

# **International Journal of Pharmacy and Pharmaceutical Sciences**

ISSN- 0975-1491

Vol 9, Issue 10, 2017

**Original Article** 

# SEQUENCE ANALYSIS AND STRUCTURAL CHARACTERIZATION OF TISSUE TRANSGLUTAMINASE 2(TG2) BY IN SILICO APPROACH

# SHIVKUMAR B. MADAGI<sup>1</sup>, PRACHI P. PARVATIKAR<sup>\*1</sup>

<sup>1</sup>Department of Bioinformatics, Akkamahadevi Women's University (Karnataka State Women's University), Vijayapura, Karnataka 586108, India

Email: prachisandeepk@gmail.com

#### Received: 14 Jun 2017 Revised and Accepted: 31 Aug 2017

#### ABSTRACT

**Objective:** TG2 is multifunctional protein. The up regulation leads into different pathological disorders. The objective of the present study was the prediction of a structural feature of TG2 (Tissue transglutaminase) protein with *in silico* approach.

**Methods:** In this study, we have investigated the structural feature of TG2 by using various biological databases (uniprot, NCBI, Pfam) and online tools such as BLAST, PDBsum, protoparam tools.

**Results:** The predicted structure of TG2 protein has shown that amino acid residues conserved throughout the sequence in selected mammals. During the course of evolution, mammalian TG2 protein is orthologus; human TG2 shares its characters with chimpanzee while mice and rat were closely related to each other. This protein was mainly cytosolic but also present in other cell organalles. It has four catalytic domains and three active sites with multiple metal binding domain specifically for calcium. The pI value was 5.11, GRAVY-0.283. The phosphorylation sites were present at serine and threonine. The structure was a monomer with 14 alpha helices and 9 sheets. Ramachandran plot showed about 98% residues in the favoured region.

**Conclusion:** Collectively, these data suggest that the predicted TG2 protein may act as a good therapeutic target. Targeting phosphorylation sites may help in down regulation of TG2. The modelled protein can be used for further work.

**Keywords:** Tissue transglutaminase 2, Metastasis, Sequence alignment, Phylogenetic analysis, BLAST, Stereochemical parameter, Phosphorylation, Hydropathicity

© 2017 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open-access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/) DOI: http://dx.doi.org/10.22159/ijpps.2017v9i10.20353

# INTRODUCTION

Tissue transglutaminase 2 (TG2) is a calcium-dependent cellular matrix protein ubiquitously expressed member of the large family of transglutaminases [1]. Transglutaminases were first isolated from the mammalian liver in 1950s [2]. Transglutaminase family consists nine members, of which TG2 is one among them. TG2 is the most extensively studied member and biologically characterized, because it performs multiple functions and present in the wide spectrum of living organisms such as microorganism, invertebrates, birds, mammals and predominantly in human beings [3]. In the humans, it is mainly found in both intracellular and extracellular locations, including cytosol, nucleus, endoplasmic reticulum, mitochondria, extracellular matrix, focal adhesion area, and as an intrinsic component of the plasma membrane [4]. TG2 interacts with both intra and extracellular proteins and alters their structure and functions.

Under the normal cellular condition, TG2 is involved in activities such as a cell death process, cell adhesion, post-translation modifications, migration, growth, apoptosis and extracellular matrix organization [5]. When apoptosis begins TG2 plays a role as protector by prevention of tissue damage, inflammation, and autoimmunity by activating macrophages [6]. Dysregulation in its function results in pathological conditions leading to Huntington's, Alzheimer's, Parkinson's diseases and various types of cancers [7].

It takes part in cross-linking of proteins by way of formation of covalent bonds with a free amine group and protein-bound glutamines leading into changes in solubility, structure and function of the proteins [8]. TG2 undergoes post-translational modifications (PTM) by phosphorylation, which take place usually at serine-216 (Ser216) and regulates the function of TG2 protein-protein interaction [9-12].

