International Standard Serial Number (ISSN): 2249-6807

International Journal of Institutional Pharmacy and Life Sciences 2(2): March-April 2012

INTERNATIONAL JOURNAL OF INSTITUTIONAL PHARMACY AND LIFE SCIENCES

Life Sciences **Research Article……!!!**

Received: 10-04-2012; Accepted: 21-04-2012

IN SILICO CHARACTERIZATION OF TYROSINE HYDROXYLASE PROTEINS USING COMPUTATIONAL TOOLS AND SERVERS

Khanam N., Nath A., Chaube R*

Department of Zoology/Bioinformatics, Mahila Mahavidyalaya, Banaras Hindu University, Varanasi-221005, U.P., India

Keywords:
Tyrosine hydroxylase, Protein sequences, Computational analysis, Property groups of amino acids, secondary structure

For Correspondence:

Dr. Radha Chaube

Department of Zoology/Bioinformatics, Mahila Mahavidyalaya, Banaras Hindu University, Varanasi-221005, U.P., India

E-mail:

chauberadha@rediffmail.com

ABSTRACT

In this paper 14 tyrosine hydroxylase protein sequences of teleost fishes were retrieved from Uniprot Database which are analysed by our own Matlab codes and *In silico* tools. Primary structure analysis shows that most of the tyrosine hydroxylase protein sequences are hydrophobic in nature due to high content of non-polar residues. The aliphatic index was computed by ExPasy's ProtParam tool. Secondary structure analysis shows that most of the tyrosine hydroxylase have predominant α-helical structures and rest of the sequences have mixed secondary structure including strands, turn, beta-sheet. The very high coil structural content are due to rich content of more flexible glutamic acid, leucine, serine and hydrophobic proline amino acids. Proline has a special property of creating kinks in polypeptide chains and disrupting ordered secondary structure. SOSUI server predicted that there is no trans-membrane region in tyrosine hydroxylase sequences. All membrane is soluble. Cleavage site computed by peptide cutter tool. The presence of disulphide bonds (S-S) is identified by using Cys-Rec Tool and visualised the 3D structure by using chimera tool.

INTRODUCTION

Presently, number of computational tools have been developed for making prediction and characterization of protein sequences such as number of amino acid, sequence length, and the physico-chemical properties of proteins such as molecular weight, atomic weight, extinction coefficient, GRAVY, aliphatic index, cleavage site etc. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function, physical and chemical properties. Sequence analysis and physiochemical characterization of protein using bio-computational tools have been done by many researches and reported.

Tyrosine hydroxylase or tyrosine 3-monooxygenase is the enzyme responsible for catalyzing the conversion of the amino acid L-tyrosine to dihydroxyphenylalanine (DOPA). It does so using tetrahydrobiopterin as a coenzyme (1). DOPA is a precursor for dopamine, which, in turn, is a precursor for norepinephrine (noradrenaline) and epinephrine (adrenaline) (2). In humans, tyrosine hydroxylase is encoded by the TH gene (3). The enzyme, an oxygenase**,** is found in the cytosol of all cells containing catecholamines. This initial reaction is the rate limiting step in the production of catecholamines. The enzyme is highly specific, not accepting indole derivatives which are unusual as many other enzymes involved in the production of catecholamines (4, 5). Mutations in this gene have been associated with autosomal recessive Segawa syndrome. Tyrosine hydroxylase can be inhibited by the drug α-methyl-para-tyrosine (11). This inhibition can lead to a depletion of dopamine and norepinepherine in the brain due to the lack of the precursor L-Dopa (L-3, 4-dyhydroxyphenylalanine) which is synthesized by tyrosine hydroxylase. This drug is rarely used and can cause depression, but it is useful in treating pheochromocytoma and also resistant hypertension (7).

In this paper, we report the *In silico* analysis and characterization studies on 14 tyrosine hydroxylase sequences of teleost fishes. Data regarding the structure and function of TH protein including possible amino acids, hydrophobicity, molecular weight, pI, extinction coefficient (EC), instability index (II) have been used for creating cDNA probes for in situ hybridization and/ or Northern blotting, quantitative measurements of TH mRNA *in situ* and *in vitro* are widely used in neurological, pharmacological and biochemical analysis of TH induction. In a similar way present study will help in designing specific biomarkers for neurological disorders in lower and higher vertebrate groups.

MATERIAL AND METHODS

Tyrosine hydroxylase sequences

Sequences are retrieved from the manually curated public protein database Swiss-Prot. Swiss-Prot is scanned for the key word tyrosine hydroxylase. The search result yielded 21 TH sequences of fishes. From this, we have retrieved 14 sequences of different teleost fishes. The sequences were retrieved in FASTA format and used for analysis (table 1).

Table 1. Tyrosine hydroxylase sequences retrieved from Swiss-Prot database.

