchy of usage is F>Y>W (Trp) in nonapolipoproteins, W>F>Y in CMTTH and F>W>Y in THR. THR are remarkable for the lack of any residue of the L/I/V/M group at column 23 (a position that features W only). Careful inspection of Figure 1 shows that CMTTH have an apparent duplication of the 5-residue core TH binding motif, N-terminal of it (columns 11-15). Compared with the canonical motif at columns 22-26, the fifth position (column 15) features the L/I/V/M group that is conserved not so strongly as in column 26.

In sum, in the THPC, CMTTH and THR the invariable positions are these five: 1 (D/E/N/Q group), 11 and 22 (F/Y/W group), 26 and 44 (I/L/M/V group) (Figure 1).

### 4.6. Comparison among the types of proteins taking into account the thyroid hormone preferentially bound/transported

As well known, all THPC bind preferentially T4 while THR bind preferentially T3, such difference contributing to the greater biological potency of T3

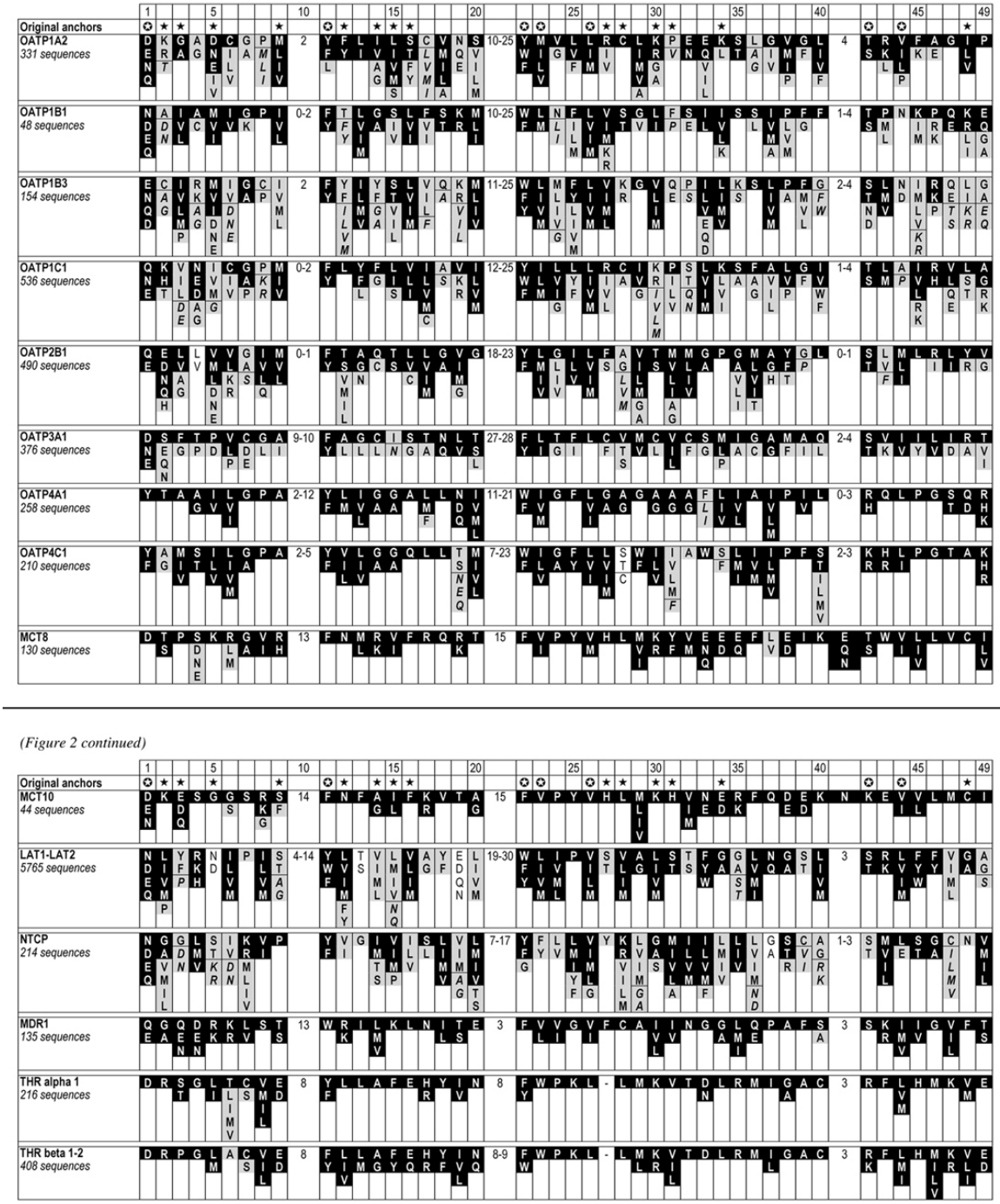


**Figure 2.** Consensus sequences for TH-binding cell proteins. Based on frequency in a given position, groups of similar amino acids are written in white boldface on black background (frequency >50%) or in black boldface on gray background (frequency 34% to 50%). When the latter case occurs for two groups in the same position but their frequencies are different, amino acids belonging to the less frequent group are written in italics. In each position of the overall consensus, amino acids belonging to the same group are shown from top to bottom in decreasing order of frequency.

compared to T4 (12). Concerning CMTTH, only for some of them the preferential transport is well-defined: MCT8 and MCT10 are preferential T3 carriers while OATP1C1 is a preferential T4 carrier (6-11). Thus, even though Figure 2 shows consensus sequences for 13 CMTTH, our analysis will be limited to OATP1C1, MCT8 and MCT10. The results of this analysis are shown in Table 2.

OATP1C1 shows more resemblance with the nonapolipoprotein than the apolipoprotein THPC (Figure 2). Indeed, OATP1C1 shares with the