TG2 catalyses  $Ca^{2+}$  dependent transamidation reaction and GTP hydrolyzing activities [13, 14]. The activity of TG2 depends upon conformation form. The transamidation reaction and binding with

GTP are responsible for a change in conformation status [15]. During transamination reaction, TG2 is found to be in open conformation, but when GTP binding occurs conformation changes to closed or folded form [16]. Closed conformational form (signalling) is the active form of TG2 which promotes cross-linking activity thereby cell adhesion [17]. The cross-linking activity of TG2 also brings about changes in solubility, structure, and function of proteins. This GTP-bound form is thought to act as a signal transducer which relays and receives signals from different receptors [18].

Due to the involvement of TG2 in different diseases, it is an important target as well as a molecular marker [19]. Selective inhibition of TG2 may help in down-regulation of pathological state [20]. For the design and development of specific inhibitors, as drug molecules, understanding of the exact structure of the active site and pathophysiology of this protein is important [21]. Till now crystal structures of the inactive form TG2 bound to GTP, ATP and with irreversible inhibitors is reported [22-24].

The present study has been undertaken to get in-depth understanding of the structure, function, and phylogeny of TG2 protein so that the information would be highly valuable for drug design, and selection of an appropriate animal model for screening drug molecule. Sequence analysis entails multiple sequence alignment, the evolutionary history of human TG2 with other mammalian organisms, physiochemical characteristics, subcellular location, phosphorylation site followed by the secondary structure prediction and the generation of the ramachandran plot. Functional analysis has also been attempted using a battery of computational tools and web servers.

#### MATERIALS AND METHODS

#### Target identification and sequence retrieval

After literature survey, the precise identification of target was performed. The sequences were retrieved from UniProt database.

From the set of the finished genome sequence, the different mammalian proteins were selected.

# Multiple sequence alignment and generating phylogenetic tree using UPGMA

Multiple sequence alignment was performed using CLUSTAL OMEGA [25] of EBI and a phylogenetic tree was analyzed by using same web tool.

#### Protein domain identification and target validation

The selected protein sequences were used as input using Pfam [26] and the results were cross-checked with NCBI CDD-tool.

To find out the catalytic domain, a significant protein match was prepared by using the BLAST [27] (Basic Local Alignment Search Tool) provided by NCBI.

# Prediction of subcellular localization, signal peptide, and physicochemical characterization

For the human TG2 protein subcellular localization prediction was carried out by CELLO tools [28]. The presence of signal peptide was checked by Signal P 4.1[29] server and physicochemical characterization was done by the ExPASy Prot Param tool [30].

# Phosphorylation profile analysis and prediction of Accessible Surface Area (ASA)

Human TG2 protein was subjected to phosphorylation profile analysis using Netphos 2.0 server [31]. ASA was predicted through Net Surf P [32] server of ExPaSy suite.

# Generation of secondary structure and Ramachandran plot

The secondary structure of human TG2 was predicted by using PDBsum server [33]. To find the percentage of favourable residues the Ramachandran plot was generated in RAMPAGE [34].

#### Prediction of tertiary structure

Tertiary structure prediction was performed by SWISS-MODEL server [35].

#### Quality assessment of protein

The quality of the human TG2 protein structure was determined by ERRAT web server [36]. The protein structure obtained from X-ray crystallography was used for verification by this server.

#### **RESULTS AND DISCUSSION**

#### Sequence retrieval

After protein database search the amino acid sequence of TG2 protein was retrieved from Uniprot. For the present study mammalian organisms; Human, Chimpanzee, Bovine, European polecat, Pig, Mice, Chick were selected (table 1). The reason for this selection for analysis is, because target protein is predominant in the mammals and the study would help to understand the conserved domain and evolutionary relationship of human TG2 with the selected mammals. Since one of the objectives of the study was to find out the suitability of different animals as models for drug screening.

#### Table 1: The different organism with uniprot ID and length of protein for the study

S. No.	Organisms	Uniprot ID	Length of protein	
1	Homo sapiens	P 21980	687	
2	Mus musculus (Mouse)	P 21981	686	
3	Cavia cutleri (Pig)	P 08587	690	
4	Bos taurus (Bovine)	P51176	687	
5	Rattus norvegicus (Rat)	AAH620621	686	
6	Pan troglodytes (Chimpanzee)	K7D7G1	687	
7	Gallus Gallus (Chicken)	Q 018413	689	
8.	Mustela Putoriufuro	G9KTJJ	687	