Accession Number	Sequence Description	Organism
B0W0A8	Tyrosine hydroxylase (Fragment)	Fathead minnow *(Pimephales promelas)*
O42091	Tyrosine 3-monooxygenase	Japanese eel *(Anguilla anguilla)*
Q6ZM47	Tyrosine hydroxylase (Fragment)	*Danio rerio*
Q6IVB6	Tyrosine hydroxylase 2	*Danio rerio*
B3DJW5	Tyrosine hydroxylase	*Danio rerio*
Q9W7K7	Tyrosine hydroxylase (Fragment)	*Danio rerio*
E9RK47	Tyrosine hydroxylase (Fragment)	*Anguilla japonica*
Q1PPW9	Tyrosine hydroxylase (Fragment)	*Acanthopagrus schlegelii*
Q6IVY6	Tyrosine hydroxylase 2	*Takifugu rubripes*
Q1LWZ5	Tyrosine hydroxylase 2	*Danio rerio*
Q6DUB2	Tyrosine hydroxylase (Fragment)	*Carassius auratus*
Q5RHI3	Novel protein similar to tyrosine hydroxylase (Th) (Fragment)	*Danio rerio*
B7U507	Tyrosine hydroxylase (Fragment)	*Epinephelus coioides*
B9W0A8	Tyrosine hydroxylase (Fragment)	*Pimephales promelas*
A0JGX6	Tyrosine 3-monooxygenase	*Takifugu rubripes*

Computational tools and servers

The amino acid composition of sequences were computed using by our own matlab codes. Percentage of hydrophobic and hydrophilic residues can be calculated from primary structure analysis. The physio-chemical parameters theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, half-life, Instability index, aliphatic index and grand average hydropathy (GRAVY) was computed using Expasy's Prot Param Prediction server.(36) The SOPMA tool (38) (Secondary Structural Content Prediction server) was used for the secondary structure prediction ((table 5). Cleavage Site computed using Peptide cutter Tool(37). The 3D structure of TH Q6IVB6 were generated by homology modelling using CPH server(39). The similar 3D structures (for the TH Q6IVB6) 2 in the Protein Data Bank (www.rscb.org) were identified by the BLASTP analysis. (http://www.ncbi.nlm.nih.gov:80/BLAST). Chimera was used to visualize 3D structures and to identify the SS bonds. The Cys-Rec Tool (22) is used for recognition of cysteine residues and -SS- bond. This Tool helps to identify number and position of cysteine bond in sequences. The three–dimensional structures of TH Q6IVB6 modelled using the PDB Blast template 1LFD are shown in figures 1, 2 and 3 respectively.

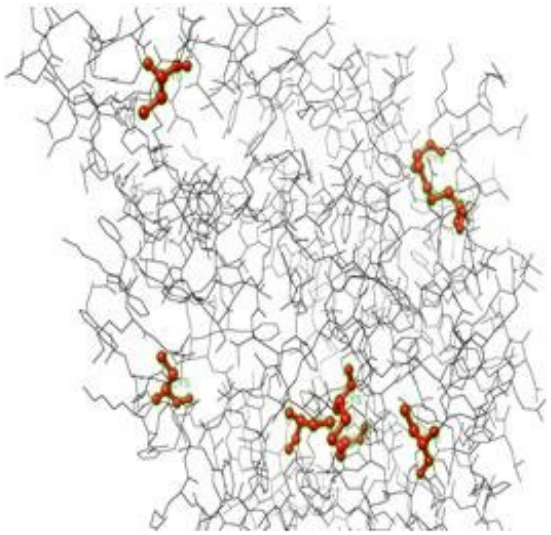

Fig 1. Chimera representation of the homology modelled 3D structure of tyrosine hydroxylase Q6IVB6 (using PDB template 1LFD). The cysteines are shown as ball and stick models (red)