**Table 2.** Matches of very highly (more than 50% of sequences) and highly (more than 33% and up to 50%) conserved amino acid groups in the consensus of thyroid hormone plasma carriers that bind preferentially T4 (apolipoproteins, nonapolipoproteins, OATP1C1) or T3 (MCT8, MCT10, THR alpha1, THR beta)

| Protein group | Apolipoproteins 10 - 18 | Nonapolipoproteins 11 - 8 | OATP1C1 39 - 40 [1,2] | MCT8 46 - 3 [1] | MCT10 47 - 3 [1] | THR alpha1 45 - 2 [1] |
|---|---|---|---|---|---|---|
| Nonapolipoproteins 11 - 8 | **7** (1, 11, 22, 23, 26, 42, 44) | | | | | |
| | **2** (3, 48) | | | | | |
| | **1** (5) | | | | | |
| OATP1C1 39 - 40 [1,2] | **7** (1, 11, 15, 22, 23, 26, 42) | **8** (1, 11, 22, 23, 24, 25, 26, 42) | | | | |
| | **4** (3, 30, 32, 47) | **1** (3) | | | | |
| | **3** (34, 38, 48) | **4** (12, 36, 37, 48) | | | | |
| MCT8 46 - 3 [1] | **8** (1, 11, 15, 22, 23, 26, 42, 44) | **7** (1, 11, 22, 23, 26, 42, 44) | **13** (1, 7, 11, 15, 22, 23, 26, 27, 29, 36, 42, 45, 47) | | | |
| | **0** | **0** | **2** (6, 37) | | | |
| | **3** (8, 12, 33) | **2** (36, 37) | **6** (2, 4, 13, 19, 25, 30) | | | |
| MCT10 47 - 3 [1] | **8** (1, 11, 15, 22, 23, 26, 43, 44) | **6** (1, 11, 22, 23, 26, 44) | **13** (1, 2, 11, 13, 15, 22, 23, 26, 27, 29, 36, 45, 47) | **29** (1, 4, 11, 12, 15, 16, 17, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 36, 38, 40, 41, 44, 45, 46, 47, 48, 49) | | |
| | **0** | **0** | **0** | **0** | | |
| | **5** (3, 9, 12, 33, 37) | **2** (3, 36) | **6** (3, 5, 8, 14, 25, 30) | **0** | | |
| THR alpha1 45 - 2 [1] | **6** (1, 11, 22, 26, 34, 44) | **6** (1, 11, 22, 26, 37, 44) | **10** (1, 2, 11, 12, 19, 22, 26, 29, 39, 48) | **15** (1, 8, 11, 13, 17, 22, 24, 26, 28, 29, 30, 33, 43, 44, 46) | **16** (1, 2, 11, 14, 17, 22, 24, 26, 28, 29, 30, 33, 35, 42, 44, 46) | |
| | **0** | **0** | **1** (6) | **1** (6) | **0** | |
| | **3** (8, 33, 48) | **2** (12, 48) | **10** (4, 5, 13, 14, 30, 31, 32, 34, 37, 45) | **1** (37) | **2** (6, 7) | |
| THR beta 44 - 2 [1] | **6** (1, 11, 22, 26, 34, 44) | **6** (1, 11, 22, 26, 37, 44) | **10** (1, 2, 11, 12, 19, 22, 26, 29, 39, 48) | **16** (1, 3, 8, 11, 13, 17, 22, 24, 26, 28, 29, 30, 33, 43, 44, 46) | **16** (1, 2, 11, 14, 17, 22, 24, 26, 28, 29, 30, 33, 35, 42, 44, 46) | **43** (1, 2, 4, 5, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49) |
| | **0** | **0** | **0** | **0** | **0** | **1** (7) |
| | **3** (8, 33, 48) | **2** (12, 48) | **10** (4, 5, 13, 14, 30, 31, 32, 34, 37, 45) | **1** (37) | **2** (6, 7) | **0** |

