Elsevier Editorial System(tm) for Molecular Immunology

Manuscript Draft

Manuscript Number:

Title: THE T-CELL RECEPTOR BETA CHAIN FROM SEA BREAM (Sparus aurata): MOLECULAR CLONING, EXPRESSION AND MODELLING OF THE COMPLEXES WITH CLASS I MHC

Article Type: Full Length Article

Section/Category:

Keywords: T-cell antigen receptor, TcRb, Sparus aurata, real time PCR, 3D structure, class I MHC, TcRb/MHC complexes

Corresponding Author: Dr Francesco Buonocore, Ph.D

Corresponding Author's Institution: University of Tuscia

First Author: Elisa Randelli, Ph.D

Order of Authors: Elisa Randelli, Ph.D; Viviana Scala; Daniela Casani; Susan Costantini, Ph.D; Angelo Facchiano, Ph.D; Massimo Mazzini; Giuseppe Scapigliati, Ph.D; Francesco Buonocore, Ph.D

Manuscript Region of Origin:

Abstract:

* Cover Letter

Dear Prof. M.R. Daha,

I am submitting you the paper: THE T-CELL RECEPTOR BETA CHAIN FROM SEA BREAM (*Sparus aurata*): MOLECULAR CLONING, EXPRESSION AND 3D STRUCTURAL ANALYSIS OF THE COMPLEXES WITH CLASS I MHC, author team: Elisa Randelli, Viviana Scala, Daniela Casani, Susan Costantini, Angelo Facchiano, Massimo Mazzini, Giuseppe Scapigliati, Francesco Buonocore for the publication on Molecular Immunology.

This paper is related to fish immunology and is focused on a molecule fundamental for adaptive immunity and for sea bream responses to virus and parasites.

Best regards

Dr. Francesco Buonocore, Ph.D

Univ. of Tuscia

Dep. Environmental Sciences

ABSTRACT

The T-cell antigen receptor (TcR) is a fundamental mediator of the adaptive immune responses, since $TCR_{\alpha\beta}$ molecules on T-cells recognize foreign structures (peptides derived from processed antigens) bound to major histocompatibility complex (MHC) antigens on another cell. In the present study we report the cDNA sequences of six TcRβ chains from sea bream (*Sparus aurata*), a fish of high economical impact in South Mediterranean aquaculture. Domains corresponding to the variable (V) region, diversity and joining (DJ) region, constant (C) Ig domain, connecting peptide (CPS), transmembrane region (TM) and cytoplasmic tail (CYT) were identified in all sequences. A multiple alignment of the predicted translation of the six sea bream molecules with other known TcRβ sequences was assembled and showed the conservation of some important features from fish to mammals. Phylogenetic analysis conducted using full-length amino acid sequences evidenced that fish sequences are in a different cluster with respect to mammalian and avian ones. Real-time PCR analysis was used to investigate TcRβ basal expression, that was maximum in the thymus followed by gut, and TcRβ expression after stimulation with LPS or PHA-L at 4 and 24 h (only the 4 h stimulation with LPS gave an significant effect). Moreover, the 3D structures of TcRβ chains and class I MHC were predicted by homology modelling with the aim to investigate the amino acid residues at the interaction surface in the TcRβ/MHC complexes.

Abstract

1

THE T-CELL RECEPTOR BETA CHAIN FROM SEA BREAM (*Sparus aurata*): MOLECULAR CLONING, EXPRESSION AND MODELLING OF THE COMPLEXES WITH CLASS I MHC

2
3
4
5

6 **Elisa Randelli, Viviana Scala, Daniela Casani, Susan Costantini[1], Angelo**
7 **Facchiano[1], Massimo Mazzini, Giuseppe Scapigliati, Francesco Buonocore***
8 Dipartimento di Scienze Ambientali, University of Tuscia, Largo dell'Università, 01100 Viterbo,
9 Italy
10 [1]CNR, Istituto di Scienze dell'Alimentazione, I-83100 Avellino, Italy
11
12
13 ***Corresponding author**: Dr. Francesco Buonocore, Dipartimento di Scienze Ambientali,
14 Università della Tuscia, Largo dell'Università s.n.c., I-01100 Viterbo, Italy.
15 Phone +39-0761-357644; Fax +39-0761-357179; Email: fbuono@unitus.it
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

ABSTRACT

The T-cell antigen receptor (TcR) is a fundamental mediator of the adaptive immune responses, since $TCR_{\alpha\beta}$ molecules on T-cells recognize foreign structures (peptides derived from processed antigens) bound to major histocompatibility complex (MHC) antigens on another cell. In the present study we report the cDNA sequences of six TcRβ chains from sea bream (*Sparus aurata*), a fish of high economical impact in South Mediterranean aquaculture. Domains corresponding to the variable (V) region, diversity and joining (DJ) region, constant (C) Ig domain, connecting peptide (CPS), transmembrane region (TM) and cytoplasmic tail (CYT) were identified in all sequences. A multiple alignment of the predicted translation of the six sea bream molecules with other known TcRβ sequences was assembled and showed the conservation of some important features from fish to mammals. Phylogenetic analysis conducted using full-length amino acid sequences evidenced that fish sequences are in a different cluster with respect to mammalian and avian ones. Real-time PCR analysis was used to investigate TcRβ basal expression, that was maximum in the thymus followed by gut, and TcRβ expression after stimulation with LPS or PHA-L at 4 and 24 h (only the 4 h stimulation with LPS gave an significant effect). Moreover, the 3D structures of TcRβ chains and class I MHC were predicted by homology modelling with the aim to investigate the amino acid residues at the interaction surface in the TcRβ/MHC complexes.

KEYWORDS : T-cell antigen receptor, TcRβ, *Sparus aurata*, real time PCR, 3D structure, class I MHC, TcRβ/MHC complexes

68 **1. INTRODUCTION**

69 Cell-mediated immunity is one of the main lines of defence that vertebrates rely on for
70 eliminating pathogen agents. T-cell receptor (TcR) antigens are present in all jawed vertebrates,
71 while they appear to be absent in jawless fish. In mammals, TcRs are heterodimeric, consisting of
72 either α/β or γ/δ polypeptide combination (Klausner et al., 1990). T-cells are activated when a TcR
73 heterodimer (αβ or γδ) in conjunction with the CD3 complex, on the surface of a T-cell, specifically
74 recognizes an external antigen (Ag). αβ TcR recognizes Ag presented by the major
75 histocompatibility complex (MHC) molecule expressed on antigen presenting cells (APC), while γδ
76 TcR, that in humans is expressed only by 1-10 % of T cells, is not MHC-restricted and appears to
77 recognize Ag in a similar manner to the direct Ag recognition processes of immunoglobulins (Ig)
78 (Chien et al., 1996).

79 The TcR chains (α, β, γ, δ) are composed of variable (V), joining (J) and sometimes diversity
80 (D) (only β) segments associated in translocon type organizations with exons specifying constant
81 domains, followed by a transmembrane domain and a short cytoplasmic tail. Recombination signal
82 sequences and recombination activator genes (RAG) are required for rearrangement of the TcR
83 segments (Marchalonis et al., 2002). The analysis of the TcR sequences led to the correct prediction
84 that TcR would share a domain organization and binding mode similar to those of antibody Fab
85 fragments (Claverie et al., 1989; Davis and Bjorkman, 1988). The first crystal structure of an αβ
86 TcR bound to a class I MHC restricted peptide was obtained in 1996 (Fields et al., 1996) and since
87 than on more data have been accumulated with 24 class I and class II TCR/MHC peptide complexes
88 available at the moment (Rudolph et al., 2006). Much less is known about γδ structures and the only
89 available data are on a quite recently obtained Vδ (Li et al., 1998) domain which shows an overall
90 architecture that closely resembles αβ TcR and Ig.

91 The structure and expression of genes corresponding to TcR chains have been characterised in
92 both teleost and cartilaginuos fishes (Hordvik et al., 1996; Partula et al., 1995; De Guerra and

93 Charlemagne, 1997; Wilson et al., 1998; Miracle et al., 2001; Wermenstam and Pilstrom, 2001;
94 Hordvik et al., 2004; Imai et al., 2005) with the Japanese flounder (*Paralichthys olivaceus*) being
95 the first species in which all the four genes (α, β, γ, δ) have been identified (Nam et al., 2003). It
96 has been demonstrated that TcR genes in teleost fish are organized in translocon type clusters (Zhou
97 et al., 2003) and, in contrast to mammals, that the TcR β loci possess a high degree of
98 polymorphism in the C region in bicolor damselfish (Kamper and Churchill McKinney, 2002).
99 Moreover, the allelic polymorphism in the TcR α and β constant domain genes is quite widespread
100 in fishes (Criscitiello et al., 2004). Finally, the expression of TcR β chain has been monitored
101 during amoebic gill disease in rainbow trout to verify its possible up/down regulation (Bridle et al.,
102 2006).

103 In the present study we report the cDNA sequences of six TcR β chains from sea bream (*Sparus*
104 *aurata*), a fish of high economical impact in South Mediterranean aquaculture, and their expression
105 pattern both at basal level and after "in vitro" stimulations. These results will add a new tool for
106 studying the effects of vaccinations and immuno-stimulations on the sea bream immune system.
107 Moreover, we analysed the cDNAs organisation and predicted, by homology modelling, the 3D
108 structures of the TcR β chains alone or complexed with the class I MHC.