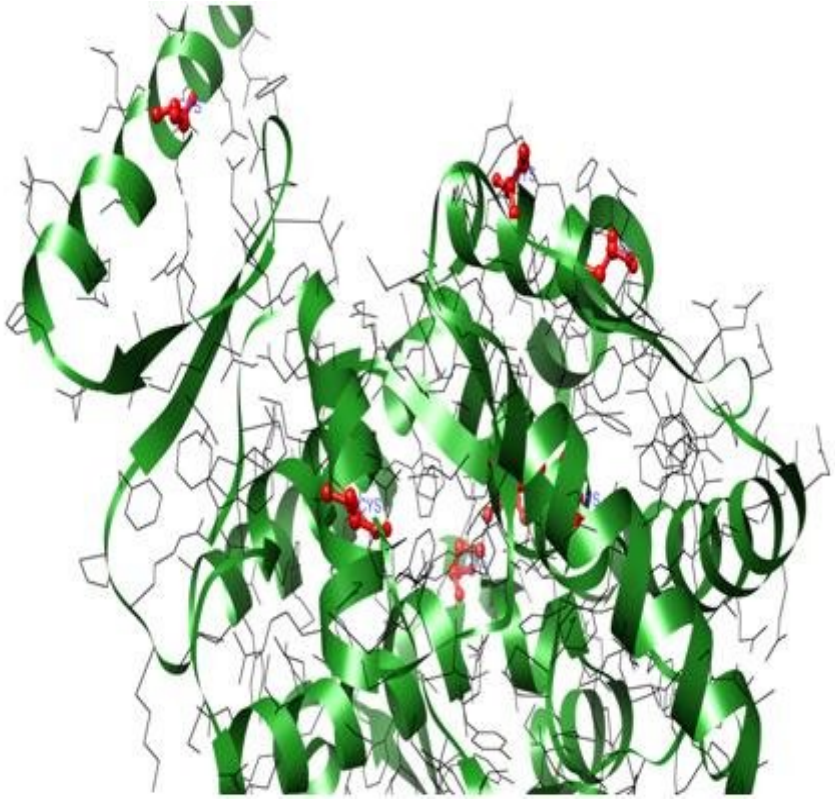

Fig 2. Chimera representation of the homology modelled 3D structure of Tyrosine hydroxylase Q6IVB6 (TH) enzyme (using PDB template 1LFD)

```
CYS_REC  Version 2. Recognition of SS-bounded cysteines

>tr|Q6IVB6|Q6IVB6_DANRE Tyrosine hydroxylase 2 OS=Danio rerio GN=th2 PE=2 SV=1
 length=471
MKSDSIAQNVPLSGRKRSLIEDARRDRESGSSSPGPSRCGDSFVFEEEKTGKITLNILFALKNEKNAGFFKAGKVFETFE
TKLLHLESRASRKSKRSSGEDLEFFMRCEVHCSDTDIFINSLKRVADDVRIVQEEKTPWFPKKISDLDTCNHLITKYDPD
LDQDHPGFSDSEYRKRRGFISEMAFNYKHGDPLPVVEYTPEENATWEVYRTLRSLYPSHACRQFLEALTQLESERLYGED
QIPQLREVSAFLKERTGFQLRPVAGLLSARDFLASLAFRVFQCTQYIRHPSAPMHSPEPDCCHELLGHIPMLADQEFAQF
SQIGLVSLGASDEDIEKLSTDYWFTVEFGLCKQNGAVKAYGAGLLSSYGELLYALSNEPQYKPFDPAETAVQPYQDQSYQ
PVYFVSESFEDAKQKLRRYSSTIQTPFSIRYDPYTCSMEVLDEPSKIQNALGQMRDDLKILHKALGRLGQK

10 cysteines are found in positions:   39  108  112  150  221  283  301  302  35
1  436


Matrix of pair scores
 POS:    39   150   221   351
  39: -999    84    67    74
 150:   84  -999    78    94
 221:   67    78  -999    99
 351:   74    94    99  -999
CYS     39 is probably SS-bounded      Score=   4.0
CYS    108 is not SS-bounded           Score= -27.5
CYS    112 is not SS-bounded           Score= -22.4
CYS    150 is not SS-bounded           Score= -18.1
CYS    221 is probably not SS-bounded  Score=  -8.2
CYS    283 is not SS-bounded           Score= -27.9
CYS    301 is not SS-bounded           Score= -28.5
CYS    302 is not SS-bounded           Score= -61.7
CYS    351 is not SS-bounded           Score= -16.1
CYS    436 is not SS-bounded           Score= -22.1

The most probable pattern of pairs: 39-150, 221-351,
```

Fig 3. Recognition of cysteine bond computed using Cys-Rec Tool.

RESULTS

The results of primary structure analysis suggest that most of TH sequences are hydrophobic in nature due to presence of high non-polar residues (table 2 and 3).

Table 2. Amino Acid Composition (%) of Tyrosine Hydroxylase Computed using Expasy Prot –Param Tool

Amino Acid (%)	O4 2091	Q6 ZM47	Q6I VB6	B3 DJW5	Q9 W7K7	E9 RK47	Q1 PP W9	Q6 IV Y6	Q1 LW Z5	Q6 DU B2	Q5 RH I3	B7 U5 07	B9 W0 A8	A0 JG X6
Ala(A)	7.8	7.1	6.2	6.7	4.5	7.4	6.1	5.9	6.2	3.5	6.1	6.5	3.1	8.4
Arg (R)	7.8	6.4	6.4	6.7	6.4	9.9	8	7.4	6.4	6.4	6.5	9.3	8.5	6.5
Asn(N)	1.6	0.2	2.3	0.8	0	1.3	1.4	2.1	2.3	1.7	2.3	0.9	0	1.4
Asp (D)	6.1	6.6	6.6	6.7	9.1	5.8	6.6	6.7	6.4	6.4	6.3	4.6	9.2	6.5
Cys(C)	1.4	2.2	2.1	1.8	1.8	1.6	2.4	2.1	2.1	2.9	2.1	2.8	1.5	2
Gln (Q)	2.9	3.9	5.1	3.5	1.8	2.6	3.3	5.3	5.1	1.7	5.1	3.7	1.5	3.7
Glu (E)	9.6	9.1	8.3	8.6	9.1	9.6	8	8	8.3	11.6	8.4	9.3	10.8	8.8
Gly (G)	4.7	4.9	5.3	4.3	1.8	3.5	4.2	5.9	5.3	2.9	5.3	5.6	3.1	4.7
His (H)	3.5	3.9	2.3	3.3	6.4	4.2	4.7	1.7	2.3	6.4	2.3	4.6	6.2	3.3
Ile (I)	4.5	3.9	4	4.7	3.6	5.1	3.3	2.9	4	5.2	4	4.6	4.6	3.9
Leu (L)	9.8	9.8	9.8	9.8	8.2	9	10.8	9.9	10	8.7	9.9	12	6.9	10.6
Lys (K)	4.7	5.9	6.2	6.1	10.9	4.8	3.3	5.5	6.2	8.7	6.1	0.9	9.2	5.5
Met (M)	0.8	0.7	1.5	0.8	0.9	1	1.4	1.5	1.5	0.6	1.5	0	0.8	1
Phe (F)	5.5	5.6	5.9	5.1	5.5	5.8	6.1	5.9	5.9	4	6.1	6.5	3.8	5.3
Pro (P)	4.7	4.9	5.5	4.9	5.5	4.5	6.6	5.9	5.5	5.8	5.5	3.7	5.4	4.9
Ser (S)	9.4	8.6	9.3	9.6	5.5	9.6	5.7	9.2	9.3	5.8	9.5	7.4	5.4	8.4
Thr (T)	5.6	5.1	4.2	5.9	6.4	6.1	6.1	4	4.2	6.9	4.2	5.6	6.9	5.3
Trp (W)	0.1	0.7	0.6	0.6	0.9	0.6	1.9	0.8	0.6	1.2	0.6	0.9	1.5	0.6
Tyr (Y)	3.5	3.9	4	3.7	3.6	2.6	4.2	4	4	4.6	4	6.5	4.6	3.7
Val (V)	4.7	6.4	4.2	6.1	8.2	5.1	5.7	5.5	4.2	5.2	4.2	4.6	6.9	5.3