[1]: two highly conserved amino acid groups are present in some positions. [2]: a highly conserved group in addition to a very highly conserved one is present in some positions. For each combination, number and (in brackets) position of the matches within the 49 position scheme (see Figure 1) are given. Matches between very highly conserved amino acid groups are written in white in black cells, matches between highly conserved amino acid groups are written in black in gray cells, matches of a group which is very highly conserved in one series of proteins and highly conserved in the other are written in black in white cells. In the row and column headings, the number of very highly or highly conserved amino acid groups in each group of proteins (see also Figure 1) are typed in white boldface on black background and black boldface on gray background, respectively.

nonapolipoprotein THPC 8 positions where the same groups of amino acids are very highly conserved (columns 1, 11, 22-26 and 42), one position where the same group of amino acids is highly conserved (column 3), and 4 positions where the very high conservation of the same group of amino acids is downgraded to a high conservation (columns 12, 36, 37 and 48). Notably, of the positions where the I/L/V/M group is very highly or highly conserved, OATP1C1 features a greater usage of Ile compared to the nonapolipoproteins (11/28 (39%) *vs* 1/9 (11%), P= 0.22) (Figure 2). In contrast, the preferential T3-carrier MCT8 shares 7 positions of very high conservation of the same group of residues with the nonapolipoproteins (columns 1, 11, 22, 23, 26, 42 and 44), none where the same group of aminoacids is highly conserved, and 2 (columns 36 and 37) where the very high conservation of the same group of amino acids is downgraded to high conservation. The corresponding numbers for MCT10 are 6 positions (no. 1, 11, 22, 23, 26 and 44), no position, and 2 positions (no. 3 and 36). Preferential Ile usage occurs at 2/16 positions (12%) or 2/12 positions (17%) for MCT8 or MCT10, respectively.

