# 3-D modeling and molecular dynamics simulation of interleukin-22 from the So-iny mullet, *Liza haematocheila*

**Zhitao Qi[1]✉ · Rongwei Shi[2] · Suyang Li[1] · Xiayan Zang[3] · Zisheng Wang[1]**

1 Yancheng Institute of Technology, Department of Ocean Technology, Key Laboratory of Aquaculture and Ecology of Coastal Pool
   in Jiangsu Province, Yancheng, China

2 East China University of Science and Technology, School of Pharmacy, Shanghai, China

3 Henan Normal University, School of Life Science, Xinxiang, China

✉ Corresponding author: qizhitao@ycit.edu.cn

## Abstract

**Background:** Interleukin-22 (IL-22) plays an important role in the regulation of immune responses. However, little is known about its function or structure in fish.
**Results:** The IL-22 gene was first cloned from So-iny mullet (*Liza haematocheila*), one of commercially important fish species in China. Then, 3-D structure model of the mullet IL-22 was constructed by comparative modeling method using human IL-22 (1M4R) as template, and a 5 ns molecular dynamics (MD) was studied. The open reading frame (ORF) of mullet IL-22 cDNA was 555 bp, encoding 184 amino acids. The mullet IL-22 shared higher identities with the other fish IL-22 homologs and possessed a conserved IL-10 signature motif at its C-terminal. The mullet IL-22 model possessed six conserved helix structure. PROCHECK, SAVES and Molprobity server analysis confirmed that this model threaded well with human IL-22. Strikingly, analysis with CastP, cons-PPISP server suggested that the cysteines in mullet IL-22 might not be involved in the forming of disulfide bond for structural stabilization, but related to protein-protein interactions.
**Conclusions:** The structure of IL-22 in So-iny mullet (*Liza haematocheila*) was constructed using comparative modeling method which provide more information for studying the function of fish IL-22.

**Keywords:** comparative modeling, Interleukin-22, *Liza haematocheila*, molecular dynamics.

## INTRODUCTION

Interleukin-22 (IL-22) was firstly described and named as IL-10-related T cell-derived inducible factor (IL-TIF) by Dumoutier et al. (2000). It is a member of IL-10 family, which shared high sequence identity and structure homology with other IL-10 family members (IL-10, IL-19, IL-20 IL-24 and IL-26) (Qi and Nie, 2008). Many kinds of immune cells, including T helper ($T_H$) cells, natural killer (NK) cells, eosnophis, mastocyte and B cells had been found to be the producer of IL-22 (Gurney, 2004). The secreted IL-22 would bind with its specific receptor, IL-22 receptor 1 and IL-10 receptor 2, and activate the JAK-STAT pathway to fulfill its function (Lejeune et al. 2002). It has been found that many antimicrobial peptides (*e.g.* β-defensin) and pro-inflammatory cytokines (*e.g.* IL-6, IL-8) could be up-regulated by IL-22 (Liang et al. 2006; Zheng et al. 2007), supporting IL-22 play an important role in regulation of immune responses during inflammatory and bacterial invasion.

Except of mammalian IL-22, IL-22 had been cloned from many lower vertebrates, such as fugu (*Takifugu rubripes*), zebrafish (*Danio rerio*) (Igawa et al. 2006), rainbow trout (*Oncorhynchus mykiss*)

(Monte et al. 2011) and the amphibian *Xenopus tropicalis* (Qi and Nie, 2008). However, these researches mainly focused on the expression pattern of IL-22. The structure of IL-22 protein remains unstudied in any non-mammalian vertebrate.

So-iny mullet (*Liza haematocheila*) is an economically important fish in China, especially in Jiangsu Province. In 2011, the total yield of this species in Jiangsu province was about 50000 tons. Recently, we have cloned and sequenced the mullet IL-22 (GenBank accession number: JF960524) and analyzed its expression pattern during bacterial infection. Our previous studies confirmed that mullet IL-22, similar to mammalian IL-22, played some anti-inflammatory roles in bacterial infection. To further understand the properties and functions of the mullet IL-22, in this report we construct its 3-D structure based on comparative modeling method. Additionally, the molecular dynamics stimulation was also carried out at 300 K to reveal the stability of model of mullet IL-22.