Table 3. Parameters computed using Expasy's Prot –Param Tool

Accession No.	Sequence Length	M. Wt.	pI	No. of Negative Residue	No. of Positive Residue	Extinction Coefficient	Instability Index	Grand Average Hydropathy	Aliphatic Index
O42091	488	55490.4	5.59	77	61	42205	49.42	-0.448	80.37
Q6ZM47	408	46504.5	5.63	64	50	40840	47.99	79.12	-0.425
Q6IVB6	471	53796.6	5.61	70	59	45435	49.11	-0.564	72.29
B3DJW5	489	55602.8	5.80	75	63	43820	49.19	-0.434	81.17
Q9W7K7	110	13012.8	6.71	20	19	11585	44.12	-0.738	74.36
E9RK47	312	35948.5	6.68	48	46	23170	53.63	-0.546	77.24
Q1PPW9	212	24764.0	5.8	31	24	35660	48.73	-0.434	77.74
Q6IVY6	476	54303.1	5.6	70	61	50935	52.82	-0.561	71.70
Q1LWZ5	471	53794.6	5.69	69	59	45435	49.11	-0.549	73.12
Q6DUB2	173	20472.1	6.11	31	26	23170	48.88	-0.805	72.66
Q5RHI3	475	54314.2	5.7	70	60	45435	49.59	-0.557	72.51
B7U507	108	12619.2	5.68	15	11	16055	39.93	-0.294	84.91
A0JGX6	489	55388.5	5.54	75	59	43320	46.93	-0.413	80.43

The presence of Cys residues indicates the presence of disulphide bridges (SS bonds) in these sequences. Moreover, the primary structure analysis suggests that the TH Q9W7K, B9W0A have no asparagine amino acid. TH B7U50 has not Met (M) amino acids. Glutamic acid, Leucine, and Serine present in high amount in almost all sequences that we have selected. The average molecular weight of tyrosine hydroxylase is 81935Da. Isoelectric point (pI) is the pH at which surface of protein is covered with charge but net charge of protein is zero. At pI protein are stable and compact. The computed pI value of all teleost fishes show acidic character (pI<7). The computed isoelectric point (pI) will be useful for developing buffer system for purification by isoelectric point focusing method. Extinction Coefficient of tyrosine hydroxylase at 280 nm is ranging from 5500 to 56560M-1cm-1with respect to concentration of Cys, Trp and Tyr. The high extinction coefficient of Q5RHI3, Q1LWZ5, Q6IVB6 indicate presence of high concentration of Cys, Trp and Tyr. The low extinction coefficient of Q9W7K, B7U507 indicate presence of low concentration of Cys, Trp and Tyr. The computed protein concentration and extinction coefficient help in quantitative study of protein and protein ligand interactions in solution.

Table 4. Physical Properties of Tyrosine Hydroxylase computed using 'Own Matlab Codes'

Accession No.	Hydrophilic Residue	Hydrophobic Residue	Acidic Residue	Basidic Residue	Aromatic Residue	Aliphatic Residue	Non Polar	Polar
O42091	25	39.7541	15.7787	15.9836	13.1148	20.082	49.1803	50.8197
Q6ZM47	25.7353	40.4412	15.6863	16.1765	14.2157	20.098	50.2451	49.7549
Q6IVB6	28.4501	38.4289	14.862	14.862	12.9512	18.0467	49.2569	50.7431
B3DJW5	25.7669	39.4683	15.3374	16.1554	12.6789	20.6544	48.6708	51.3292
Q9W7K7	30.9091	37.2727	18.1818	23.6364	16.3636	20	44.5455	55.4545
E9RK47	24.0385	38.141	15.3846	18.9103	13.141	19.2308	46.1538	53.8462
Q1PPW9	22.6415	41.9811	14.6226	16.0377	16.9811	19.8113	52.8302	47.1698
Q6IVY6	27.521	38.4454	14.7059	14.4958	12.395	18.2773	50.2101	49.7899
Q1LWZ5	28.2378	38.6412	14.6497	14.862	12.9512	18.259	49.4692	50.5308
Q6DUB2	30.0578	35.8382	17.9191	21.3873	16.185	19.0751	44.5087	55.4913
Q5RHI3	28.2105	38.5263	14.7368	14.9474	13.0526	18.1053	49.2632	50.7368
B7U507	19.4444	44.4444	13.8889	14.8148	18.5185	21.2963	53.7037	46.2963
B9W0A8	30.7692	33.8462	20	14.8148	16.1538	18.4615	42.3077	57.6923
A0JGX6	25.9714	40.8998	15.3374	15.3374	16.1538	19.8364	50.5112	49.4888