MCT8 and MCT10 are by far more alike to the THR, which are also preferential T3-binders. MCT8 shares with THR alpha1 15 positions where the same group of amino acids are very highly conserved (no. 1, 8, 11, 13, 17, 22, 24, 26, 28-30, 33, 43, 44 and 46), similarly to the 16 positions concerning MCT10 (no. 1, 2, 11, 14, 17, 22, 24, 26, 28-30, 33, 35, 42, 44 and 46). The corresponding numbers for the comparison of MCT8 or MCT10 with THRβ are 16 positions (no. 1, 3, 8, 11, 13, 17, 22, 24, 26, 28-30, 33, 43, 44 and 46) or 16 positions (no. 1, 2, 11, 14, 17, 22, 24, 26, 28-30, 33, 35, 42, 44 and 46). MCT8 or MCT10 shares with THR alpha1 one (column 6) or no position where the same group of amino acids are highly conserved and none or none concerning the equivalent comparison of MCT8 or MCT10 with THRβ. In the comparison of MCT8 or MCT10 with THR alpha1, there is one or two positions (column 37 or columns 6-7) where the very high conservation of the same group of amino acids is downgraded to a high conservation; in the corresponding comparison of MCT8 or MCT10 with THR beta, the number of positions is one (no. 37) or two (no. 6 and 7) (Figure 2). The greater resemblance existing between the T3-binders (MCT8 and MCT10 on one side, and THRs on the other side) is highlighted further by the sole or preferential presence of D (Asp) in column 1, the peculiar presence of P (Pro) in column 24 and the peculiar string LMK (Leu, Met, Lys or conservative variation thereof) at columns 28-30 (Figure 2).

### 4.7. Overall consensus (based on 9,741 human and animal sequences)

As shown in Figure 1, the basic schema that had resulted from 426 sequences of human and animal THPC (3) is confirmed on a more than a 20-fold greater number of sequences that also contain TH binders localized in the plasma membrane or in the nucleus.

Seven of the 8 positions with a very high representation of a given group of residues in the 426 sequences (columns 1, 11, 22, 23, 26, 42 and 44) are confirmed in the 9,741 sequences. Other positions where representativity was high in the consensus from the 426 THPC sequences are confirmed in the consensus from 9,741 sequences (columns 15, 37), or are upgraded (column 24). Furthermore, other positions where representativity was modest are now upgraded to high (columns 8) or even very high (column 12).

### 4.8. Comparison of consensus sequences of two different thyroid hormone binders

We compared consensus sequences of two different thyroid hormone binders and found that they show high and low similarity in consensus sequences (Table 3).

We first examined ranking that is based on the number of positions where the same group of amino acids are conserved at very high degree (columns with the single letter amino acid-code typed boldface white against a black background). In the comparison between CMTTH and THPC, the pairs MDR1/ApoC-III and MCT10/Apo-E rank first, in that they share 15 positions out of the maximum possible of 46. In contrast, the pairs MCT8/CBG, MCT10/CBG, OATP4A1/CBG, OATP4C1/TTR, OATP4C1/CBG, OATP4C1/alpha1-AGP, NTCP/albumin and NTCP/alpha1-AGP rank last, in that they share only 4 positions. In the comparison between CMTTH and THR, the pairs MCT8/THR beta, MCT10/THR alpha1 and MCT10/THR beta rank first, in that they share 16 positions out of the maximum possible of 46. In contrast, the T4-carrier OATP1C1 shares 10 positions with either THR, while NTCP, MDR1 or OATP1B3 rank last with only 7 to 9 positions. Finally, in the comparison between THR and THPC, the pairs THR alpha1/Apo-E and THR beta/Apo-E rank first, in that they share 12 positions. In contrast, the pairs THR alpha1/TBG, THR alpha1/alpha1-AGP, THR beta/TBG and THR beta/alpha1-AGP rank last, in that they share only 4 positions.

Next, we examined a composite ranking that also takes into account conservation of the same group of amino acids at high degree (columns with the single letter amino acid-code typed against a gray background) or downgrade from a very high to a high degree of conservation. In the comparison between CMTTH and THPC, OATP1A2 and alpha1-AGP rank first, in that they share 31 positions out of the maximum possible of 46. In contrast, the pair OATP4C1/

**Table 3.** Highest and lowest results in the matching of very highly (>50% of sequences) and highly (>33% to 50%) conserved amino acid groups between the consensus motifs of the proteins belonging to thyroid hormone plasma carriers (apolipoproteins and nonapolipoproteins), cell surface thyroid hormone transporters and thyroid hormone nuclear receptors.