ERCDLELETINGENHEN ALCEEKLUWR BGO EFKLITHEGENN ERCDLELETINGENHEN ALCERLUWR BGO EFKLITHEGENN ERCDLELENNGENHEN ALCERLUWR BGO EFKLITHEGENN ERCDLELENNGENHEN ALCERLUWR BGO EFKLITHEGEN ERCDLELGENNGENHEN ALCERLUWR BGO EFKLITHEGES ERCDLELGENGENHEN ALCEGEKLURR BGO EFKLITHEGES ERCDLELGENGENHEN ALCEGEKLURR BGO EFKLITHEGES ERCDLELGENGENHEN ALCEGEKLURR BGO EFFLITHEGES ERCDLELGENGENHEN ELGEGEKLURR BGO EFFLITHEGES
ERC DLE LETINGRICH HTA DLCRE KL VVR. RCG PFWL THE FEGNNY ERC DLE LEANGRICH HTA DLCRE KL VVR. RCG PFWL THE FEGNNY ERC DLE LEVINGRICH HTA ELCRE KL VVR. RCG PFWL THE FEGNSY ERC DLC ELCYNGRICH HTA DLCRE KL VLR. RCG PFWL THE FEGNSY ERC DLE I GANGRICH HTA DLC QCKL VLR. RCG RTFRI TLY FEGNSY ERC DLE I GANGRICH HTA DLC QCKL VLR. RCG RTFRI TLY FEGNSY
ERC DLE LEANGR DHHTA DLC RE RL VAR RGÖPFNL TLH FEG RN Y NGC DLE LEVNGR DHHTA BLC RE RL VAR RGÖPFNL TLH FEG RS Y ERC DL QL EVNGR DH RTA DLC RE RL VLR RGØPFNL TLH FEG RG Y ERC DLE I QANGR DH HTA DLC QQK VLR RGØPFNL TLY FEG RG Y
ERC DLE LEANGR DHHTA DLC RE RL VAR RGÖPFNL TLH FEG RN Y NGC DLE LEVNGR DHHTA BLC RE RL VAR RGÖPFNL TLH FEG RS Y ERC DL QL EVNGR DH RTA DLC RE RL VLR RGØPFNL TLH FEG RG Y ERC DLE I QANGR DH HTA DLC QQK VLR RGØPFNL TLY FEG RG Y
KNC DLELEVN GR DH HTAELCRERL VVR RGQ PFWL TLH FEGRS Y ERC DLQLEVNGR DH RTADLCRERL VLR RGQ PFWL TLH FEGRG Y ERC DLE I QANGR DH HTADLCQE KL VLR RGQR FRL TLY FEGRG Y ERC DLE I QANGR DH HTADLCQQ KL VLR RGQR FRL TLY FEGRG Y
ERC DLQLEVNGR DH RTA DLC RE RLVLR RGQ PF WLTLH FEGRG Y ERC DLE I QANGR DH HTA DLC QE KLVLR RGQ RF RLTLY FEGRG Y ERC DLE I QANGR DH HTA DLC QQ KLVLR RGQ RF RLTLY FEGRG Y
ERC DLQLEVNGR DH RTA DLC RE RLVLR RGQ PF WLTLH FEGRG Y ERC DLE I QANGR DH HTA DLC QE KLVLR RGQ RF RLTLY FEGRG Y ERC DLE I QANGR DH HTA DLC QQ KLVLR RGQ RF RLTLY FEGRG Y
ERC DLEI QANGR DH HTA DLC QE KLVLR RGORF RLTLY FEGRG Y ERC DLEI QANGR DH HTA DLC QQ KLVLR RGORF RLTLY FEGRG Y
ERC DLEI QANGR DHHTA DLC QQ KLVLR RGQ RFRL TLY FEGRG Y
ERC DLEI QANGR DHHTA DLC QQ KLVLR RGQ RFRL TLY FEGRG Y
TFSVVTG PAPSQEAGTKARF PLRDAVE EGDWT AT VVDQQDCT L
TES WITG PAPSOEAGTKARE PLRDAVDEGDWT AT WDOODCT L
TFC AVTG PDPSE EAGTKALFRL SDATE EGAWA AVAADORD STL
TFS AVTG PAPSE EAGTKARF PL SS AME EGAWT AVVMD QQD GVL
TFN AV TG PDPSE EA GTM ARF SL SS AVE GGT WS AS AVD QQD ST V
TFG AV TG PDPSE EAGTKARFSLSDNVE EGS WS AS VLD QQDNVL
TFG AV TG PDPSEEA GTK ARFSL SD DVE EGS WS AS VLD QQDNVL