**MATERIALS AND METHODS**

**Software and server**

Computational studies were carried out on an Inspur, 12 GHz PC equipped with the Red hat 6.0 environment. Sequences alignment was performed using Clustal W (Thompson et al. 1994) and was decorated with Genedoc (Nicholas et al. 1997). Comparative modeling was generated using the SWISS-Model software. Molecular dynamics stimulation, energy minimization, and trajectory analysis were performed using GROMACS (Groningen Machine for Chemical Simulations) 4.5.5 (Van Der Spoel et al. 2005). The model was validated by PROCHECK (Laskoswki et al. 1993) and Molprobity (Davis et al. 2004). Protein visualization and superimposition was performed using DeepView software. Pairwise comparison of protein structures were done at Dali server. Pocket information for mullet IL-22 was predicted by CASTp (Dundas et al. 2006) and cons-PPISP (Chen and Zhou, 2005) to locate the binding-site amino acids.

**Modeling of mullet IL-22**

The partial cDNA sequence of IL-22 of mullet was cloned by using degenerate primers and the full length of mullet IL-22 was obtained by RACE (5' and 3') method. The mullet IL-22 amino acid sequence was deduced using the Expasy translation tool. The signal peptide was predicted using SignalP 3.0 (Bendtsen et al. 2004) and was excluded from the model building. The mullet IL-22 except of signal peptide was used for template searching by GeneSilico Metaserver (Kurowski and Bujnicki, 2003), Pcons.net software (Wallner and Elofsson, 2005). Meanwhile, a sequence-structure comparison between the mullet IL-22 and template was also investigated using FUGUE (Find Homologs of Uncharacterized Gene Products Using Environment-specific substitution tables) program (Shi et al. 2001).

**Molecular dynamics stimulation**

MD simulations were conducted for the modeled system and crystal structure (PDB code: 1M4R) in explicit solvent using the GROMACS 4.5.5 package. The model and crystal structure were solvated by 6942 and 7679 water molecules in an octahedral box with edges that were 1.0 nm from the molecular boundary, respectively. To obtain a neutral system, three $NA^+$ ions were added (charge +3.00) to the mullet IL-22 model (which has a net negative charge of -3.00). The solvated systems were then subjected to further energy minimizations (maximum number of steps: 4000) to remove steric conflicts between the protein and water molecules, using the steepest descent integrator. Convergence was achieved in the energy minimization when the maximum force was smaller than 1000 kJ $mol^{-1}$ $nm^{-1}$. The energy-minimized models were subjected to position-restrained MD under NPT conditions, keeping the number of particles (N), the system pressure (P) and the temperature (T) constant. This was carried out for 100,000 steps for a total of 200 ps. The reference temperature for coupling (via V-rescale temperature coupling) was 300 K, and a pressure of 1 atm was maintained by the Parrinello–Rahman algorithm. Snapshots of the trajectory were taken every 1 ps. The final MD of 2,500,000 steps was carried out for 5,000 ps (5 ns) using the particle mesh Ewald (PME) electrostatics method under NPT conditions.

## RESULTS AND DISCUSSION

### Mullet IL-22 sequence analysis

The obtained mullet IL-22 cDNA was 1070 bp in length and contained a 5' un-translated region (UTR) of 150 bp, an open reading frame (ORF) of 555 bp and a 3' UTR of 365 bp. There are six mRNA instability motifs (ATTTA) in the 3' UTR and a polyadenylation signal (AATAAA) at 14 bp upstream of the polyA tail. The mullet IL-22 cDNA was translated into 184 amino acids (aa) with a 36 aa signal peptide predicted using SignalP 3.0. Thus, the mature mullet IL-22 molecule was 148 aa. Multiple alignment of the mullet IL-22 with other known IL-22 molecules showed a reasonable level of conservation. The conserved IL-10 family signature (G-X2-KA-X2- [D,E]-X-D[ILV]-[FLY]-[FILMV]-X2-[ILMV][EKQR]) (Qi and Nie, 2008), was found at the 3' end of the molecules. The phylogenetic tree analysis showed that the mullet IL-22 clustered well with teleost IL-22 (data not shown). All these analyses confirmed the sequence we obtained was exactly the IL-22 homologue. Meanwhile, the mullet IL-22 shared 22.5% identity and 50.5% similarity with human IL-22 which suggested that the human IL-22 could be used as template for constructing the 3-D structure of mullet IL-22 by comparative modeling method (Sali and Blundell, 1993).