| Protein group | THPC | CMTTH | THR |
|---|---|---|---|
| **Highest homology** | | | |
| THPC | --- | **15** MDR1/ApoC-III, MCT10/Apo-E | **12** THR alpha1/Apo E, THR beta/Apo E |
| | --- | **31** OATP1A2/alpha1-AGP | **20** THR alpha 1/alpha1-AGP, THR beta/alpha1-AGP |
| CMTTH | **15** MDR1/ApoC-III, MCT10/Apo-E | --- | **16** MCT8/THR beta, MCT10/THR alpha1, MCT10/ THR beta |
| | **31** OATP1A2/alpha1-AGP | --- | **22** OATP1A2/THR alpha1 |
| THR | **12** THR alpha1/Apo E, THR beta/Apo E | **16** MCT8/THR beta, MCT10/THR alpha1, MCT10/THR beta | --- |
| | **20** THR alpha1/alpha1-AGP, THR beta/alpha1-AGP | **22** OATP1A2/THR alpha1 | --- |
| **Lowest homology** | | | |
| THPC | --- | **4** MCT8/CBG, MCT10/CBG, OATP4A1/CBG, OATP4C1/TTR, OATP4C1/CBG, OATP4C1/alpha1-AGP, NTCP/albumin, NTCP/alpha1-AGP | **4** THR alpha1/TBG, THR alpha1/alpha1-AGP, THR beta/TBG, THR beta/alpha1-AGP |
| | --- | **5** OATP4C1/TTR | **8** THR beta/Apo C-I |
| CMTTH | **4** MCT8/CBG, MCT10/CBG, OATP4A1/CBG, OATP4C1/TTR, OATP4C1/CBG, OATP4C1/alpha1-AGP, NTCP/albumin, NTCP/alpha1-AGP | --- | **7** NTCP/THR alpha1, NTCP/THR beta |
| | **5** OATP4C1/TTR | --- | **9** MDR1/THR alpha1, MDR1/THR beta |
| THR | **4** THR alpha1/TBG, THR alpha1/alpha1-AGP, THR beta/TBG, THR beta/alpha1-AGP | **7** NTCP/THR alpha1, NTCP/THR beta | --- |
| | **8** THR beta/Apo C-I | **9** MDR1/THR alpha1, MDR1/THR beta | --- |

CMTTH: Cell surface thyroid hormone transporters. THPC: thyroid hormone plasma carriers. THR: Thyroid hormone nuclear receptors. THPC consists of apolipoproteins and nonapolipoproteins. Matches between a very highly conserved group of residues in a given column of one protein and the same very highly conserved group of residues in the corresponding column of the other protein are written in white in black cells. Data written in black in white cells indicate any of (i) the aforesaid match, (ii) match between a highly conserved group of residues in a given column of one protein and the same highly conserved group of residues in the corresponding column of the other protein, (iii) a match between a group of residue that is very highly conserved or highly conserved in one protein and the same group of protein that is, respectively, highly conserved or very highly conserved in the other protein. In each cell, the number of matches and the pairs of proteins which have that number of matches are shown.

TTR ranks last, in that the two proteins share only 5 positions. In the comparison between CMTTH and THR, OATP1A2 ranks first with 22 positions shared with THR alpha1, while MDR1 ranks the last with only 7 positions shared with either THR. In the comparison between nuclear receptors and plasma carriers, the pairs THR alpha1/alpha1-AGP and THR beta/alpha1-AGP rank first, in that they share 20 positions. In contrast, the pair THR beta/Apo C-I rank last, in that the two proteins share only 8 positions.

In summary, using the stringent criterion of comparison based on the highest degree of similarity (that is, considering solely columns with the single letter amino acid-code typed boldface white against a black background), a consistent finding appears. The finding is that, regardless of isoform, THR (which have a preferential binding for T3) share the greatest number of positions featuring the same group of very highly conserved residues with MCT8 and MCT10 (MTC8 and MCT10 being the CMTTH that transport T3 preferentially). In contrast, regardless of isoform, THR share the lowest number of such positions with THPC (apo C-I, TBG, alpha-1 acid glycoprotein), which are known to transport preferentially T4.