109

110
111
112
113
114
115
116
117
118
119
120
121
122
123

124

125 ## 2. MATERIALS AND METHODS

126 ***2.1 Sea bream TcRβ cloning and sequencing***

127 Two degenerate primers (TCNERV: 5'- GCNGTKTAYTWCTGTGC -3' and TCNEFR: 5'-
128 SACRTGGTCRGGRTARAA- 3' where N = A,C,G,T ; W = A,T ; R = A,G ; K = G,T ; S = G,C ;
129 Y = C,T) corresponding to highly conserved regions of known TcRβ genes were used in RT-PCR
130 on total RNA extracted with Tripure (Roche) solution from a juvenile sea bream (150 g of weight)
131 thymus. The leukocyte cells were obtained following the procedures described in Scapigliati et al.
132 (2001). RT-PCR was performed using Ready-To-Go RT-PCR Beads (Amersham Pharmacia). For
133 cDNA synthesis, 1 μg of total RNA and 0.5 μg of random primers [pd(N)$_6$] were used in each
134 reverse transcription reaction in a total volume of 50 μl. Reactions were conducted using the
135 Mastercycler personal (Eppendorf). The cycling protocol was one cycle of 94°C for 5 min, 35
136 cycles of 94°C for 45 s, 48°C for 45 s, 72°C for 45 s, followed by one cycle of 72 °C for 10 min.
137 PCR products (15 μl) were visualised on 1% (w/v) agarose gels containing ethidium bromide (10
138 ng/ml) using hyperladder IV (Bioline) as size marker. Controls for the presence of DNA
139 contamination were performed using the RNA samples as template. DNA amplified by PCR was
140 purified using the QIAquick Gel Extraction Kit (QIAgen), inserted into the pGEM-T Easy vector
141 (Promega) and transfected into competent JM109 *Escherichia coli* cells. Plasmid DNA from at least
142 ten independent clones was purified using the Wizard Plus SV Minipreps DNA Purification System
143 (Promega) and sequenced using MWG DNA Sequencing Services. Sequences generated were
144 analysed for similarity with other known sequences using the FASTA (Pearson and Lipman, 1988)
145 and BLAST (Altschul et al., 1990) programs.

146 Further primers were designed based on the initial sea bream TcRβ sequence for 5'- and 3'-
147 rapid amplification of cDNA ends (RACE)-PCR (TCRORAFR1: 5'- CCACCAACAGTGAAAGTGC -3'
148 and TCRORAFR2: 5'-CGTTGCTTTGTGTGATCAGC-3'; TCRORARV1: 5'-
149 GCTGATCACACAAAGCAACG -3' and TCRORARV2: 5'-GCACTTTCACTGTTGGTGG-3').

150 cDNA was synthesised from total thymus RNA with the First-strand cDNA Synthesis kit
151 (Amersham Pharmacia) following the manufacturers instructions. For 3' RACE-PCR, cDNA was
152 transcribed using an oligo-dT adaptor primer (5'-CTCGAGATCGATGCGGCCGCT$_{15}$-3'). PCR
153 was performed initially with the TCRORAFR1 primer and the oligo-dT adaptor primer, followed
154 by a semi-nested PCR using TCRORAFR2 primer and the adaptor primer (5'-
155 CTCGAGATCGATGCGGCCGC-3'). For 5' RACE-PCR, cDNA was transcribed from total RNA
156 using the oligo-dT primer, treated with *E. coli* RNase H (Promega), purified using a PCR
157 Purification Kit (QIAgen), and tailed with poly(C) at the 5' end with terminal deoxynucleotidyl
158 transferase (TdT, Promega). PCR was performed initially with TCRORARV1 primer and an Oligo-
159 dG primer (5'-GGGGGGIGGGIIGGGIIG-3'), and then semi-nested with TCRORARV2 and the
160 oligo-dG primers. Sequencing and similarity searches were as described above.
161 The obtained cDNA sequences were analysed for the presence of a signal peptide, using SignalP
162 software (Nielsen et al., 1997), and of N- (with the NetNGlyc 1.0 Server) and O-linked
163 glycosylation sites (Julenius et al., 2005). Comparison of the sea bream TcRβ amino acid sequences
164 to their counterparts from other species was carried out using the MEGA 3.1 Software (Kumar et
165 al., 2004). A phylogenetic tree was constructed by the "neighbour-joining" method using MEGA
166 3.1 Software (Kumar et al., 2004) on full-length amino acid sequences and bootstrap values
167 calculated.

168 ***2.2 Basal TcRβ expression analysis***

169 To study the TcRβ basal expression, ten sea bream juveniles were sampled and leucocytes from
170 different tissues (thymus, spleen, liver, gills, head kidney, gut) obtained as described in Scapigliati
171 et al., 2001. Total RNA was isolated from each tissue separately with Tripure (Roche) following the
172 manufacturer's instructions, resuspended in DEPC treated water and used for real-time quantitative
173 PCR without pooling the samples coming from the different fishes. Controls for the presence of
174 DNA contamination were performed using using β-actin primers that bracket an intron.

175 For reverse transcription, the BioScript RNase H minus (Bioline) enzyme was used with the
176 following protocol: 2 μg of total RNA was mixed with 1 μl of random hexamer (0.2 μg/μl;
177 Amersham Pharmacia) and nuclease free water was added to a final volume of 12 μl. This mixture
178 was incubated at 70° C for 5 min and then cooled on ice. Successively, 0.4 μl of a reaction mix
179 containing 100 mM dNTPs (25mM each; Promega), 4 μl of 5X Reaction buffer, nuclease free water
180 to a final volume of 19.75 μl and 0.25 μl of BioScript at 200 u/μl were added and the solution
181 incubated at 25 °C for 10 min and then at 37 °C for 60 min. Finally, the reaction was stopped by
182 heating at 70 °C for 10 min.

183 The expression level of TcRβ was determined with a Mx3000P™ real time PCR system
184 (Stratagene) equipped with version 2.02 software and using the Brilliant SYBR Green Q-PCR
185 Master Mix (Stratagene) following the manufacturer's instructions, with ROX as internal passive
186 reference dye. The reference dye is not reactive during real time PCR and therefore can be used to
187 normalize slight differences in the volume of the added real time PCR reaction, transparency of the
188 plastic caps and other sources of well-to-well differences. Specific PCR primers were designed for
189 the amplification of about 200 bp products from both TcRβ (selected in the Ig constant region) and
190 β-actin, used as an house-keeping gene. The primers were: RQTCRORAFW2: 5'-
191 GTCGACGCCAACAAATGG- 3' AND RQTCRORARW2: 5'-AGACGCAGCCCTTGATGA-3',
192 RTACTFR2: 5'-ATGTACGTTGCCATCC-3' and RTACTRV2: 5'-GAGATGCCACGCTCTC-3',
193 respectively. Approximately 20 ng of cDNA template was used in each PCR reaction. The PCR
194 conditions were 95 °C for 10 min, followed by 35 cycles of 95 °C for 45 s, 52 °C for 45 s and 72 °C
195 for 45 s. Triplicate reactions were performed for each template cDNA and the template was
196 replaced with water in all blank control reactions. Each run was terminated with a melting curve
197 analysis which resulted in a melting peak profile specific for the amplified target DNA. The
198 analysis was carried out using the endpoints method option of the Mx3000P™ software that causes
199 the collection of the fluorescence data at the end of each extension stage of amplification. A relative

200 quantitation has been performed, comparing the levels of the target transcript (TcRβ) to a reference
201 transcript (calibrator, the tissue with the lowest TcRβ expression, in this case the head kidney) with
202 an expression level of the gene of interest (TcRβ) defined as 1.0. A normalizer target (β-actin) is
203 included to correct for differences in total cDNA input between samples. The quantitative
204 experiment is based on threshold cycle (Ct) determination, defined as the cycle at which a
205 statistically-significant increase in fluorescence (above background signal) is detected. The results
206 are expressed as the mean ± SD of the results obtained from the ten considered fishes. The real-time
207 PCR products from the different tissues were examined successively by agarose gel electrophoresis
208 to investigate their specificity and size.

209 ***2.3 In vitro TcRβ expression analysis***

210 The *in vitro* TcRβ expression was studied by real time PCR using different stimulating
211 conditions on head kidney (HK) leucocytes obtained from ten sea bream juveniles as described
212 above. HK leucocytes were adjusted to 1 x 10^5 cells/ml and incubated at 18 °C for 4 h and 24 h with
213 5 μg/ml of lipopolysaccharide (LPS from *E. coli* 0127:B8, Sigma) in PBS or with 1 μg/ml of lectin
214 from *Phaseolus vulgaris* Leucoagglutinin (PHA-L from Sigma) in PBS. The control was stimulated
215 with PBS only at the same time points.

216 Total RNA was isolated with Tripure (Roche) following the manufacturer's instructions,
217 resuspended in DEPC treated water and used for real-time quantitative PCR without pooling the
218 samples coming from the different fishes. The primers and the real time PCR conditions were the
219 same as described above, except that the calibrators for these experiments were the time 0 control.
220 Data were expressed as the mean ± SD of the results obtained from the ten considered fishes and the
221 differences from the control have been considered significant if $p < 0.05$ using the standard *t* student
222 test.

223 ***2.4 3D Modelling of sea bream TcRβ chains***

224 The 3D models of the six sea bream TcRβ sequences (region 21-140) were created according to
225 the homology modelling strategy and using as templates the experimental human and murine
226 structures of TcRβ complexed with HLA-A2 (PDB code: 2GJ6, chain E) (Gagnon et al., 2006) and
227 I-Ak proteins (PDB code: 1G6R, chain B) (Degano et al., 2000), respectively. As the sequence
228 identities between the sea bream TcRβ chains and the human and murine homologous templates
229 were close to 30%, we used an accurate procedure for the modelling strategy (Facchiano et al.,
230 2001; Scapigliati et al., 2004; Costantini et al., 2005; Buonocore et al., 2006; Buonocore et al.,
231 2007; Costantini et al., 2007). The search for sequence similarity within databases was performed
232 with the BLAST program (Altschul et al., 1990). The alignment of the protein sequences was made
233 with CLUSTALW program (Thompson et al., 1994) and a few manual refinements were added to
234 account for the position of secondary structures.

235 The MODELLER module (Sali and Blundell, 1993) implemented in the Quanta molecular
236 simulation package (Accelrys Inc., San Diego, CA, USA) was used to build 10 full-atom models of
237 each TCRβ chain by setting 4.0 Angstroms as RMS deviation among initial models and by full
238 optimization of models. To select the best model, we used the PROCHECK program (Laskowski et
239 al., 1993) to evaluate their stereochemical quality and structural packing quality, and ProsaII
240 program (Sippl, 1993) to check the fitness of the sequences to the obtained structures and to assign
241 a scoring function. Secondary structures were assigned by the DSSP program (Kabsch and Sander,
242 1983). Search for structural classification was performed on CATH database (Orengo et al., 1997;
243 Pearl et al., 2000). Molecular superimposition, RMSD values and figures were obtained with the
244 Insight II package (Accelrys Inc., San Diego, CA, USA).