On the basis of instability index Expasy's ProtParam classifies the B7U507 are stable (instability index<40) and A0JGX6, Q5RHI3, Q6DUB2, Q1LWZ5, Q6IVY6, Q1PPW9, E9RK47, Q9W7K7, B3DJW5, Q6IVB6, Q6ZM47, O42091 are unstable (instability index>40). The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains (A, V, I, AND L) is regarded as positive factor for the increase of thermal stability of globular protein. The lower thermal stability is indicative of a more flexible structure when compared to others tyrosine hydroxylase (table 3). The very high aliphatic index of all tyrosine hydroxylase infers that tyrosine hydroxylase may be stable for a wide range of temperature. Grand Average hydropathy (GRAVY) Index of tyrosine hydroxylase are ranging from -0.1 to 0.9.The very low GRAVY index of all sequences infers that tyrosine hydroxylase of teleost fishes could result in better interection with water. The secondary structure predicted with help of programs SOPM and SOPMA (table 5) infers that the TH have high lysine (L) and glutamic acid aminoacids and mostly have alfa-helices, rest of TH have mixed amount of alfa-helices, beta strand, coils. The very

high coil structural content of Q6IVY6 (59%), Q1LWZ5 (53%) are due to rich content of more flexible glycine and hydrophobic proline amino acids. Proline has a special property of creating kinks in polypeptide chains and disrupting ordered secondary structure. The SOSUI server classifies the TH are soluble proteins. B7U507, Q6ZM47 have high amount of alpha helix while B9W0A8, Q9W7K7 have high amount of random coil. There is no occurance of 3_{10} helix, Pi helix, Beta bridge, Bend region (table 5).

Table 5. Secondary structure analysis computed using SOPMA Tool

Accession No.	Alpha Helix	310 Helix	Pi Helix	Beta Bridge	Extended Strand	Beta Turn	Bend Region	Random Coil
O42091	52.87	0	0	0	10.45	4.10	0	32.58
Q6ZM47	57.35	0	0	0	8.58	3.92	0	30.15
Q6IVB6	52.23	0	0	0	9.13	3.18	0	35.46
B3DJW5	53.99	0	0	0	9.61	4.7	0	31.70
Q9W7K7	47.27	0	0	0	10	1.82	0	40.91
E9RK47	50	0	0	0	10.58	6.09	0	33.33
Q1PPW9	55.66	0	0	0	5.66	5.19	0	33.49
Q6IVY6	49.16	0	0	0	9.87	3.36	0	37.61
Q1LWZ5	52.65	0	0	0	9.34	4.03	0	33.97
Q6DUB2	54.34	0	0	0	10.4	4.05	0	31.21
Q5RHI3	53.05	0	0	0	9.47	5.05	0	32.42
B7U507	57.41	0	0	0	8.33	7.41	0	26.85
B9W0A8	46.92	0	0	0	11.54	3.85	0	37.69
A0JGX6	54.40	0	0	0	10.63	3.27	0	31.70

Chymotrypsin-low specificity (112 for O42091) have more cleavage site than Chymotrypsin-high specificity (46 for O42091). Lysine C and Lysine N both have same cleavage site. Pepsin (pH1.3) have more cleavage site than pepsin (pH>2). Q6IVY6, Q1LWZ5, Q5RHI3 have seven CNBr cleavage site which is larger than rest of the sequence while B7U507 have no CNBr cleavage site (table 6).

Table 6. Cleavage Site computed using Peptide cutter Tool.

Accession No.	Chymotrypsin-high specificity	Chymotrypsin-low specificity	Formic acid	Trypsin	Lys C	Lys N	Pepsin (pH1.3)	Pepsin (pH>2)	CNBr	Proline-endopeptidase
O42091	46	112	30	57	23	23	137	110	4	4
Q6ZM47	41	97	27	47	24	24	123	98	3	4
Q6IVB6	48	109	31	57	29	29	140	113	7	4
B3DJW5	45	109	33	59	30	30	137	110	4	5
Q9W7K7	10	25	10	17	12	12	29	23	1	3
E9RK47	27	68	18	43	15	15	78	67	3	3
Q1PPW9	25	59	14	23	7	7	69	52	3	2
Q6IVY6	49	109	32	57	26	26	142	115	7	5
Q1LWZ5	48	110	30	57	29	29	142	115	7	4
Q6DUB2	16	41	11	24	15	15	46	34	1	3
Q5RHI3	49	111	30	58	29	29	142	116	7	4
B7U507	15	33	5	10	1	1	41	31	0	1
B9W0A8	12	28	12	21	12	12	32	21	1	3
A0JGX6	46	116	32	56	27	27	143	117	5	4

The 3D structures of TH Q6IVB6 were modelled using PDB templates 1LFD selected from hits obtained through the BLASTP analysis and modelled structures were evaluated. The cysteine and the SS bonds identified using three dimensional structures of TH Q6IVB6 are shown in figures 1 and 2 respectively. Number and position of cysteine residues are shown in table 7.