## 5. DISCUSSION

We have validated the hypothesis that a large TH-binding motif, and in particular a 5-residue long core motif that was representative of 426 THPC sequences from *Homo sapiens* and another 94 animal species, would have been more or less conserved in thyroid hormone cell membrane transporters (CMTTH) and thyroid hormone nuclear receptors (THR). Previously, when we probed, with the same procedure as THPC, 2,918 sequences of proteins that are not known as TH binders, no TH-binding motif was detectable (3). These control sequences, which belong to *Homo sapiens* and another 55 species, were alpha- and beta-tubulin, eosinophil cationic protein, endothelin-1, -2 and -3, IgG receptor, tropomyosin, Wnt inhibitory factor 1, erythropoietin, haptoglobin and insulin (3).

Concerning the 5-residue core motif, the amino acids of the group I/L/M/V at position no. 3 (corresponding to position 24 in the large, 49-position motif) that are highly conserved in THPC become very highly conserved in CMTTH and in the consensus sequence representative of all 9,741 proteins. What changes is the hierarchy of amino acid usage in the first, second, third and fifth position (columns 22, 23, 24 and 26). Comparing THPC with CMTTH, consistent is the least usage of Met in the second and fifth position (columns 22 and 26) and the predominant usage of Leu in the second position. Perphaps, because THR include only three isoforms (THR alpha1, THR beta1 and THR beta2) that have an extremely high degree of homology between each other, at all five positions except the first only one amino acid is used. In contrast, CMTTH and THPC include 8,691 and 426 different sequences, respectively.

Upon perusing the consensus sequences from the 426 THPC (3), we were tempted to conclude that a less well-conserved copy of the 5-residue TH-binding motif is present N-terminal of it (columns 11-16). However, even though the first position continued to be very highly conserved (F/W/Y, column 11), we refrained from writing this conclusion, because the amino acids in the second position (L/V/I/M, column 12) were modestly conserved and those in the fifth position (V/L/I/M, column 16) were highly conserved, as opposed to both being very highly conserved in the original motif (columns 22-26). However, now one can observe (Figure 1) that the amino acids in the second position (column 12) are very highly conserved in both CMTTH and THR and, in CMTTH, with the same hierarchy of usage (L>V>I>M); amino acids in the fifth position are maintained in CMTTH only, though with a different usage of the four amino acids. Of interest, the fourth position (column 14) is occupied, with a modest degree of conservation, by Ala/Gly in THPC. Now, one can also observe that Ala/Gly are very highly conserved amino acids at that position in THR. Even the number of residues at the three buffer positions (columns 10, 21 and 41) is consistent, and even more so upon comparing the THPC with the CMTTH.

In summary, a variety of plasma, cell membrane and cell nucleus proteins so diverse genetically and functionally share the property of binding/transporting TH because they share a local region of homology, which we refer to as a large TH binding motif and that incorporates the 5-residue hydrophobic core motif originally detected in a few apolipoproteins of a few species (1, 2). The high degree of conservation of this domain across different proteins and species is confirmed by the remarkable E values ($10^{-15}$ to $10^{-62}$) obtained when comparing the homologous segments with the BLAST software. It is likely that the TH-binding domain of proteins originated from a common ancestor gene, with “variations on the theme” at certain positions dictated by localization in the body (fluids, cell, nucleus) and overall function of the single protein. Not unexpectedly, the large TH-binding motif in CMTTH and the one in THR are more similar to each other than either one is to the large TH-binding motif of THPC. Why both the large and short TH binding motifs remained so conserved during evolution and why so many proteins need to have TH-binding property are questions unanswered thus far. Similarly unanswered are the questions as to why deficiency or mutations of any of the THPC are unconsequential in terms of clinical thyroid abnormalities and biochemically measurable TH function, while inactivating mutations of at least one CMTTH (*viz*, MCT8) have devastating clinical consequences and measurable biochemical abnormalities of TH function, and inactivating mutations of any of the three THR isoforms also carry clinical consequences and measurable biochemical abnormalities of TH function.