245 ***2.5 3D Modelling of sea bream class I MHC***

246 The modelling of sea bream MHC class I (20-294 amino acid region, accession number
247 DQ211541) was performed using as template the experimental human and murine structures of
248 MHC class I complexed with TcRβ (PDB code: 2GJ6, chain A, and 1G6R, chain H) (Gagnon et al.,

249 2006; Degano et al., 2000). The 3D structure of sea bream MHC class I was modelled and analysed
250 with the same procedures and software used for TcRβ chains.

251 ***2.6 Simulation of sea bream class I MHC/TcRβ complexes***

252 The sea bream MHC class I/TcRβ complexes were created using as reference the experimental
253 models of human and murine MHC class I complexed with the related TcRβ chains (Gagnon et al.,
254 2006; Degano et al., 2000).

255 CVFF force field developed for the Insight II software was adopted to assign potentials and
256 charges for the complexed proteins. The complexes were minimized by using 500 steps of energy
257 minimization under conjugate gradient algorithm in order to optimise side chain conformations and
258 avoid sterical clashes (Scapigliati et al., 2004; Costantini et al., 2005; Chambery et al., 2007;
259 Costantini et al., 2007; Gianfrani et al., 2007). To compare the protein-protein interaction in the
260 complexes, the "Protein - Protein Interaction Server" (Jones and Thornton, 1996) and the program
261 NACCESS (Hubbard et al., 1991) were used to evaluate the interface surface area and to identify
262 the amino acids at the protein-receptor interface. H-bonds were calculated with Hbplus program
263 (McDonald and Thornton, 1994), that identifies H-bonds within a distance of 2.5 Å and a minimum
264 angle of 90°. Moreover, the energy of interaction between MHC and TcRβ was also calculated by
265 using the Energy/Intermolecular tool in the Docking module of Insight II and the binding free
266 energy by using the DCOMPLEX program (Liu et al., 2004).

267

268

269

270

271

272

273

274

275 # 3. RESULTS

276 ***3.1 Sea bream TcRβ sequences analysis***

277 PCR with degenerate primers TCNEFR and TCNERV gave a product of the expected size (219
278 bp) that when sequenced showed high homology for other known TcRβ sequences (data not
279 shown). 3'-RACE-PCR performed with TCRORAFR2 (based on the initial 219 bp sequence) and
280 the adaptor primer gave a product of about 927 bp that contained the 3'-end of the gene. 5'-RACE-
281 PCR performed with TCRORARV2 (based on the initial 219 bp sequence) and oligo- dG gave a
282 product of about 476 bp that contained the 5'-end of the gene, with some differences between the
283 selected clones. The six full-length nucleotide sequences (EMBL accession numbers **AM261209**;
284 **AM261210**; **AM490435**; **AM490436**; **AM490437**; **AM490438**) are comprised of 1428 bp from the
285 three overlapping products and have been confirmed by PCR using primers that amplify the coding
286 sequence (data not shown). They encode for 308 amino acids, with a predicted 21 amino acid signal
287 peptide, a 44 bp 5'-UTR and a 482 bp 3'- UTR. Finally, the 3'-UTR contained a polyadenylation
288 signal (AATAAA) 18 bp upstream of the poly(A) tail.

289 A multiple alignment of the predicted translation of the six sea bream molecules with other
290 known TcRβ sequences was assembled (Fig. 1) to investigate the conservation of characteristic
291 amino acid residues involved in structural domains found within TcRs. Domains corresponding to
292 the variable (V) region, diversity and joining (DJ) region, constant (C) Ig domain, connecting
293 peptide (CPS), transmembrane region (TM) and cytoplasmic tail (CYT) were identified according
294 to Hein (1994). An analysis of the *Sparus aurata* clone 1 sequence (Fig. 1) showed the presence of
295 four putative N-glycosylation sites, one in the DJ region and three in the C domain, and no O-
296 glycosylation sites. A similar situation for the putative glycosylation sites has been found in the
297 channel catfish (*Ictalurus punctatus*) (Wilson et al., 1997), in the Atlantic salmon (*Salmo salar*)
298 (Hordvik et al., 1996) and in the Japanese flounder (*Paralichthys olivaceus*) (Nam et al., 2003).

299 Some amino acid residues are conserved in all the considered sequences (Fig. 1) from fish to
300 mammals and this should indicate their importance for the functional activity of TcRβ molecules.

301 The two cysteine residues that in human (C-42/C-110) are involved in an intrachain disulfide bond
302 fundamental for the assembly of the V domain (Wilson and Garcia, 1997) are conserved in sea
303 bream, together with the sequence Phe^{122}-Gly^{123}-X^{124}-Gly^{125} in the DJ region, typical of most TcRs
304 and Ig light chains (Wilson et al., 1998). The constant Ig domain is the region that presents the
305 highest homology between all considered sequences. The two cysteine residues involved in human
306 (C-162/C-227) in the formation of another intrachain disulfide bond (Wilson and Garcia, 1997), the
307 sequence Ser^{211}-Arg^{212}-Leu^{213}, linked to the TcRα/β interaction (Arnaud et al., 1997), and a Glu
308 residue (E-151 in human), which is thought to form an hydrophilic bond with the TcRα chain
309 (Garcia et al., 1996), are conserved in sea bream. On the contrary, the Cys residue that in mammals
310 (Wilson and Garcia, 1997) is known to form an interchain disulphide bond with TcRα chain is not
311 present, as already observed in all known teleost sequences (Hordvik et al., 1996; Wilson et al.,
312 1998; Wermenstam and Pilstrom, 2001; Nam et al., 2003). The CPS shows no conserved amino
313 acids, while in the TM region the residue K^{280}, that should be important for the TcR cell surface
314 expression and for the assembly of the complex between TcRβ and CD3 (Morley et al., 1988;
315 Alcover et al., 1990), is present in sea bream.
316 Phylogenetic analysis (Fig. 2) conducted using full-length amino acid sequences showed that
317 fish sequences are in a different cluster with respect to mammalian and avian sequences. The six sea
318 bream TcRβ sequences are all in the same cluster but divided in four different groups: *Sparus*
319 *aurata* clone 2 and clone 3 (with 93 % sequence identity), and *Sparus aurata* clone 1 and clone 6
320 (with 94 % sequence identity) are in the same group, whereas *Sparus aurata* clone 4 and *Sparus*
321 *aurata* clone 5 are in two different groups.
322 ***3.2 Basal TcRβ expression analysis***
323 The expression analysis of TcRβ in organs and tissues of un-stimulated sea bream is shown in
324 Figure 3. Real-time PCR products were loaded on agarose gels to exclude the formation of non-
325 specific amplicons and single bands of the expected sizes were obtained. Moreover, to take into

326 consideration the individual genetic variability ten different fish were sampled and analysed
327 separately. The highest TcRβ expression was detected in thymus, followed by gut. Lower TcRβ
328 mRNA levels were observed in spleen, liver and gills. The head kidney showed the lowest
329 expression level.

330 ***3.3 In vitro TcRβ expression analysis***

331 To investigate whether TcRβ expression level could be modulated with LPS and PHA-L, *in*
332 *vitro* stimulation of HK leucocytes for a short (4 h) and a longer (24 h) time was studied; RNA was
333 extracted from the HK leucocytes of ten individuals analysed separately. Real-time PCR products
334 were loaded on agarose gels and single bands of the expected sizes were obtained. The results are
335 shown in Figure 4. Using LPS (Fig. 4), a dramatic increase of TcRβ expression was detected at 4 h
336 ($p<0,05$), whereas at 24 h ($p<0,05$) it was similar to the control. When using PHA for *in vitro*
337 treatment (Fig. 4), a little decrease was observed both at 4 h and 24 h, although being not
338 statistically significant.

339 ***3.4 3D Modelling of sea bream TcRβ chains***

340 The crystallographic structures of human and mouse TcRβ chains, complexed with HLA-A2
341 (PDB code: 2GJ6, chain E) (Gagnon et al., 2006) and I-Ak (PDB code: 1G6R, chain B) (Degano et
342 al., 2000) respectively, were selected as template models as the sequence identity percentages
343 between sea bream and man/mouse TcRβ molecules ranged from 28% to 35%. The alignments of
344 the sea bream TcRβ sequences with human and murine ones and the related experimental structures
345 (data not shown) were used to create structural models for each sea bream TcRβ chain for the
346 region between the amino acid 21 and 140. In Figure 5 the 3D structure for the TcRβ *Sparus*
347 *aurata* clone 3 (the sequence with the highest similarity with human and murine templates) with its
348 secondary structure elements is shown. About 50% of the residues in each sea bream TcRβ model
349 folds in beta conformation, thus defining the global structure as “mainly beta” with
350 immunoglobulin-like topology in agreement to the structural classification reported by CATH

351 database (Orengo et al., 1997) for the reference structures of human and murine TCRβ. In each
352 TCRβ molecule the position and orientation of two Cys residues make it possible that the structure
353 is stabilized by an intrachain disulfide bond. The short 3_{10} helix observed in both crystallographic
354 human and murine structures is present also in all sea bream models.

355 The TCRβ models were compared by superimposition of structurally conserved regions,
356 obtaining root mean square deviation (RMSD) values related to the differences in terms of
357 secondary structures. RMSD values range between 0.42 and 0.99 Angstroms, indicating that these
358 conformations present some structural differences, even if the overall 3D structure is conserved.