Table 7. Recognition of SS-bounded cysteines by Cys-Rec Tool

Accession No.	No. of Cysteine Residue	Position Of Cysteine Residue
O42091	7	125 , 167, 239, 253, 301, 319, 370
Q6ZM47	9	54, 98, 140, 212, 226, 241, 274, 292, 343
Q6IVB6	10	39, 108, 112, 150, 221, 283, 301, 302, 351, 436
B3DJW5	9	82 , 126, 168, 240, 254, 269, 302, 320 , 371
Q9W7K7	2	33, 75
E9RK47	5	125, 167, 239 , 253, 301

Q1PPW9	5	13, 85, 99, 147, 165
Q6IVY6	10	111, 153, 225, 239 , 287, 305, 306, 356, 441, 450
Q1LWZ5	10	39, 108, 112, 150 , 221, 283 , 301, 302, 351, 436
Q6DUB2	5	30 , 72, 144, 152, 158
Q5RHI3	10	41, 110 , 114, 152, 224, 286 , 304, 305 , 355, 440
B7U507	3	43 , 57, 105
B9W0A8	2	30 , 72
A0JGX6	9	110, 126, 168, 240, 254, 269, 302, 320, 371

DISCUSSION AND CONCLUSION

Fourteen teleost TH sequences were selected to study their physico-chemical properties, which can be extremely useful model for screening compounds at several stages of drug discovery process. Teleost fish, which roughly make up half of the extant vertebrate species, exhibit an amazing level of biodiversity affecting their morphology, ecology and behaviour as well as many other aspects of their biology (35). This huge variability makes fish extremely attractive for the study. The last common ancestor between fish and mammals dates back to the very origin of the vertebrate lineage and today, half of modern vertebrates are fish. It is thus not surprising that several fish species have played important roles in recent years to advance our understanding of vertebrate genome evolution, to inform us on the structure of human genes, and, somewhat more unexpectedly, to provide leads to understanding the function of genes involved in human diseases. Humans and fishes share many developmental pathways, organ systems, and physiological mechanisms, making conclusions relevant to human biology. The respective advantages of zebrafish, medaka, *Tetraodon*, or *Takifugu* have been well exploited so far with bioinformatics analyses and molecular biology techniques (34). There are various reports on the efficacy models in zebrafish, its utility as an early stage screening tool for the discovery of CNS therapeutics should enable increased throughput of *in vivo* analysis of novel compounds for neurological disorders. Zebrafish have recently become a focus of neurobehavioral studies since larvae display learning, sleep, drug addiction, and other neurobehavioral phenotypes that are quantifiable and relate to those seen in man (4, 5, 6, 7). Furthermore, the organization of the zebrafish genome and the genetic pathways controlling signal transduction and development are highly conserved between zebrafish and man (24,25,26,27,28). We can describe the application of the teleost fishes models to neurodegenerative disorders, schizophrenia and

learning and memory. Several of these diseases affect an aging population and are adult onset raising questions as to the appropriateness of using a rapidly developing larval system. As more information is reported on neurological assays in teleost fishes, the utility of this model organism as an early stage screening tool for CNS disorders should help increase in vivo throughput and ameliorate the cost associated with drug screening in mammals. Patients with schizophrenia and other conditions including Huntington's disease have been shown to exhibit impaired prepulse inhibition (PPI) (29). PPI is a neurological phenomenon in which a weak prestimulus or prepulse suppresses the response to a subsequent startling stimulus and is highly conserved among vertebrates. Again, using this paradigm will allow the screening of novel therapeutics for schizophrenia at an earlier stage of the drug discovery process. The main pharmacological approach to studying PD in animal models has been using 1-methythl-4-phenyl-1, 2, 3, 6-tetrahydropyridine (MPTP) which reproduces some of the effects of idiopathic Parkinson's disease (30). Zebrafish embryos treated with MPTP demonstrated a loss of TH and DAT- expressing neurons which could be rescued using the monoamine oxidase-B inhibitor deprenyl (31). A reduction was also seen in the locomotor activity mimicking motor effects seen in Parkinson's disease patients; however this did not always occur simultaneously with the reduction in dopaminergic cells (32) and in the adult a reduction is seen in the locomotor activity without an effect being seen in dopaminergic cells (33). The utility of these various models will aid screening for novel compounds for both the hereditary and the idiopathic forms of Parkinson's disease. Primary structure analysis reveals that most of the TH sequences under study are hydrophobic in nature and contain disulphide linkages. In general, disulfide bonds are formed in the lumen of the RER (rough endoplasmic reticulum) but not in the cytosol. This is due to the oxidative environment of the ER and the reducing environment of the cytosol. Thus disulfide bonds are mostly found in secretory proteins, lysosomal proteins, and the exoplasmic domains of membrane proteins. Disulfide bonds play an important role in the folding and stability of some proteins, usually proteins secreted to the extracellular medium (18). Since most cellular compartments are reducing environments, in general, disulfide bonds are unstable in the cytosol, with some exceptions as noted below, unless a sulfhydryl oxidase is present (17). Disulfide bonds in proteins are formed between the thiol groups of cysteine residues. The other sulfur-containing amino acid, methionine, cannot form disulfide bonds. A disulfide bond is typically denoted by hyphenating the abbreviations for cysteine, e.g., when referring to Ribonuclease A the "Cys26-Cys84 disulfide bond", or "26-84 disulfide bond", or most simply as "C26-C84" (19)