## 6. REFERENCES

1. S. Benvenga, H.J. Cahnmann, D. Rader, M. Kindt, A. Facchiano, J. Robbins: Thyroid hormone binding to isolated human apolipoproteins A-II, C-I, C-II, and C-III: homology in

thyroxine binding sites. *Thyroid* 4, 261-267 (1994)
DOI: 10.1089/thy.1994.4.261
PMid:7833661

2. S. Benvenga: Thyroid hormone binding motif is evolutionarily conserved in apolipoproteins. *Thyroid* 7, 605-611 (1997)
DOI: 10.1089/thy.1997.7.605
PMid:9292950

3. S. Benvenga, F. Guarneri: Conservation in the phylum of the local homology of apolipoproteins with the thyroid hormone plasma carriers. *Rev Endocr Metab Disord* 17, 537-544 (2016)
DOI: 10.1007/s11154-016-9379-7
PMid:27549767

4. P.M. Yen, G.A. Brent: Genomic and nongenomic actions of thyroid hormones. In: Werner and Ingbar's The Thyroid: a fundamental and clinical text, 10th ed. Eds: LE Braverman, RD Utiger. Wolters Kluwer, Lippincott, Williams & Wilkins, Philadelphia, 127-138 (2013)

5. F. Guarneri, B. Guarneri: Bioinformatic analysis of HLA-linked genetic susceptibility to immunoallergic diseases: the MotiFinder software. *Ann Ital Dermatol Allergol* 64, 69-75 (2010)

6. E.C. Friesema, G.G. Kuiper, J. Jansen, T.J. Visser, M.H. Kester: Thyroid hormone transport by the human monocarboxylate transporter 8 and its rate-limiting role in intracellular metabolism. *Mol Endocrinol* 20, 2761-2772 (2006)
DOI: 10.1210/me.2005-0256
PMid:16887882

7. E.C. Friesema, J. Jansen, J.W. Jachtenberg, W.E. Visser, M.H. Kester, T.J. Visser: Effective cellular uptake and efflux of thyroid hormone by human monocarboxylate transporter 10 (MCT10) *Mol Endocrinol* 22, 1357-1369 (2008)

8. K. Fujiwara, H. Adachi, T. Nishio, M. Unno, T. Tokui, M. Okabe, T. Onogawa, T. Suzuki, N. Asano, M. Tanemoto, M. Seki, K. Shiiba, M. Suzuki, Y. Kondo, K. Nunoki, T. Shimosegawa, K. Iinuma, S. Ito, S. Matsuno, T. Abe: Identification of thyroid hormone transporters in humans: different molecules are involved in a tissue-specific manner. *Endocrinology* 142, 2005-2012 (2001)
DOI: 10.1210/endo.142.5.8115
PMid:11316767

9. W.M. van der Deure, P.S. Hansen, R.P. Peeters, K.O. Kyvik, E.C. Friesema, L. Hegedüs, T.J. Visser: Thyroid hormone transport and metabolism by OATP1C1 and consequences of genetic variation. *Endocrinology* 149, 5307-5314 (2008)
DOI: 10.1210/en.2008-0430
PMid:18566113

10. F. Pizzagalli, B. Hagenbuch, B. Stieger, U. Klenk, G. Folkers, P.J. Meier: Identification of a novel human organic anion transporting polypeptide as a high affinity thyroxine transporter. *Mol Endocrinol* 16, 2283-2296 (2002)
DOI: 10.1210/me.2001-0309
PMid:12351693

11. E.C. Friesema, S. Ganguly, A. Abdalla, J.E. Manning Fox, A.P. Halestrap, T.J. Visser: Identification of monocarboxylate transporter 8 as a specific thyroid hormone transporter. *J Biol Chem* 278: 40128-40135 (2003)

12. S. Benvenga. Thyroid hormone transport proteins and the physiology of hormone binding. In: Werner and Ingbar's The Thyroid: a fundamental and clinical text, 10th ed. Eds: LE Braverman, RD Utiger. Wolters Kluwer, Lippincott, Williams & Wilkins, Philadelphia, 93-103 (2013)

**Key Words**: Thyroid hormones, Cell membrane thyroid hormone carriers, Thyroid hormone nuclear receptors, Amino acid sequence homology, Motifs

**Send correspondence to**: Fabrizio Guarneri, Department of Clinical and Experimental Medicine, Dermatology, via Consolare Valeria, Gazzi, 98125 Messina, Italy, Tel: 39 090 2212890, Fax: 39 090 2927691, E-mail: f.guarneri@tiscali.it