### Model building

The model of mullet IL-22 was performed using comparative modeling method. By using sequence comparison and functionally conserved domain search methods, the human IL-22 crystal structure (PDB code: 1M4R) at 2.0 Å resolution (Nagem et al. 2002) was evaluated as the best template for modeling. The Pdbblast, FFA3 and Sp³ program showed that the mullet IL-22 shared ~28% identity with 1M4R. In FUGUE analysis, the top Z-score against the cut-off score (Z-score > 6.0) was considered to be the best template for modeling. Our result showed that mullet IL-22 shared the highest Z-Score of 16.64 with 1M4R, followed by 2ILK (*Homo sapiens* IL-10) with a Z-score of 16.25. Those results confirmed that the 1M4R could be used as template for modeling mullet IL-22.

The secondary structure of the mullet IL-22 amino acid sequence was predicted by PSIPRED (McGuffin et al. 2000). The PSIPRED verified the helices, coils and strands by calculating hydrogen-bond energies and main chain dihedral angles. The result showed that the mullet IL-22 contained the six conserved α-helices structures, similar to the structure of human IL-22. The comparison of 2-D structure between mullet IL-22 and human IL-22 also confirmed that the human IL-22 (1M4R) was a good choice for using as template (Figure 1).

### Validation of mullet IL-22 model

To validate the mullet IL-22 model, a Ramachandran plot was drawn by PROCHECK. The result showed that the phi-psi angles of 88.2% (105/134) of the residues were in favored regions, 7.6% (9/134) were in additionally allowed regions, 4.2% (5/134) were in generously allowed regions, and no residue was in the disallowed regions (Figure 2a). Pairwise structural alignment was performed in the Dali server (Holm and Sander, 1995) and the result showed the highest Z-score value (23.9) between the mullet IL-22 model and 1M4R. The model of mullet IL-22 was also checked in SAVES server. The verifying 3-D analysis showed that 71.54% of the residues had an averaged 3D-1D score of >0.2. ERRAT program showed the overall quality factor of mullet IL-22 model were 97.521 (a value of ~95% shows high resolution) (Figure 2b). All those results indicated that the mullet IL-22 protein model was valid. The MolProbity server predicted that residues with bad bonds were 0%, residues with bad angles were <0.1%, Cβ deviations >0.25 Å was 0%, and ramachandran favoured residues were >98%, which further confirmed the reliability of mullet IL-22 model.

The homology model of mullet IL-22 was compatible with the six α-helices model of the human monomer (Figure 3a). The overlaying of mullet and human IL-22 showed that there were low deviations between the mullet and human IL-22. The helix A, C and E is shorter, while the AB and DE loop is longer in mullet IL-22 (Figure 3b).

**Molecular dynamics simulation**

The stability and properties of the structure of the homolog was studied by explicit solvent MD stimulation. The root mean square deviation (RMSD) analysis not only reflects the change of protein backbone versus simulation time, but also indicates the divergence of two structures. The RMSD of homolog became stable at 2.5 ns. The RMSD values of human IL-22 and mullet IL-22 were 0.19 nm and 0.25 nm, and their standard deviations were 0.028 nm and 0.039 nm, respectively (Figure 4). This result indicated that an accepted structure was obtained by the simulation and it was reliable for further analyses.