359 ***3.5 3D Modelling of sea bream class I MHC***

360 The search in the non-redundant database evidenced that the 3D structure of HLA-A2 (PDB
361 code: 2GJ6, chain A) (Gagnon et al., 2006) and I-Ak (PDB code: 1G6R, chain H) (Degano et al.,
362 2000) can be used to obtain the 3D model of sea bream class I MHC as the sequence identity
363 percentages between sea bream and man/mouse resulted of 31 and 34%, respectively. Starting from
364 the alignment of these three sequences (data not shown), ten structural models were created for sea
365 bream class I MHC 20-294 amino acid region and the best model (Figure 6) was selected as
366 described before (see the section: "Materials and Methods). This model has a classical organization
367 in three distinct domains (i.e. α1, α2 and α3) in agreement with the structural classification
368 reported by CATH database (Orengo et al., 1997; Pearl et al., 2000) for the crystallographic
369 structures of human and murine class I MHC molecules. The α1 and α2 domains are classified as
370 "alpha-beta" and each consists of an alpha helical region and four strands of beta sheet in an anti-
371 parallel orientation. The C-terminal domain (α3) has a "mainly beta" fold and is characterized by an
372 immunoglobulin-like beta-sandwich made of two anti-parallel sheets, each consisting of three main
373 strands and few shorter strands, organized in greek-keys motifs. In the sea bream class I MHC
374 model four cysteine residues are located in the same positions of the human and murine templates
375 and they should form two intrachain disulfide bonds, one in the α2 domain and the other in the α3

376 domain. The sea bream class I MHC model was compared by structural superimposition, as for
377 TCRβ models, with the two experimental structures used as templates, but the presence of gaps in
378 the alignment made it difficult to perform a complete structural comparison of the three models
379 (RMSD values of 1.12 and o.96, respectively). However, the comparison of the secondary
380 structures evidenced that the helices and β-strands are conserved, with few external residues
381 resulting added or excluded in some secondary structure elements and that the 3_{10} helix present in
382 human and murine structures is not conserved in the sea bream class I MHC model.

383 ***3.6 Simulation of sea bream class I MHC/TcRβ complexes***

384 On the basis of the crystallographic structures of the human and murine class I
385 MHC/TcRβ complexes we simulated the interaction between the six TcRβ chains and class I MHC
386 in sea bream according to the procedure already described (Scapigliati et al., 2004; Costantini et al.,
387 2005; Chambery et al., 2007; Costantini et al., 2007; Gianfrani et al., 2007). For each complex we
388 evaluated the interaction residues, the number of intrachain H-bonds and the interface surface area
389 (Table I). The interaction regions between the class I MHC and TCRβ in each complex are well
390 conserved (data not shown) and, in particular, the amino acids of TcRβ chains at the interface with
391 class I MHC are located in loop regions, whether the class I MHC ones in helical regions. The
392 number of interaction residues is different in the several complexes, due to the amino acid
393 differences between the six sea bream TcRβ amino acid sequences. The value of interface surface
394 area and the number of intrachain H-bonds resulted the highest for the complex between class I
395 MHC and TcRβ clone 3 (see Table I); the other complexes show little differences in the interface
396 surface area values and variable numbers of intrachain H-bonds.

397
398
399
400
401
402

403 **4. DISCUSSION**

404 The heterodimeric, membrane anchored TcR endows the cell with antigen specific reactivity
405 and this molecule has been extensively studied in mammals. Fish, like all jawed vertebrate, can
406 develop cell mediated immunity and the knowledge on fish TcR has increased rapidly in recent
407 years.

408 In this paper we report the homology cloning of TcRβ chain from sea bream (*Sparus aurata*),
409 and analysed its expression and 3D structure, with a particular interest to the interactions with class
410 I MHC.

411 Six different TcRβ cDNAs were isolated from the thymus of a juvenile sea bream and the
412 predicted amino acid sequences were studied to evidence the presence of conserved structures. The
413 J segment motif (Moss et al., 1995), F-G-X-G, and the four Cys residues involved in mammals in
414 two intrachain disulfide bonds are conserved in all sea bream sequences (see Figure 1). The Cys
415 that in mammals is involved in an interchain disulfide bond with TcRα is not present in sea bream,
416 but some authors (Arnaud et al., 1997) demonstrated that it may be not important for the
417 dimerisation of the TcRαβ heterodimer. On the contrary, the amino acid motif (YCLSSRLRVSA)
418 that the same authors proposed to be involved in the αβ chain interaction and cell surface
419 expression, is partially conserved in all sequences. Therefore, the amino acids (the Y and the stretch
420 SRL) that in this motif are retained from fish to mammals should be fundamental for the
421 structure/function relationship in TcRβs. The conserved antigen receptor transmembrane (CART)
422 motif (Campbell et al., 1994), made mainly of polar or aromatic amino acids, is present in sea
423 bream TM sequences like in all teleost TcRβ. It should encode a structural unit with an important
424 role in the assembly and/or signalling of the TcR/CD3 complex (Campbell et al., 1994). In
425 particular, the two tyrosine residues (Y-265 and Y-275) that have been identified to have a
426 fundamental role in the signalling of TcRβ towards the ξ subunit of CD3 (Fuller-Espie et al., 1998)
427 are conserved in sea bream.

428 Phylogenetic analysis (see Fig. 2) evidenced that, although some features are conserved, fish are
429 in a different cluster compared to mammals and sea bream TcRβ sequences are inside this cluster,
430 with bicolor damselfish and rainbow trout as the most closely related.

431 Real-time PCR analysis detected high levels of constitutive TcRβ expression in thymus (see
432 Fig. 3), as already observed in cod (Wermenstam and Pilstrom, 2001) and rainbow trout (Partula et
433 al., 1995), followed by gut, in contrast to what was detected in cod (Wermenstam and Pilstrom,
434 2001) where no expression was found in the intestine. However, the gut is a tissue with a certain
435 number of T-cells (Romano et al., 2007) and deeply involved in fish immune responses and
436 leukocytes ontogeny as demonstrated in rainbow trout (Bernard et al., 2006), sea bass and carp
437 (Rombout et al., 2005). The *in vitro* TcRβ expression was studied using sea bream head kidney
438 leukocyte cells and the selected stimulants were LPS, to simulate a pathogen infection, and the cell
439 mitogen agent PHA-L. LPS induced an high TcRβ expression in sea bream after 4 h as observed
440 also *in vivo* on mice (Tough et al., 1997), whether PHA-L showed no effect after both at 4 h and 24
441 h. Probably, this last result was due to the stimulation period that was too short, as in the *Pagrus*
442 *auratus* (Morrison et al., 2004) leukocyte proliferation was induced significantly by PHA only after
443 72 h.

444 3D structural analysis of sea bream TcRβ chains showed that the beta-strands are well
445 conserved in all models, although some little differences could be evidenced. In fact, few external
446 residues result added or excluded in the β-strands, and, in TCRβ clone 1 and clone 2 a further short
447 β-strand in the 57-58 residue region is un-recognized (data not shown). The simulation of the
448 complexes between the TcRβ chains and the class I sea bream MHC gave us the opportunity to
449 study the amino acid residues at the interaction surface. Furthermore, for each complex we have
450 evaluated the binding energy and interaction energy (Fig. 7) and the best values have been obtained
451 for the complex with TcRβ clone 3: this complex showed also the highest value of interface surface
452 area and the highest number of intrachain H-bonds (see Table I).

In conclusion, the availability of sea bream TcRβ sequences will add a new tool to analyse fish immune responses to viral and parasitic infections and will be of relevance to study the phylogeny and evolution of specific antigen receptors. Moreover, the investigation on the TcRβ/class I MHC complexes will help to have information about the putative amino acid residues at the interaction surfaces. The step forward will be the search of an experimental testing to assess the 3D structures reliability: probably it could be possible to make some synthetic peptides that should block the interactions between TcRβ and class I MHC, based on the conserved amino acid residues at the interaction surfaces, and investigate if these peptides are able to modulate sea bream immune responses.

ACKNOWLEDGEMENTS

Authors are indebted to Dr. C. Magugliani and Dr. E. Vela (Nuova Azzurro, Civitavecchia) for the supply of fish. This work was supported by the European Commission within the project IMAQUANIM (EC contract number FOOD-CT-2005-007103).

479 **REFERENCES**

480 Alcover, A., Mariuzza, R.A., Ermonval, M., Acuto, O., 1990. Lysine 271 in the transmembrane
481 domain of the T-cell antigen receptor β chain is necessary for its assembly with the CD3 complex
482 but not for α/β dimerization. J. Biol. Chem. 265, 4131-4135.

483 Altschul, S.F., Gish, W., Miller, W., Myers, E., Lipman, D.J., 1990. Best local alignment search
484 tool. J. Mol. Biol. 215, 403-410.

485 Arnaud, J., Huchenq, A., Vernhes, M.-C., Caspar-Bauguil, S., Lenfant, F., Sancho, J., Terhost,
486 C., Rubin, B., 1997. The interchain disulfide bond between TCRαβ heterodimers on human T cells
487 is not required for TCR-CD3 membrane expression and signal transduction. Inter. Immunol. 9, 615-
488 626.

489 Bernard, D., Six, A., Rigottier-Gois, L., Messiaen, S., Chilmonczyk, S., Quillet, E., Boudinot,
490 P., Benmansour, A., 2006. Phenotypic and functional similarity of gut intraepithelial and systemic T
491 cells in a teleost fish. J. Immunol. 176, 3942-3949.

492 Bridle, A.R., Morrison, R.N., Nowak, B.F., 2006. The expression of immune-regulatory genes
493 in rainbow trout, *Oncorhynchus mykiss*, during amoebic gill disease (AGD). Fish Shellfish
494 Immunol. 20, 346-364.

495 Buonocore, F., Randelli, E., Bird, S., Secombes, C.J., Costantini, S., Facchiano, A., Mazzini,
496 M., Scapigliati, G., 2006. The CD8α from sea bass (*Dicentrarchus labrax* L.): cloning, expression
497 and 3D modelling. Fish Shellfish Immunol. 20, 637-646.

498 Buonocore, F., Randelli, E., Casani, D., Costantini, S., Facchiano, A., Scapigliati, G., Stet, R.J.,
499 2007. Molecular cloning, differential expression and 3D structural analysis of the MHC class-II
500 beta chain from sea bass (*Dicentrarchus labrax* L.). Fish Shellfish Immunol., in press.

501 Campbell, K.S., Backstrom, B.T., Tiefenthaler, G., Palmer, E., 1994. CART: a conserved
502 antigen receptor transmembrane motif. Semin. Immunol. 6, 393-410.