where the disulfide bond is understood and does not need to be mentioned. The prototype of a protein disulfide bond is the two-amino-acid peptide cystine, which is composed of two cysteine amino acids joined by a disulfide bond (shown in Figure 2) in its unionized form). The structure of a disulfide bond can be described by its χ_{ss} dihedral angle between the $C^{\beta} - S^{\gamma} - S^{\gamma} - C^{\beta}$ atoms, which is usually close to ±90°. It holds two portions of the protein together, biasing the protein towards the folded topology. That is, the disulfide bond destabilizes the unfolded form of the protein by lowering its entropy. The disulfide bond may form the nucleus of a hydrophobic core of the folded protein, i.e., local hydrophobic residues may condense around the disulfide bond and onto each other through hydrophobic interactions. Physico-chemical characterization studies give a good idea about the properties such as pI, EC, AI, GRAVY and Instability index that are essential and vital in providing data about the proteins and their properties. The extinction coefficient indicates how much light a protein absorbs at a certain wavelength (15). EC helps in an estimation of protein by a spectrophotometer during its purification. It is possible to estimate the molar extinction coefficient of a protein from knowledge of its amino acid composition. The instability index provides an estimate of the stability of your protein in a test tube. Statistical analysis of 12 unstable and 32 stable proteins has revealed (14) that there are certain dipeptides, the occurence of which is significantly different in the unstable proteins compared with those in the stable ones. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable. The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine). It may be regarded as a positive factor for the increase of thermostability of globular proteins. (16). GRAVY (Grand Average of Hydropathy) value for a peptide or protein is calculated as the sum of hydropathy values (20,21) of all the amino acids, divided by the number of residues in the sequence. Secondary structure analysis(24) predicts that most of them contain only alfa-helices and remaining of them contain mixed structure. Helices observed in proteins can range from four to over forty residues long, but a typical helix contains about ten amino acids (about three turns). In general, short polypeptides do not exhibit much alpha helical structure in solution, since the entropic cost associated with the folding of the polypeptide chain is not compensated for by a sufficient amount of stabilizing interactions. In general, the backbone hare readily attacked by the ambient water molecules. However, in more hydrophobic environments such as the plasma membrane, or in the presence of co-solvents such as trifluoroethanol (TFE), or

isolated from solvent in the gas phase (10). Oligopeptides readily adopt stable α-helical structure. Furthermore, crosslinks can be incorporated into peptides to conformationally stabilize helical folds. Crosslinks stabilize the helical state by entropically destabilizing the unfolded state and by removing enthalpically stabilized "decoy" folds that compete with the fully helical state (9). Tyrosine hydroxylase activity is regulated chronically (days) by protein synthesis (12). Tyrosine hydroxylase activity is increased in the short term by phosphorylation. The regulatory domain of tyrosine hydroxylase contains at least three serine residues that are phosphorylated by multiple protein kinases (13). The presence of Cys residues in fish indicates the presence of disulfide bridges which is also confirmed using MATLAB tools and Chimera tools. Protein contain cysteine may protect themselves from irreversible damage due to oxidative stress. During times of oxidative stress, such as when damaged by exposure to the sun, some proteins can modify their cys residues through a process called S-thiolation. The cysteine thiol group is nucleophilic and easily oxidized. The reactivity is enhanced when the thiol is ionized, and cysteine residues in proteins have pK_a values close to neutrality, so are often in their reactive thiolate form in the cell (8). Because of its high reactivity, the thiol group of cysteine has numerous biological functions.

REFERENCES

1. Kaufman S., "Tyrosine hydroxylase: Advances in Enzymology - and Related Areas of Molecular Biology *Adv. Enzymol. Relat. Areas Mol. Biol.*. 1995, 70:103–220.

2. Nagatsu T., "Tyrosine hydroxylase: human isoforms, structure and regulation in physiology and pathology". *Essays Biochem.* 1995, 30: 15–35.

3. Grima B, Lamouroux A, Boni C, Julien J-F, Javoy-Agid F and Mallet J. A single human gene encoding multiple tyrosine hydroxylases with different predicted functional characteristics. *Nature*. 1987, 326:707–711.

4. Sivkumar K, Balaji S and Gangaradhakrishnan. *In silico* characterization of antifreeze proteins using computational tools and servers. *J. Chem. Sci.* 2007, 119: 571-579.

5. Cristobal S, Zemla A, Fischer D, Rychlewski L and Elosfsson A. A study of quality measures for protein threading models. *BMC Bioinformatics* 2001, 2- 5

6. Ginns E., Pauls D., Falls K., Keith T., and Paul S. Search for DNA markers linked to manic depressive illness in the Old Order Amish. J. Psychiatr. Res. 1992, 26**:** 305–308.