The root mean square fluctuation (RMSF) reflects the mobility of a certain residue around its mean position, which is another tool for studying the dynamics stability of the system (Cheng et al. 2012). The mean RMSFs of human IL-22 and mullet IL-22 were 0.115 nm and 0.127 nm, and their standard deviations were 0.04 nm and 0.05 nm, respectively. Although there are some deviations among the trajectories, those data suggest that minute fluctuations highlighted the reliability of the structure (Figure 5).

Using the CASTp server, 27 pockets were predicted in the mullet IL-22. The largest pocket was formed by the amino acids in loop AB, and helix F with 225.4 areas and 214 volumes. Results of the 1M4R analyzed using CASTp server showed that there were 20 pockets in 1M4R and the largest pocket was formed by the amino acids in helix A, loop AB and helix F with 591.5 areas and 713.4 volumes. These analysis indicated these pockets might be the domain for protein interactions (Jones and Thornton, 1996). Interestingly, the cysteine residues that found in mullet and human IL-22 were predicted to form the pockets (Figure 6a, b). Further analysis in cons-PPISP server showed that these cysteine residues all occurred in the clusters that may take part in the protein interaction. These data suggested those cysteine residues might not be involved in the forming of disulfide bond.

**CONCLUDING REMARKS**

IL-22, one of members of IL-10 family, play a crucial role in regulation of immune response, especially in pro-inflammation. In this study, the obtained mullet IL-22 shared higher identities with its homologs from other fish species and possessed the IL-10 signature motif at its C-terminal. The 3-D modeling data showed that mullet IL-10 showed some structural differences from human IL-10, especially the cysteine residues. Further studies showed that the cysteine residues in mullet did not form any significant bonds involved in structural stabilization, but was related to protein-receptor interaction. So we speculated that the cysteine residues in higher vertebrates were formed by mutation during the evolution. Further studies are required to investigate those results.

**REFERENCES**

BENDTSEN, J.D.; NIELSEN, H.; HEIJNE, G.V. and BRUNAK, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology*, vol. 340, no. 4, p. 783-795. **[CrossRef]**
CHEN, H.L. and ZHOU, H.X. (2005). Prediction of interface residues in protein-protein complexes by a consensus neural network method: Test against NMR data. *Proteins: Structure, Function and Bioinformatics,* vol. 61, no. 1, p. 21-35. **[CrossRef]**
CHENG, W.Y.; CHEN, J.Z.; LIANG, Z.Q.; LI, G.H.; YI, C.H.; WANG, W. and WANG, K.Y. (2012). A computational analysis of interaction mechanisms of peptide and non-peptide inhibitors with MDMX based on molecular dynamics simulation. *Computational and Theoretical Chemistry*, vol. 984, p. 43-50. **[CrossRef]**
DAVIS, I.W.; MURRAY, L.W.; RICHARDSON, J.S. and RICHARDSON, D.C. (2004). MolProbity: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research*, vol. 32, no. 2, p. 615-619. **[CrossRef]**

DUMOUTIER, L.; LOUAHED, J. and RENAULD, J.C. (2000). Cloning and characterization of IL-10-related T cell derived inducible factor (IL-TIF), a novel cytokine structurally related to IL-10 and inducible by IL-19. *Journal of Immunology*, vol. 164, no. 4, p. 1814-1819.

DUNDAS, J.; OUYANG, Z.; TSENG, J.; BINKOWSKI, A.; TURPAZ, Y. and LIANG, J. (2006). CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research*, vol. 34, no. 2, p. 116-118. **[CrossRef]**

GURNEY, A.L. (2004). IL-22, a Th1 cytokine that targets the pancreas and select other peripheral tissues. *International Immunopharmacology*, vol. 4, no. 5, p. 669-677. **[CrossRef]**

HOLM, L. and SANDER, C. (1995). Dali: A network tool for protein structure comparison. *Trends in Biochemical Sciences*, vol. 20, no. 11, p. 478-480. **[CrossRef]**

IGAWA, D.; SAKAI, M. and SAVAN, R. (2006). An unexpected discovery of two interferon gamma-like genes along with interleukin (IL)-22 and -26 from teleost: IL-22 and -26 genes have been described for the first time outside mammals. *Molecular Immunology*, vol. 43, no. 7, p. 999-1009. **[CrossRef]**