A FDVET GPC PVET SGTRSHFT LT DCPEE GTWSAVLQQQDG ATL
· · · · · · · · · · · · · · · · · · ·
ANA PIGLYRLSLEA STGYQG SSFVLGHFILLFNAWCPADAVYL
ANA PIGLYRLSLEA STGYQG SS FVLGH FILLFNA WCP ADA VYL
ANA PVGHYRLSLEA STGYQG SS FMLGQ FTLLFNSWCPADAVYL
ANA FVGLYRLTLEA STGYQG SS FVLGH FTLLFNT WCP ADA VYL
A DE AT MET AT DE AT
ADA PIGLYRLSLEA STGYQGSS FVLGHFILLYNPRCPADAVYM
ANA PIGLYRLSLEA STGYQGSS FVLGH FILLYNA WCPADDVYL
ANA PV GQ YRLSLET STG YQG SS FMLGH FILLF NAWCP ADD VY L
s iarvgryr ltle a <mark>st gy qgs sp</mark> hlg dp vll pn awh pe da vyl
EYVLT QQGFIYQGS AKF IKN IPWN FGQFEDGILD ICL ILLDVN
EYVLT QQ GFIYQ GS AKF IKN IPWN FGQ FEDGILD ICL ILL DVN
EYVLT QQGFIYQG3 AKF IKN IP WN FGQ FEE GILD ICLMLL DVN
EYVLT QQ GFIYQ GSVKC IKS IPWN FGQ FEDGILD ICL ALL DT N
EYVLT QQGFIYQGS AKF ING IP WN FGQFEDGILD ICLMLLDIN
EYVLT QQGFIYQGSVKF IKSVPWN FGQFEDGI LDTCLMLL DMN
EYVLT QQGFIYQGSVKF IKSVPWN FGQFEDGILDACLMLLDVN
LIVEIQUGEIIQUS VALIAS VEWELOUTED GILDACLALEDVA
EYVLSQQGLIYMG SRDYITST PWNFGQFEDE IL AICLEMLDIN
********** ** . * ******************
GRD CSRR SSPVYVGRVV SGMVN CND-DQGVLLGRWDNNYG DGV
GRD C3RR33PVYVG RVV 3GMVN CN D-D QGV LL GR WDN NYG DG V
GRD CS RR SSPVY VG RVV SGMVN CN D-D QGV LL GR WDN NYA DG I
GRDCSRRSSPVYMGRVVSAMVNCND-DQGVLLGRWDNDYKDGV
GOD CS RR SRPVY VG RVV SAMVN CN D-D QGV LQ GR WDN NYS DG V
SRDCSRR SSPIYVGRVV SAMVN CND-DQGVLLGRWDNNYGDGI
SRDCSRRSSPIYVGRVVSGMVNCND-DOGVLLGRWDNNYGDGI
N LD CSR RND PVYI GRVVS AMVNC NDE DHGVL LG RWDNH YE DGM
A COCKERD FUTIGRUUS ANUNC BDE DEGUL DO RADANTE DOM
SVD ILRRWHNHG CORVKYGOCWVF AAV ACT VLRC LGI PIR VVT
SVD ILRR WHNHG CORVEY YOO CWVF AAV ACT VLRC LGI PTR VVT
SVD ILRRWKRDG CORVKYGO CWVF AAV ACT VLRC LGI PTR VVT
SVD ILRRWKTSG CORVKYGO CWVF AAV ACT VLRC LGI PTR VVT
SVD ILRRWKDYG CQRVK YGQ CWVF AAV ACT VLRC LGI PTR VVT
SVDILRRWHEHGCOOVKYGOCWVFAAVACTVLRCLGIPTRVVT
SVD ILRRWKEHG CO OVKYGO CWVF AAV ACT VLRC LGI PTR VVT
SVDILKRWRRLGCOPVKYGOCWVFAAVACTVMRCLGVPSRVVT
QNSNLLIEYFRNEFGEI QGDKSEM IWN FHCWVESWMTRPDLQP
QNSNLLIEYFRNEFGEI QGDKSEM IWN FHCWVESWMTRPDLQP
QNSNLLIEYFRNEFGEI QSDKSEM IWN FHCWVESWMTRPDLQP
QNSNLLIEYVYNEFGEM QRDKSEM IWN YHCWTESWMSRPDLQP
QNSNLLIEYFRNESGEIEGNKSEMIWN FHCWVESWMTRPDLEP
AND MT T T D V DAN D D ADT D ON UG DU TIM DUG MT D AT D A D A D
QNSNLLIEYFRNEFGELESNKSEM IWN FHCWVESWMTRPDLQP
QNSNLLIEYFRNEYGELESNKSEMIWN FHCWVESWMTRPDLQP
ING NLV IDR YLSE TGMEE RRS TDMIWNF HOWVE OWNTR PDLAP
A. AA. A