7. Wolters E. C. and Calne D. B. (1989) Parkinson's disease. *CMAJ* 1989, 140 (5):507-514**.**

8. Bulaj, G., Kortemme T., Goldenberg, D.P. "Ionization-reactivity relationships for cysteine thiols in polypeptides". Biochemistry. 1998 23;37(25):8965-8972.

9. Kutchukian, P. S; Yang J. S, Verdine G. L, Shakhnovich E. I "All-Atom Model for Stabilization of alpha-Helical Structure in Peptides by Hydrocarbon Staples". *Journal of the American Chemical Society* 2009,131 **(**13) : 4622–4627

10. Hudgins, "Helix Formation in Unsolvated Alanine-Based Peptides: Helical Monomers and Helical Dimers". *Journal of the American Chemical Society* 1999, 121:3494-3501

11. Hedstrand H, Landgren E, Betterle C, Perheentupa J, Gustafsson J, Husebye E, Rorsman F, Kämpe O."Identification of tyrosine hydroxylase as an autoantigen in autoimmune polyendocrine syndrome type I". *Biochem. Biophys. Res. Commun.* 2000, 267 (1): 456–61

12. Bobrovskaya L, Gilligan C, Bolster E.K, Flaherty J.J, Dickson P.W, Dunkley P.R "Sustained phosphorylation of tyrosine hydroxylase at serine 40: a novel mechanism for maintenance of catecholamine synthesis". *J. Neurochem.* 2007, 100 (2): 479–89

13. Haycock J.W "Phosphorylation of tyrosine hydroxylase in situ at serine 8, 19, 31,and 40". *J. Biol. Chem.* 1990, 265 (20): 11682–91.

14. Guruprasad, K., Reddy, B.V.B. and Pandit, M.W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng*. 1990, 4:155-161.

15. Gill, S.C. and von Hippel, P.H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem*. 1989, 182:319-326.

16. Ikai, A.J. Thermostability index of globular proteins. *J. Biochem*. 1980, 88: 1895-1898

17. Hatahet F, Nguyen V.D, "Disruption of reducing pathways is not essential for efficient disulfide bond formation in the cytoplasm of E. coli". *MCF* 2010, 9 (67): 67.

18.Sevier, C.S and Kaiser, C.A. "Formation and transfer of disulphide bonds in living cells". *Nature Reviews Molecular and Cellular Biology* 2002, 3 (11): 836-847.

19.Ruoppolo M, Vinci F, Raines RT, Marino G. " Contribution of individual disulphide bonds to the oxidative folding of ribonuclease A" *Biochemistry*, 2000 39 : 12033-42.

20. Abeliovich A, Parkinsonism genes: culprits and clues. *J Neurochem.* 2006**,** 99:1062–72.

21. Kyte J, Method for displaying hydropathic character of protein. J.Mol.Biol.157(1):105-32

22. http://sun1.softberry.com/berry.phtml?topic=cys_rec&group=help&subgroup=propt

23.Takatsugu Hirokawa, Shigeki Mitaku, classification and secondary structure prediction system for membrane proteins *Bioinform. Appl. Note* 1998, 14 : 378-379

24. Zhdanova IV, Melatonin promotes sleep-like state. *Brain Res.* 2001, 903:263–8

25. Ninkovic J, Folchert A, Genetic identification of AChE as a positive modulator of addiction to the psychostimulant D-amphetamine in zebrafish. *J Neurobiol.* 2006**,** 66:463–75.

26. Orger MB, Behavioral screening assays in zebrafish. *Methods Cell Biol.* 2004**, 77**:53–68.

27. Guo S. Linking genes to brain, behavior and neurological diseases: what can we learn from zebrafish? *Genes Brain Behav.* 2004,3(2):63-74.

28. Cahill GM. Clock mechanisms in zebrafish. *Cell Tissue Res.* 2002, 309:27–34.

29. Postlethwait J.H, Woods IG, Ngo-Hazelett P, et al. Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res.* 2000, 10:1890–902.

30. Swerdlow N.R, Impaired prepulse inhibition of acoustic and tactile startle response in patients with Huntington's disease. *J Neurol Neurosurg Psychiatry.* 1995, 58:192–200.

31. Smeyne,MPTP model of Parkinson's disease.*Brain Res Mol Brain Res.* 2005, 134:57–66.

32. Bretaud S, Lee S, Guo S. Sensitivity of zebrafish to environmental toxins implicated in Parkinson's disease. *Neurotoxicol Teratol***.** 2004,26:857–64.

33. Salzmann J, Anichtchik OV, Best JD, Effect of MPTP treatment on locomotor activity and dopaminergic cells in the zebrafish larvae. Workshop – Neurodegeneration; basic mechanisms of motor and cognitive dysfunctions; September 2–4; Cracow, Poland. 2006

34. Anichtchik OV, Kaslin J, Peitsaro N, et al. Neurochemical and behavioural changes in zebrafish Danio rerio after systemic administration of 6-hydroxydopamine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. *J. Neurochem.* 2004, 88**:**443–53.

35. Hugues Roest Crollius and Jean Weissenbach *Fish genomics and biology,* 2005

36. http://web.expasy.org/protparam.

37. http://web.expasy.org/peptide_cutter.

38. www.scribd.com/api_user_11797_abbaspu/d/.../42-SOPMA

39. www.cbs.dtu.dk › CBS Prediction Servers