503 Chambery, A., Pisante, M., Di Maro, A., Di Zazzo, E., Costantini, S., Colonna, G., Parente, A.,
504 2007. Invariant Ser211 is involved in the catalysis of PD-L4, type I RIP from (*Phytolacca dioica*)
505 leaves. Proteins: Structure, Function and Bioinformatics 67, 209-218.
506 Chien, Y.H., Jores, R., Crowley, M.P., 1996. Recognition by γ/δ T cells. Annu. Rev. Immunol.
507 14, 511-532.
508 Claverie, J.M., Prochnicka-Chalufour, A., Bougueleret, L., 1989. Implications of a Fab-like
509 structure for the T-cell receptor. Immunol. Today 10, 10-14.
510 Costantini, S., Colonna, G., Facchiano, A.M., 2007. Simulation of conformational changes
511 occurring when a protein interacts with its receptor. Comput. Biol. Chem. 31, 196-206.
512 Costantini, S., Colonna, G., Rossi, M., Facchiano, A.M., 2005. Modelling of HLA-DQ2 and
513 simulations of its interaction with gluten peptides to explain molecular recognition in celiac disease.
514 J. Mol. Graph. Model. 23, 419-431.
515 Criscitiello, M.F., Wermenstam, N.E., Pilstrom, L., McKinney, E.C., 2004. Allelic
516 polymorphism of T-cell receptor constant domains is widespread in fishes. Immunogenetics 55,
517 818-824.
518 Davis, M.M., Bjorkmann, P.J, 1988. T-cell antigen receptor genes and T-cell recognition.
519 Nature 334, 395-402.
520 Degano, M., Garcia, K.C., Apostolopoulos, V., Rudolph, M.G., Teyton, L., Wilson, I.A., 2000.
521 A functional hot spot for antigen recognition in a superagonist TCR/MHC complex. Immunity 12,
522 251-261.
523 De Guerra, A., Charlemagne, J., 1997. Genomic organisation of the TcR β-chain diversity (*Dβ*)
524 and joining (*Jβ*) segments in the rainbow trout: presence of many repeated sequences. Mol.
525 Immunol. 34, 653-662.
526 Facchiano, A.M., Stiuso, P., Chiusano, M.L., Caraglia, M., Giuberti, G., Marra, M., Abruzzese,
527 A., Colonna, G. 2001. Homology modelling of the human eukaryotic initiation factor 5A (eIF-5A).
528 Protein Engineering 14, 881-890.

529 Fields, B.A., Malchiodi, E.L., Li, H., Ysern, X., Stauffacher, C.V., Schlievert, P.M.,
530 Karjalainen, K., Mariuzza, R.A., 1996. Crystal structure of a T-cell receptor β-chain complexed
531 with a superantigen. Nature 384, 188-192.

532 Fuller-Espie, S., Hoffman Towler, P., Wiest, D.L., Tietjen, I., Spain, L.M., 1998.
533 Transmembrane polar residue of TCR β chain are required for signal transduction. Inter. Immunol.
534 10, 923-933.

535 Gagnon, S.J., Borbulevvch, O.Y., Davis-Harrison, R.L., Turner, R.V., Damirjian, M.,
536 Wojnarowicz, A., Biddison, W.E., Baker, B.M., 2006. T cell receptor recognition via cooperative
537 conformational plasticity. J. Mol. Biol. 363, 228-243.

538 Garcia, K.C., Degano, M., Stanfield, R.L., Brunmark, A., Jackson, M.R., Peterson, P.A.,
539 Teyton, L., Wilson, I.A., 1996. An αβ T cell receptor structure at 25 Å and its orientation in the
540 TCR-MHC complex. Science 274, 209-219.

541 Gianfrani, C., Siciliano, R., Facchiano, A.M., Camarca, A., Mazzeo, F.M., Costantini, S.,
542 Savati, V.M., Maurano, F., Mazzarella, G., Iaquinto, G., Bergamo, P., Rossi, M., 2007.
543 Transamidation of wheat flour inhibits the response to gliadin of intestinal T cells in celiac disease.
544 Gastroenterology, in press.

545 Hein, W.R., 1994. Structural and functional evolution of the extracellular regions of T cell
546 receptors. Semin. Immunol. 6, 361-372.

547 Hordvik, I., Jacob, A.L.J., Charlemagne, J., Endresen, C., 1996. Cloning of T-cell antigen
548 receprot beta chain cDNAs from Atlantic salmon (*Salmo salar*). Immunogenetics 45, 9-14.

549 Hordvik, I., Torvund, J., Moore, L., Endresen, C., 2004. Structure and organisation of the T cell
550 receptor alpha chain genes in Atlantic salmon. Mol. Immunol. 41, 553-559.

551 Hubbard, S.J., Campbell, S.F., Thornton, J.M., 1991. Molecular recognition. Conformational
552 analysis of limited proteolytic sites and serine proteinase protein inhibitors. J. Mol. Biol. 220, 507-
553 530.

554 Imai, E., Ishikawa, J., Moritomo, T., Tomana, M., 2005. Characterisation of T cell antigen
555 receptor α chain isotypes in the common carp. Fish Shellfish Immunol. 19, 205-216.
556 Jones, S., Thornton, J.M., 1996. Principles of protein-protein interactions derived from
557 structural studies. Proc. Natl. Acad. Sci. USA 93, 13-20.
558 Julenius, K., Molgaard, A., Gupta, R., Brunak, S., 2005. Prediction, conservation analysis and
559 structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology 15, 153-
560 164.
561 Kabsch, W, Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of
562 hydrogen-bonded and geometrical features. Biopolymers 22, 2577-2637.
563 Kamper, S.M., Churchill McKinney, E.C., 2002. Polymorphism and evolution in the constant
564 region of the T-cell receptor beta chain in an advanced teleost fish. Immunogenetics 53, 1047-1054.
565 Klausner, R.D., Lippincott-Schwartz, J., Bonifacino, J.S., 1990. The T-cell antigen receptor:
566 insights into organelle biology. Annu. Rev. Cell Biol. 6, 403-431.
567 Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated Software for Molecular
568 Evolutionary Genetics Analysis and Sequence Alignment. Briefings in Bioinfiormatics 5, 150-163.
569 Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK - A
570 program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283-291.
571 Li, H., Lebedeva, M.I., Llera, A.S., Fields, B.A., Brenner, M.B., Mariuzza, R.A., 1998.
572 Structure of the Vδ domain of a human γδ T-cell antigen receptor. Nature 291, 502-506.
573 Liu, S., Zhang, C., Zhou, H., Zhou, Y., 2004. A physical reference state unifies the structure-
574 derived potential of mean force for protein folding and binding. Proteins 56, 93-101.
575 Marchalonis, J.J., Jensen I., Schulter S.F., 2002. Structural, antigenic and evolutionary analyses
576 of immunoglobulins and T cell receptors. J. Mol. Recognit. 15, 260-271.
577 McDonald, I.K., Thornton, J.M., 1994. Satisfying hydrogen bonding potential in proteins. J.
578 Mol. Biol. 238, 777-793.

579 Miracle, A.L., Anderson, M.K., Litman, R.T., Walsh, C.J., Luer, C.A., Rothenberg, E.V.,
580 Litman, G.W., 2001. Complex expression patterns of lymphocyte-specific genes during the
581 development of cartilaginous fish implicate unique lymphoid tissues in generating an immune
582 repertoire. Intern. Immunol. 13, 567-580.

583 Morley, B.J., Chin, K.N., Newton, M.E., Weiss, A., 1988. The lysine residue in the membrane-
584 spanning domain of the β chain is necessary for cell surface expression of the T-cell antigen
585 receptor. J. Exp. Med. 168, 1971-1978.

586 Morrison, R.N., Lyons, A.B., Nowak, B.F., Hayball, J.D., 2004. Snapper (*Pagrus auratus*)
587 leucocyte proliferation is synergistically enhanced by simultaneous stimulation with LPS and PHA.
588 Fish Shellfish Immunol. 16, 307-319.

589 Moss, P.A., Bell, J.I., 1995. Sequence analysis of the human αβ T-cell receptor CDR3 region.
590 Immunogenetics 42, 10-18.

591 Nam, B-H., Hirono, I., Aoki, T., 2003. The four TCR genes of teleost fish: the cDNA and
592 genomic DNA analysis of Japanese flounder (Paralichthys olivaceus) TCR α-, β-, γ-, and δ-chains.
593 J. Immunol. 170, 3081-3090.

594 Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and
595 eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1-6.

596 Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B., Thornton, J.M., 1997.
597 CATH- A Hierarchic Classification of Protein Domain Structures. Structure 5, 1093-1108.

598 Partula, S., De Guerra, A., Fellah, J.S., Charlemagne, J., 1995. Structure and diversity of the T
599 cell antigen receptor β-chain in a teleost fish. J. Immunol. 155, 699-706.

600 Pearl, F.M.G, Lee, D., Bray, J.E, Sillitoe, I., Todd, A.E., Harrison, A.P., Thornton, J.M.,
601 Orengo, C.A., 2000. Assigning genomic sequences to CATH. Nucl. Acids Res. 28, 277-282.

602 Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. Proc.
603 Natl. Acad. Sci. USA 85, 2444-2448.

604 Romano, N., Rossi, F., Abelli, L., Caccia, E., Piergentili, R., Mastrolia, L., Randelli, E.,
605 Buonocore, F., 2007. Majority of TcRβ(+) T-lymphocytes located in the thymus and midgut of the
606 bony fish, *Dicentrarchus labrax* (L.). Cell Tissue Res., in press.
607 Rombout, J.H., Huttenhuis, H.B., Picchietti, S., Scapigliati, G., 2005. Phylogeny and ontogeny
608 of fish leucocytes. Fish Shellfish Immunol. 19, 441-455.
609 Rudolph, M.G., Stanfield, R.L., Wilson, I.A., 2006. How TCRs Bind MHCs, peptides and
610 coreceptors. Annu. Rev. Immunol. 24, 419-466.
611 Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial
612 restraints. J. Mol. Biol. 234, 779-815.
613 Scapigliati, G., Buonocore, F., Bird, S., Zou, J., Pelegrin, P., Falasca, C., Prugnoli, D.,
614 Secombes, C.J., 2001. Phylogeny of cytokines: molecular cloning and expression analysis of sea
615 bass *Dicentrarchus labrax* interleukin-1 beta. Fish Shellfish Immunol. 11, 711-726.
616 Scapigliati, G., Costantini, S., Colonna, G., Facchiano, A., Buonocore, F., Bossù, P., Holland,
617 J.W., Secombes, C.J., 2004. Modelling of fish interleukin 1 and its receptor. Dev. Comp. Immunol.
618 28, 429-441.
619 Sippl, M.J., 1993. Recognition of errors in three-dimensional structures of proteins. Proteins 17,
620 355-362.
621 Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Clustal W: Improving the sensitivity of
622 progressive multiple sequence alignment through sequence weighting, position-specific gap
623 penalties and weight matrix choice. Nucl. Acids Res. 22, 4673-4680.
624 Tough, D.F., Sun, S., Sprent, J., 1997. T cell stimulation *in vivo* by lipopolysaccharide (LPS). J.
625 Exp. Med. 185, 2089-2094.
626 Wermenstam, N.E., Pilstrom, L., 2001. T-cell antigen receptors in Atlantic cod (*Gadus morhua*
627 L.): structure, organisation and expression of TCR α and β genes. Dev. Comp. Immunol. 25, 117-
628 135.