JONES, S. and THORNTON, J.M. (1996). Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 1, p. 13-20. **[CrossRef]**

KUROWSKI, M.A. and BUJNICKI, J.M. (2003). GeneSilico protein structure prediction meta-server. *Nucleic Acids Research*, vol. 31, no. 13, p. 3305-3307. **[CrossRef]**

LASKOSWKI, R.A.; MACARTHUR, M.W.; MOSS, D.S. and THORNTON, J.M. (1993). PROCHECK: A program to check the stereo chemical quality of protein structures. *Journal of Applied Crystallography*, vol. 26, no. 2, p. 283-291. **[CrossRef]**

LEJEUNE, D.; DUMOUTIER, L.; CONSTANTINESCU, S.; KRUIJER, W.; SCHURINGA, J.J. and RENAULD, J.C. (2002). Interleukin-22 (IL-22) activates the JAK/STAT, ERK, JNK and p38 MAP kinase pathways in a rat hepatoma cell line: Pathways that are shared with and distinct from IL-10. *Journal of Biological Chemistry*, vol. 277, no. 37, p. 33676-33682. **[CrossRef]**

LIANG, S.C.; TAN, X.Y.; LUXENBERG, D.P.; KARIM, R.; DUNUSSI-JOANNOPOULOS, K.; COLLINS, M. and FOUSER, L.A. (2006). Interleukin (IL) -22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *Journal of Experimental Medicine*, vol. 203, no. 10, p. 2271-2279. **[CrossRef]**

MCGUFFIN, L.J.; BRYSON, K. and JONES, D.T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, vol. 16, no. 4, p. 404-405. **[CrossRef]**

MONTE, M.M.; ZOU, J.; WANG, T.H.; CARRINGTON, A. and SECOMBERS, C.J. (2011). Cloning, expression analysis and bioactivity studies of rainbow trout (*Oncorhynchus mykiss*) interleukin-22. *Cytokine*, vol. 55, no. 1, p. 62-73. **[CrossRef]**

NAGEM, R.A.P.; COLAU, D.; DUMOUTIER, L.; RENAULD, J.C.; OGATA, C. and POLIKARPOV, I. (2002). Crystal structure of recombinant human interleukin-22. *Structure*, vol. 10, no. 8, p. 1051-1062. **[CrossRef]**

NICHOLAS, K.B.; NICHOLAS, H.B.J. and DEERFIELD, D.W.II. (1997). GeneDoc: Analysis and visualization of genetic variation. *EMBNEW NEWS*, vol. 4, p. 14.

QI, Z.T. and NIE, P. (2008). Comparative study and expression analysis of the interferon gamma gene locus cytokines in *Xenopus tropicalis*. *Immunogenetics*, vol. 60, no. 11, p. 699-710. **[CrossRef]**

SALI, A. and BLUNDELL, T.L. (1993). Comparative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology*, vol. 234, no. 3, p. 779-815. **[CrossRef]**

SHI, J.; BLUNDELL, T.L. and MIZUGUCHI, K. (2001). FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology*, vol. 310, no. 1, p. 243-257. **[CrossRef]**

THOMPSON, J.D.; HIGGINS, D.G. and GIBSON, T.J. (1994). Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, vol. 22, no. 22, p. 4673-4680. **[CrossRef]**

VAN DER SPOEL, D.; LINDAHL, E.; HESS, B.; GROENHOF, G.; MARK, A.E. and BERENDSEN, H.J.C. (2005). GROMACS: Fast, flexible and free. *Journal of Computational Chemistry*, vol. 26, no. 16, p. 1701-1718. **[CrossRef]**

WALLNER, B. and ELOFSSON, A. (2005). Pcons5: Combining consensus, structural evaluation and fold recognition scores. *Bioinformatics*, vol. 21, no. 23, p. 4248-4254. **[CrossRef]**

ZHENG, Y.; DANILENKO, D.M.; VALDEZ, P.; KASMAN, I.; EASTHAM-ANDERSON, J.; WU, J.F. and OUYANG, W.J. (2007). Interleukin-22, a Th17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis. *Nature*, vol. 445, no. 7128, p. 648-651. **[CrossRef]**
.