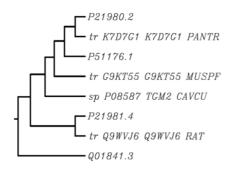
Fig. 1: The multiple sequence alignment of human, chimpanzee, bovine, European polecat, Pig, Mice, Chick. Note that conserved amino acid sequence in different organisms indicated in yellow

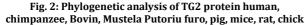
## Multiple sequence alignment

The result of multiple sequence alignment by using the CLUSTAL OMEGA showed that the amino acid residues of TG2 protein were conserved throughout the sequence. Across that 11 motifs showed maximum conserved regions. The alignment score of 98914 suggested that this protein can act as a good target protein for further work (fig. 1).

### **Phylogenetic analysis**

After multiple sequence alignment, the phylogenetic tree was constructed by using CLSTAL OMEGA. The entire sequences were used to understand the overall evolutionary pathway for diversification. The method taken for this analysis was UPGMA rooted tree; neighbour-joining tree without distance correction. The results reveal that TG2 protein is orthologs; may evolve same or new function. It has two clusters with one related to higher (Human, chimpanzee, bovine, European polecat, and pig) and other to lower (Mice, rat, chick) classes of mammals (fig. 2)





# Conserved domain identification for function prediction and protein family search

The search carried out at Pfam showed that human TG2 protein contains four domains and belongs to different families. At N-terminal transglutaminase–N–(4-122 residues) belongs to transglutaminase family, and core domain (256-359 residues) belongs to transglutaminases like superfamily. Two transglutaminase at C-terminal (473-573 and 587-686 residues) belongs to transglutaminase C-terminal Ig family. NCBI CDD server also showed the same result as that of Pfam. The BLAST search was carried out with TG2 against the non-redundant database. The query sequence (*Homo sapiens*) compared with top 100 organisms. The results indicated that, protein has highest scoring hits, 99% significant match and 0 e-value.

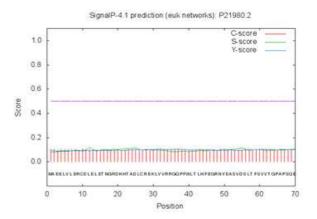
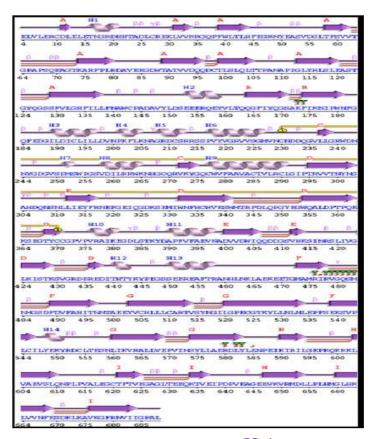


Fig. 3: Signal p output showed that human TG2 cleavage site (Cscore 0.109), signal score (S-score 0.121) and combined cleavage Site(Y-Score 0.105)