629 Wilson, I.A., Garcia, K.C., 1997. T-cell receptor structure and TCR complexes. Curr. Opin. 630 Struct. Biol. 7, 839-848.

631 Wilson, M.R., Zhou, H., Bengten, E., Clem, L.W., Stuge, T.B., Warr, G.W., Miller, N.W., 1998. 632 T-cell receptors in channel catfish: structure and expression of TCR α and β genes. Mol. Immunol. 633 35, 545-557.

634 Zhou, H., Bengten, E., Miller, N.W., Clem, L.W., Wilson, M., 2003. The T cell receptor β locus 635 of the channel catfish, *Ictalarus punctatus*, reveals unique features. J. Immunol. 170, 2573-2581.

636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659

660 **FIGURE CAPTIONS**

661 **Figure 1.** Alignment of the predicted sea bream TcRβ amino acid sequences with other known
662 TcRβ molecules. Regions corresponding to the putative signal peptide, variable region,
663 diversity and joining region, constant Ig domain, connecting peptide (CPS),
664 transmembrane region (TM) and cytoplasmic tail (CYT) are shown above the sequences.
665 Conserved cysteine residues are evidenced in bold, conserved amino acid residues are
666 indicated with an asterisk, the putative N-glycosylation sites are in bold and underlined.
667 Accession numbers: *Sparus aurata* (sea bream) clone 1 AM261209; *Sparus aurata* clone
668 2 AM261210; *Sparus aurata* clone 3 AM490435; *Sparus aurata* clone 4 AM490436;
669 *Sparus aurata* clone 5 AM490437; *Sparus aurata* clone 6 AM490438; *Gadus morhua*
670 (Atlantic cod) AJ133848; *Stegastes partitus* (bicolor damselfish) AAG46047;
671 *Paralichthys olivaceus* (bastard halibut) AB053414; *Oncorhynchus mykiss* (rainbow
672 trout) AF329700; *Anas platyrhynchos* (mallard) AY039002; *Rattus norvegicus* (Norway
673 rat) BC099166; *Mus musculus* (house mouse) AY188690; *Homo sapiens* (human)
674 DQ341458.

675 **Figure 2.** Phylogenetic tree showing the relationship between the six sea bream TcRβ sequences
676 and other known TcRβ molecules. The tree was constructed by the "neighbour-joining"
677 method and was bootstrapped 10000 times. 0.1 indicates the genetic distance.

678 **Figure 3.** Sea bream TcRβ basal expression in different tissues. TcRβ mRNA levels were expressed
679 as a ratio relative to β-actin levels in the same samples after real-time PCR analysis
680 using the tissue with the lowest expression (head kidney) as calibrator.

681 **Figure 4.** *In vitro* sea bream TcRβ expression analysis. LPS: TcRβ mRNA levels expressed as a
682 ratio relative to β-actin levels in the same samples after real-time PCR analysis of HK
683 leucocytes stimulated with PBS (control) and with 5 μg/ml LPS for 4 and 24 h and
684 normalised against the non-stimulated controls. PHA: TcRβ mRNA levels expressed as

685 a ratio relative to β-actin levels in the same samples after real-time PCR analysis of HK

686 leucocytes stimulated with PBS (control) and with 1 μg/ml PHA-L for 4 and 24 h and

687 normalised against the non-stimulated controls. Controls for 4 and 24 h of incubation

688 with PBS only are also shown in the graphs. Data were expressed as the mean ± SD and

689 asterisks indicates when $p<0.05$ with respect to the time 0 control.

690 **Figure 5.** 3D model of *Sparus aurata* TcRβ clone 3. The backbone ribbon and secondary structure

691 topology are shown: yellow arrows represent β-strands. Amino and carboxy terminal

692 ends are indicated. Green and yellow sticks evidenced the presence of a putative

693 intrachain disulfide bond.

694 **Figure 6.** 3D model of sea bream class I MHC. The backbone ribbon and secondary structure

695 topology are shown: yellow arrows represent β-strands and red cylinders represent α-

696 helices. Green and yellow sticks indicate the presence of two putative intrachain disulfide

697 bonds.

698 **Figure 7.** The bars represent the binding energies (A) and the interaction energies (B) (expressed in

699 kcal/mol) evaluated for each simulated sea bream MHC/TCRβ complex and for the

700 experimental human and murine complexes.

701

702 **TABLE CAPTIONS**

703 **Table I.** Analysis of the MHC/TcRβ complexes in terms of interface surface area, intrachain H-

704 bonds and number of interaction residues.