## How to reference this article:

# Figures



**Fig. 1 Secondary structure of mullet IL-22 predicated using PSIPRED.** H denotes a helical region and C denotes a coil region.
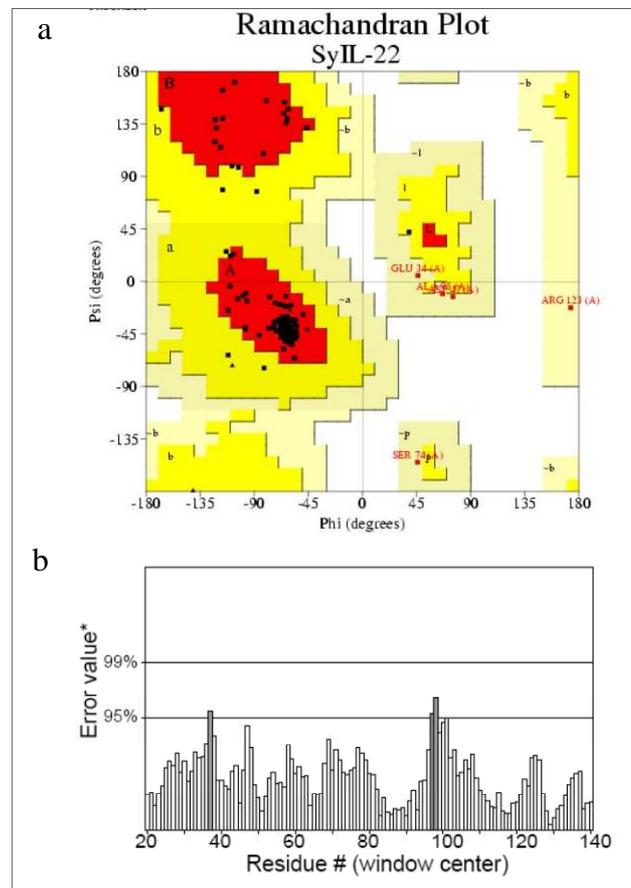
**Fig. 2 Validation of mullet IL-22 model.** (a) Ramachandran plot of the mullet IL-22 model. The plot was calculated using PROCHECK program. (b) The overall quality factor of mullet IL-22 model checked with ERRAT program.
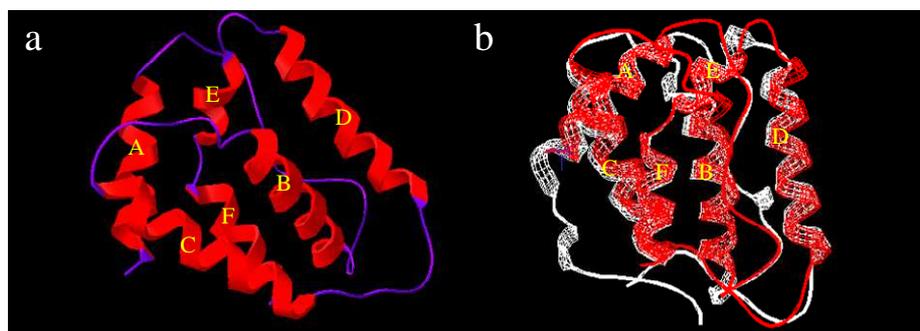