705

706

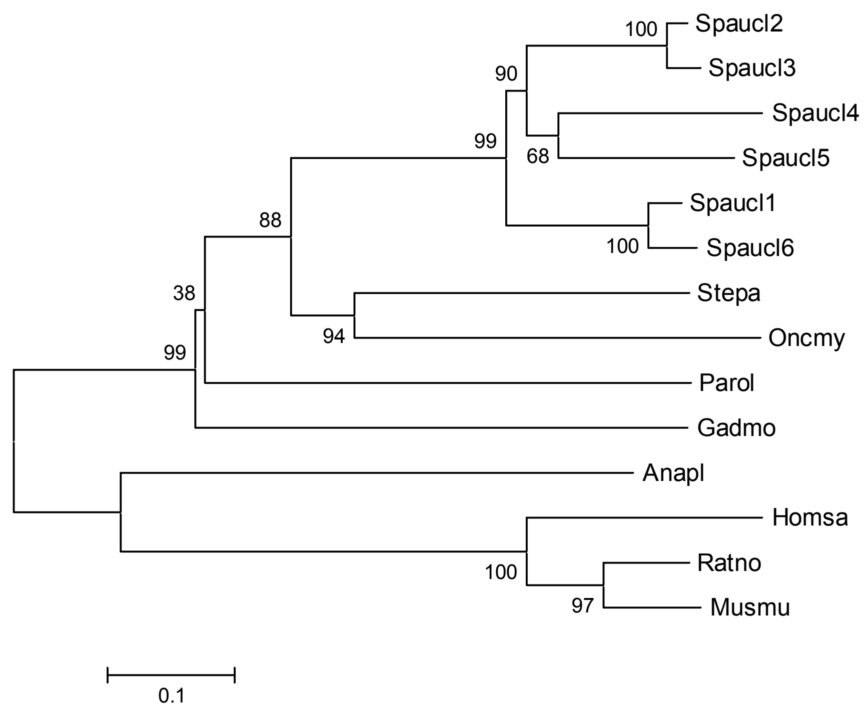
Figure

Figure 1

```
        <     signal peptide  ><                 variable region
Spaucl1 -MILIFLSIH LNNIL----V SGLSLNDQVH QTPADMFKRP GGEAKINCFH TI-QSYNRIL WYKQTNE--Q LQFLGYMNIN
Spaucl2 -.SSVMRIFG .LL.C----F YSQVSSVTFQ EF.LQIVNES S-.VQ.Q.S. DD-STKTQM. .FQ.KKDSLV .TLI..VYGQ
Spaucl3 -.SSLMRIFG .LL.C----F YSHVTAITFQ EF.PQIVSES S-.VH.Q.S. DD-TGKLQM. .FQ.KKDSLV .TLI.FVYGQ
Spaucl4 MIRN.YRIT- .LLLW----L .CR.QSV..D .G.SSILGS. N.S.T.S.N. SH-SLF.V.. ..Q.PTGDAG .KLI.LTQYT
Spaucl5 -.TAGLITLS AALLC----I I..IDGSD.T ...LLWMMEG Q-S.TM..S. .KGT..YEMY ..R.LPG-QR MKEIVFTTQS
Spaucl6 -......... .....----. .....SE... ........Q. .E...F.... ..-SN.DK.. .......--. ..........
Gadmo   -.CQFIC.LG IIFVC----F THEVQSVIIK .SS.KIVRKG AKGIQ.D.S. DD-S..PLMY ..QRKD.SPS .TLI.FGYES
Stepa   -.KHVLIITG .CFTFNIIL. ..S..S.K.D .A.T.IYGKQ .ET.E.T.S. K.-DN..... ....L.--RN ......L...
Parol   -..PSLNTLT FFVLW----A A.V.HSVLIT .W.HYISRF. S.S.EMH.YQ ND-TD.PYMY ..R.QRG-KE P.LVV.LAGS
Oncmy   -..R.LI..T MGYRA---WA A.S.PSN... .G...LY.NQ .EL..ME.S. S.-ST..V.. ....S.Y-RE .V.....QLK
Anapl   ------MGMW TAWCVATFFF GAR---AKIT ..S-SLVLKE D...TLK.SQ ND--NH.YMS ..L.QPG-KG ..L.Y.SIGA
Ratno   ---------- ---------- ----MEAA.T .S.RNRVAVT ...VTLS.RQ .N--NHDYMY ..R.NMG-DE .RLIH.SYDV
Musmu   ------MGSR .FFVV-LILL CAKHMEAA.T .S.RSKVAVT ..KVTLS.HQ .N--NHDYMY ..R.DTG-HG .RLIH.SYVA
Homsa   ------MT.R .LCYMGFYFL GAGLMEADIY ...RYLVIGT .KKITLE.SQ .M--GHDKMY ..Q.DPG-ME .HLIH.SYGV
                                                           *              *
                                ><    diversity and joining region               ><
Spaucl1 NGYPENGAG- -VKIEGNANT DQNCTLTIEE LKLS-SSAVY FCAART---- --GEYEAYFG PGTKLTVLGK DDKITPPT-V
Spaucl2 GSPTYE.EM- EKQFKLTKEG TTKG..I.HS A...-D.... ....TD---- -NNNQP.... Q......... ........-.
Spaucl3 GSPTYE.EM- EKQFKLTKEE TTKG..I.RR A...-D.... ....KA---G INQANP.... Q......... ........-.
Spaucl4 -SPTLEEPF- KEHFKVTGDG SSKSE.HVQ. .R.PED.GM. Y...PG---- ---G..PI.. G......V.. ........-.
Spaucl5 PPHQYESGFS TE.FPATKKD ALTGS..V.R .L.N-D.G.. ...VTS---- -G.NT..H.. G......... ........-.
Spaucl6 .........- -..V..G.KK .........R ...N-..... L....R---- --TGR..... G......... ........-.
Gadmo   STQNYEDRF- EERLNITRES VLQG..VLT. AAE.-D.... ....SMG--- EG.SEP.F.. K.......EP GCIVS...-.
Stepa   K....D.VDV --T.D.D..K GR......NS .SV.-..... ....SYGTG- GPQTEP.... K.......ET .RTV....K.
Parol   SANF.E.FK- -SGF.AEIVQ KKKWS.K.PS IQEK-DE... L...SH---- -RNTQP.... Q.......EP GQAVKS.I-.
Oncmy   T.F..V.FD- ---...D..A GGTS....KQ .TPN-..... Y...---TG- TKNYNP.F.. A.......DP NI.V.E..-.
Anapl   DQEAVGDTH- -PGYKATRLN LSDFH.V.KP V.MN-H..D. ...SSP---- NR.SNTQ... E...I...E. N.V.K..A-.
Ratno   .RTEKGDVS- -SGYKASRPS QEDFS.IL.S AS..-QT... ...S.D---- TRNT-.VF.. K..R...VED LKTV...K-.
Musmu   DSTEKGDIP- -DGYKASRPS QE.FS.IL.L AS..-QT... ...SSE---- NR.A-.QF.. ...R....ED LRNV...K-.
Homsa   .STEKGDLS- -SESTVSRIR TEHFP..L.S ARP.-HTSQ. L..SSV---- GGDQ-PQH.. D..R.SI.ED LN.VF..E-.
                                   *              *  **                **  **                 *
                                          constant Ig domain
Spaucl1 KVLEPSEKEC RNKVEKEKRK KTLLCVISRF YPDHVNVTWK INN------- -EEMSKGVAT DNMPAQP--- NDGKFYKITS
Spaucl2 .......... .......... .......... .......... ...------- -......... .......--- ..........
Spaucl3 .......... .......... .......... .......... ...------- -......... .......--- ..........
Spaucl4 .......... .......... .......... .......... ...------- -......... .......--- ..........
Spaucl5 .......... .......... .......... .......... ...------- -......... .......--- ..........
Spaucl6 .......... .......... .......... .......... ...------- -......... .......--- ..........
Gadmo   V..P...... .DR--..QL. ...V..A.G. .....G.S.T V.G------- -QSVI....S .HPALRV--- D.K--.Q...
Stepa   .IFP..A... ...K-DDI.. ...V..A.G. .....S.S.E K.GKVVPDSE AKDRQEKYGV ATDS.AK--- RV.E..R...
Parol   ..FR..S... ..PIDN.RE. -..V..A.D. .....S.Y.Q .IQLNVTSG- -VNVTR..T. .EAALR---- -KD.V.T...
Oncmy   ...A..A... EDR--NK.K. ...V...AT.. ......T.F.Q V..------- -VNRTE.AG. ..RALWD--- K..-L.S...
Anapl   AIFS..KQ.I QE.-----S. A..V.LA.G. ...TL.LV.. V.G------- -A.RTE..G. .ETSTS---- -YENT.SL..
Ratno   SLF....A.I AD.-----Q. A..V.LARG. F....ELS.W V.G------- -K.IRN..S. .PQAYK---E SNNIT.CLS.
Musmu   SLF...KA.I A..-----Q. A..V.LARG. F....ELS.W V.G------- -K.VHS..S. .PQAYK---E SN-YS.CLS.
Homsa   A.F....A.I SHT-----Q. A..V.LATG. F....ELS.W V.G------- -K.VHS..S. .PQ.LKEQPA LNDSR.CLS.
            **  *           *  ** *    *  **     *                                        *   *
                                                                  ><     CPS      ><      TM
Spaucl1 RLKVDANKWF DPENEFKCIA SFFNGTGTTY HENGTRG--- -----IEAPK TGQNITTEAY LKRSQTAKLS YGVLIIKGCV
Spaucl2 .......... .......... .......... .......--- -----..... .......... .......... ..........
Spaucl3 .......... .......... .......... .......--- -----..... .......... .......... ..........
Spaucl4 .......... .......... .......... .......--- -----..... .......... .......... ..........
Spaucl5 .......... .......... .......... .......--- -----..... .......... .......... ..........
Spaucl6 .......... .......... .......... .......--- -----..... .......... .......... ..........
Gadmo   ..R.E.R..Y TGG.I.T.NV .Y...ND.I. TSAEVY.GGD VRW--.KTEP D.E--.R.EF V.VT...... .I.M.V.NI.
Stepa   ..R.P.AHYN T.G.T.T..V ..Y...QNV- LRHASID--- --S--.KGES E.G-M.R.K. ..HT.S.... .....V.S.I
Parol   ....S.ED.Y K..WN.E..V R....HD.. .KDSIS.--- --------EQ GPDIL.R.K. .RITRQ.... .S.....SS.
Oncmy   ..R.P..E.H K...R.T..V ..YD..DNIR VT.D.ISG-- --D--LQGQS G.E-...DY. V.ST.....A .SIF.A.STF
Anapl   ..RISSQE.. N.L.R.E.V. N..KNG---- --------TQ ESIHRFIYGD A.CI.FK.N. QRSATAG.FL .IM..L.SIL
Ratno   ..R.S.PF.H N.R.H.R.QV Q.YGL.EEDN WSEDSPKPVT QNISAEAWGR ADCG..SAS. QQGVLS.TIL .EI..G.ATL
Musmu   ..R.S.TF.H N.R.H.R.QV Q.HGLSEEDK WPE.SPKPVT QNISAEAWGR ADCG..SAS. HQGVLS.TIL .EI.LG.ATL
Homsa   ..R.S.TF.Q N.R.H.R.QV Q.YGLSENDE WTQDRAKPVT QIVSAEAWGR ADCGF.SVS. QQGVLS.TIL .EI.LG.ATL
        **              * *                                                          *     *
                     ><  CYT  >
Spaucl1 YGAFVMFLVW KLPGSSGKRN N--
Spaucl2 .......... .......... .--
Spaucl3 .......... .......... .--
Spaucl4 .......... .......... .--
Spaucl5 .......... .......... .--
Spaucl6 .......... .......... .--
Gadmo   ..V..TI.A. ..GLGRSHAT AKK
Stepa   ....IG.... ..Q.....H. .--
Parol   .....A.... R.QS.TE.Q. ---
Oncmy   ..LV..VMI. .FQ...E.QI ---
Anapl   ..I...GMML RSK------- ---
Ratno   .AVL.ST..V MAMVKRKSS- ---
Musmu   .AVL.SG..L MAMVKKKNS- ---
Homsa   .AVL.SA..L MAMVKRKDF- ---
        *
```

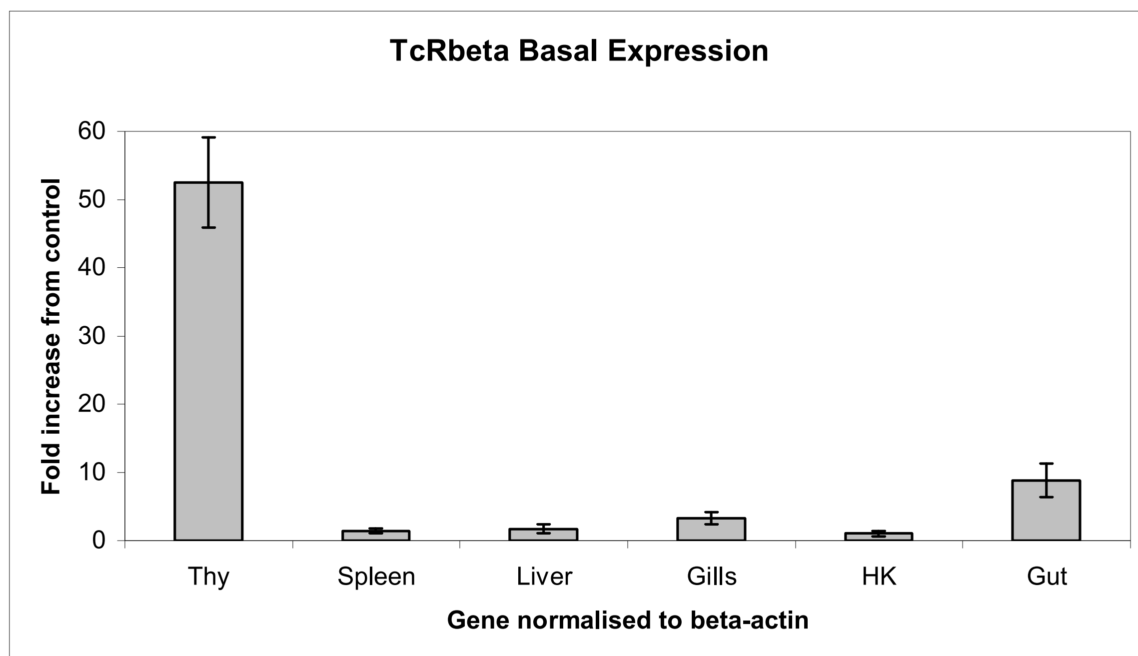
Figure

Figure 2



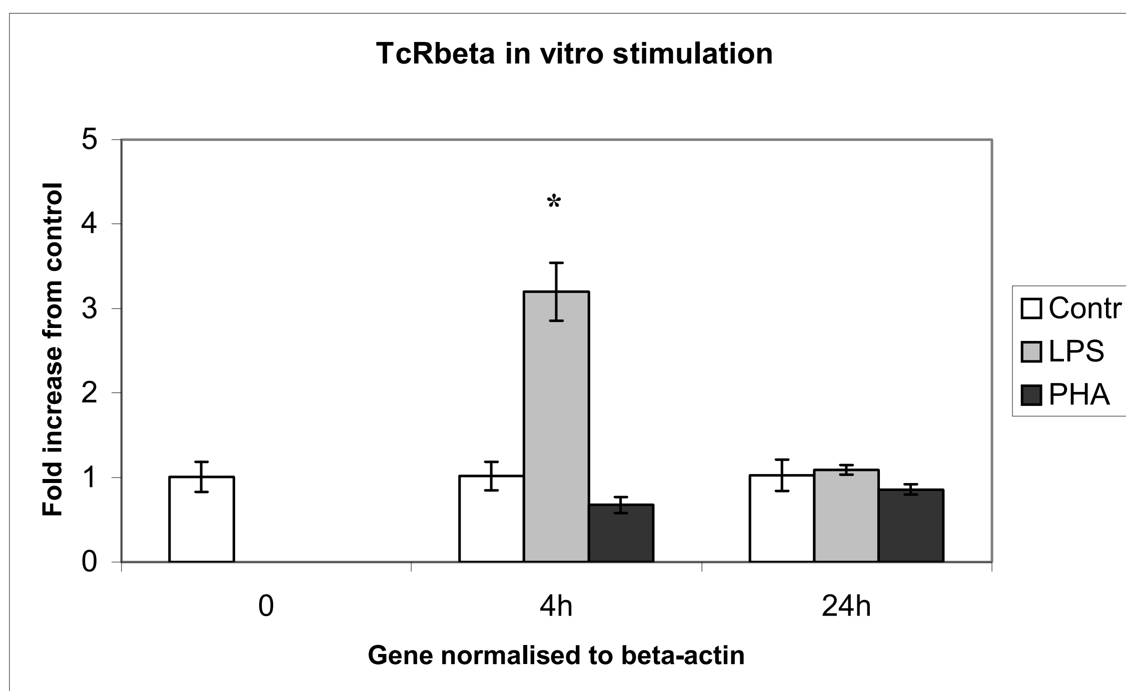
Figure

Figure 3



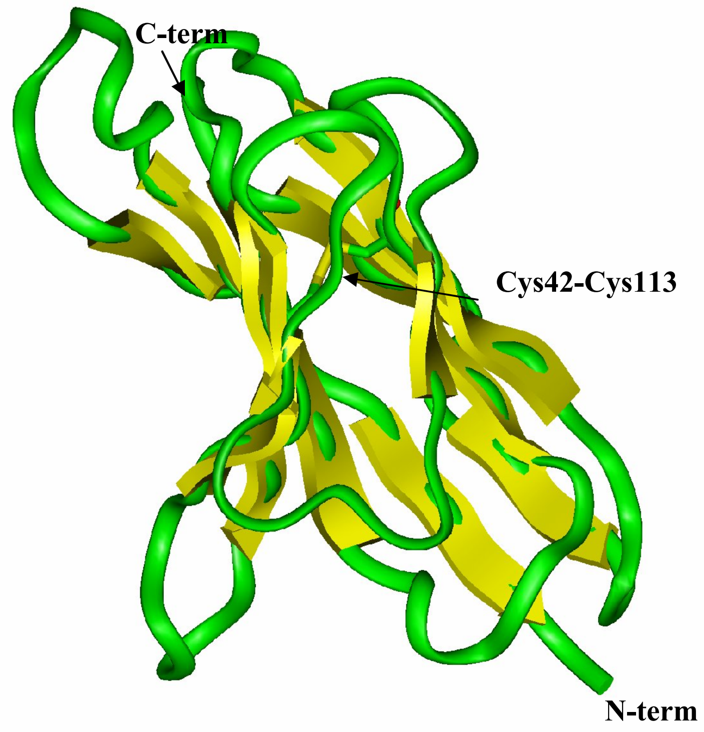
Figure

Figure 4



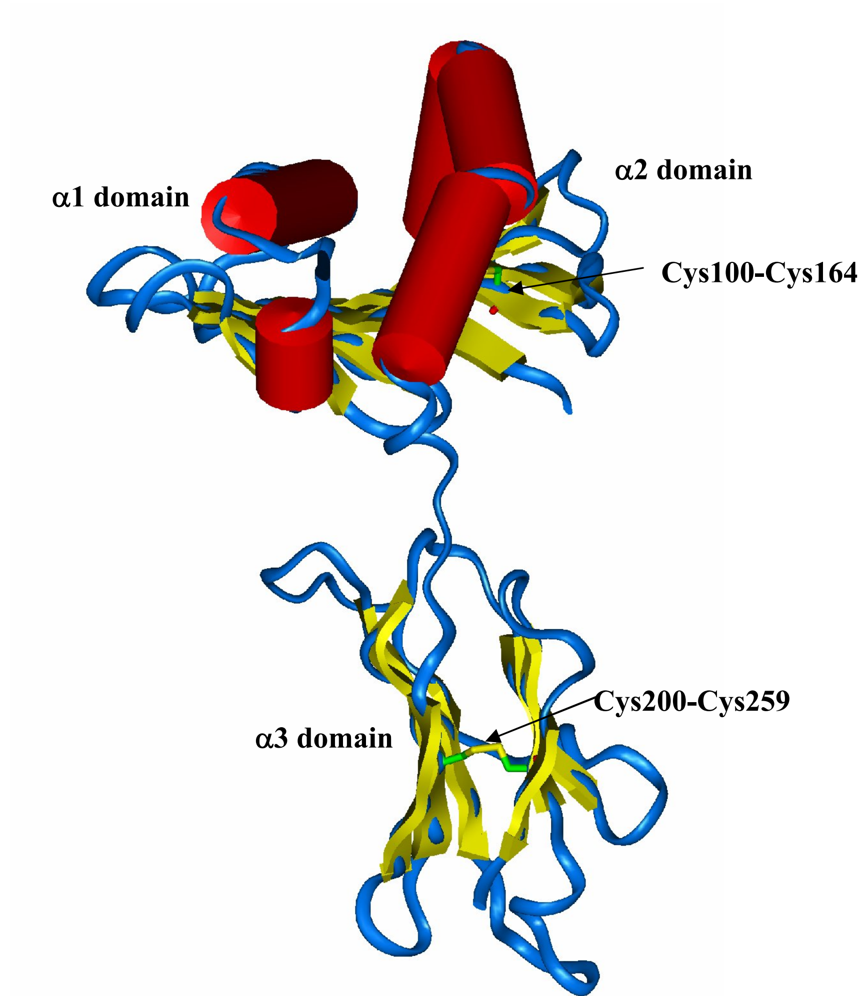
Figure

Figure 5



Figure

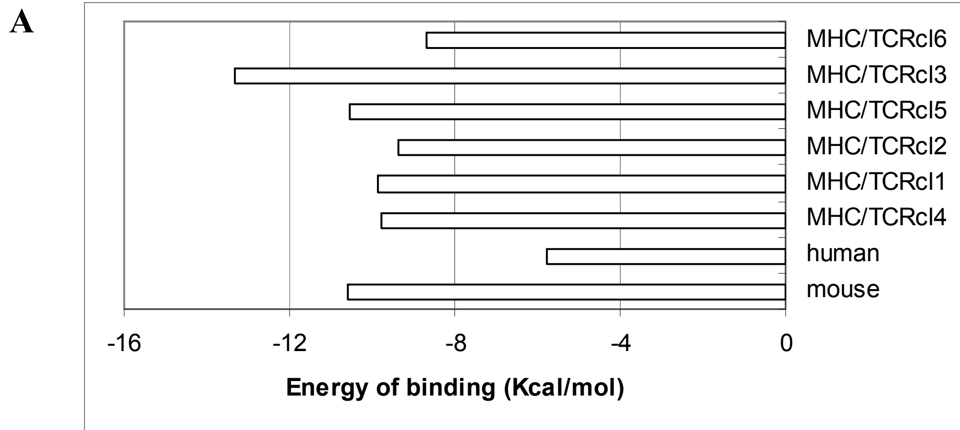
Figure 6



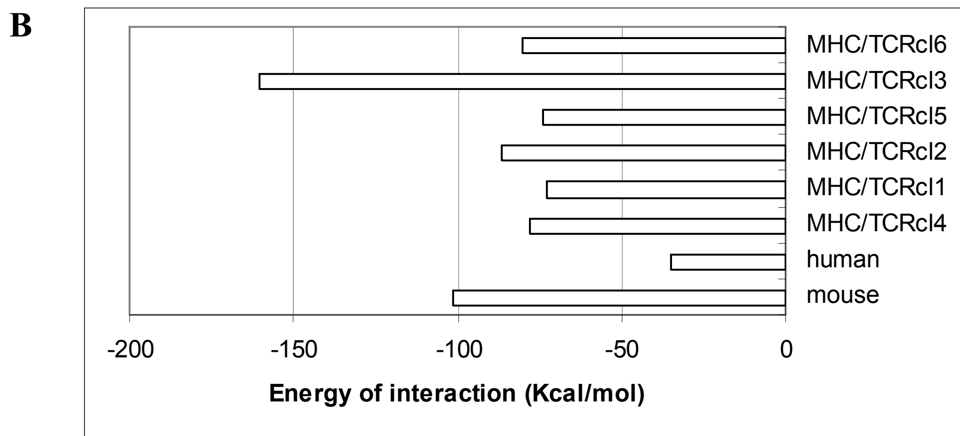
Figure

Figure 7

A



B



Table

Table I

	Interface Surface Area	Intrachain H-bonds	Interaction residues
MHC murine	490.06	3	15
TcRβ murine	500.53	3	18
MHC human	262.30	2	10
TcRβ human	300.76	2	9
MHC	642.91	3	18
TcRβ clone 4	649.22	3	16
MHC	672.71	2	19
TcRβ clone 1	670.04	2	18
MHC	726.4	6	19
TcRβ clone 2	682.4	6	19
MHC	622.88	4	18
TcRβ clone 5	640.17	4	18
MHC	1056.8	11	18
TcRβ clone 3	997	11	18
MHC	618.35	5	21
TcRβ clone 6	608.93	5	17