**Fig. 3 3-D structure of mullet IL-22 (a) and superimposition of the mullet IL-22 model onto the template (human IL-22, PDB ID: 1M4R) (b).** The mullet IL-22 is shown in red and the template are shown in white.
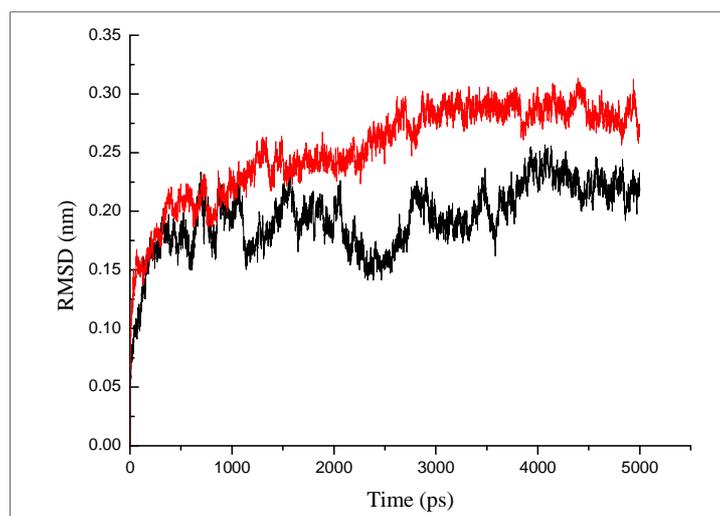
**Fig. 4 Comparison of the root mean square deviations (RMSDs) between human and mullet IL-22 versus simulation time.** The RMSDs of the Cα backbone atoms in human and mullet are shown in black and red, respectively.
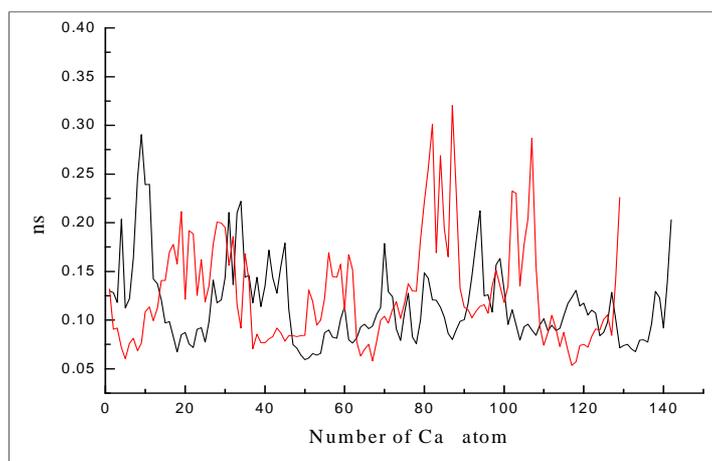


**Fig. 5 Comparison of the root mean square fluctuations (RMSFs) of human and mullet IL-22.** The RMSFs of the Cα backbone atoms in human and mullet IL-22 are shown in black and red, respectively.
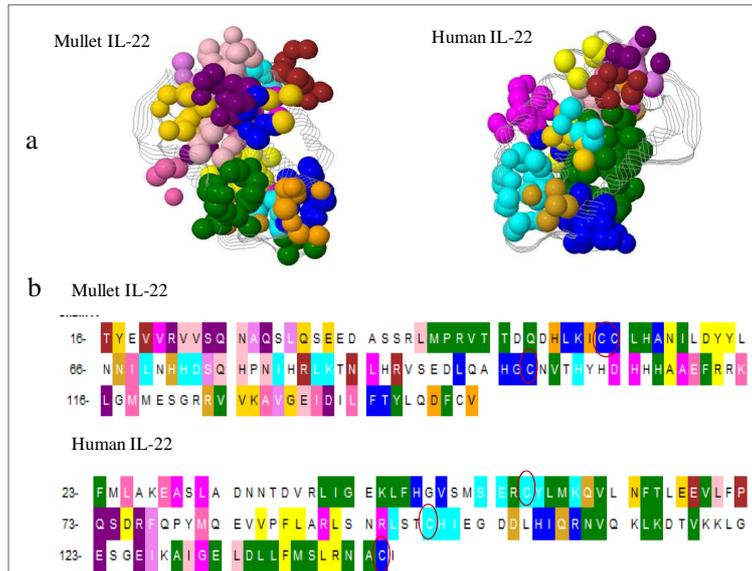
**Fig. 6 Pockets analysis using CASTp server.** (a) The predicted pockets in mullet and human IL-22. (b) Amino acids in mullet and human IL-22 predicted to form pockets. Amino acids with different colours denote pockets predicted by CASTp.