# Prediction of subcellular localization, signal peptide, and physicochemical characterization

Analysis of human TG2 protein carried out with CELLO predicted this protein to be cytosolic. Apart from being cytosolic, it is also located in extracellular, plasma membrane, Golgi complex and other cellular components. TMHMM server also predicted that protein to have a major portion cytosolic protein. The Protparam tool identified this protein have molecular weight about 77 kDa. The in vitro studies have shown that the molecular weight ranges in between 76-80kDa [42]. pI (isoelectric point) of TG2 is 5.11, a total number of negatively charged residue (Asp+Glu) 96, positively charged (Arg+Lys) are 71. Grand average of hydropathicity (GRAVY) is-0.283, aliphatic index 87.79. The negative GRAVY value of this protein indicates that it is a protein consisting of hydrophilic residues, the pI value indicates that it is possibly an acidic protein. SignalP (V 4.1) tool, indicated this protein to be a cytoplasmic membrane-associated protein and was found to have signal peptide within 1-12 residues (0.121) with cleavage site predicted between 1-30 residues (0.109) (fig. 3).

### Phosphorylation profile analysis, ASA prediction

Phosphorylation generally occurs on serine, threonine, tyrosine and histidine residues in eukaryotic proteins. Regions of human TG2 sequence showed extensive phosphorylation on serine and threonine residues, while low phosphorylation capability was predicted at tyrosine residues.

As per NetSurfP tool found that TG2 has a combination of both buried and exposed amino acid residues which signify the presence of transmembrane segments in this human TG2 protein. The RSA (Relative Surface Accessibility) value ranges from 0.021 to 0.716 indicating that this protein is present at both intra and extracellular membrane.

# Prediction of secondary structure and ramachandran plot

PDBsum analysis showed 9 sheets, 1 beta-alpha unit, 14 betahairpins, 13 beta bulges, 36 strands, 14 helices, 16 helix-helix interfaces, 57 beta turns, 6 gamma turns and 1 disulphide bond (fig. 4). The  $\Phi$  and  $\Psi$  distributions of the Ramachandran plots of glycine, proline residues are summarized (fig. 5). Altogether 98 % of the residues are in the favored region and remaining 2% in the allowed region.

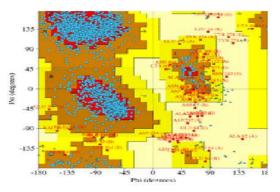


Fig. 5: Plot showing the quality of the modeled structure of the human TG2 protein 98% favoured and 2% allowed region

#### Structure prediction validation

Structure modelling of human TG2 protein (P 21980) was carried out by SWISS-MODEL server. For structure modelling, all useful parameter was considered as the template for predicting the structure. The modelled structure was opened in SWISS-PDB VIEWER (fig. 6).

## Quality of the protein

Structure validation of the predicted structure was done by feeding the predicting high-resolution crystal structures selected from the PDB into the ERRAT protein verification server. The overall quality factor obtained was 96.486. Using this technique we can differentiate between correctly and incorrectly determined regions of protein structure based on the atomic structure of the protein. It provides a useful tool for model-building and structure verification. It appears to be sensitive to errors in backbone positions on the order of 1.5 A.

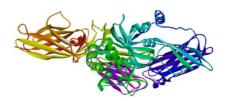


Fig. 6: Tertiary structure of human TG2 protein

#### DISCUSSION

TG2 plays various roles, but the exact mechanism of action is still not completely understood. As it has significant involvement in the pathogenesis of human diseases including autoimmune [37], neoplastic [38] and chronic inflammatory diseases [39], eventually ending into tissue fibrosis. So the rigorous investigation is in progress on this protein by both *in vitro* and *in silico* study. These observations have raised great interest in transglutaminase as a potential drug target [40].

Over the years, evidence from in vitro studies have been shown that this multifunctional protein undergoes transfers among different cell compartments are taking place such as the cytosol, the mitochondria, at both the intracellular and the extracellular surface and nucleus itself. This subcellular redistribution depends on ligands [41] and conformational changes which influence deeply the activity [42] and the physicochemical properties of the protein [43]. The in vitro study of both normal and transformed cell line models showed that TG2 promotes cell-ECM interaction which plays a critical role in cell growth, survival, migration and invasion [44, 45]. In the present study, the signal P and ASA prediction tool also showed that TG2 present in both locations. The phylogenetic analysis showed that in mammals TG2 protein may share common evolutionary characteristic both functional and structural. Human and chimpanzee are more closely related as depicted in the tree and can be shown as per the distances between these two seems to be very less. While other species were found to be distantly related with respect to the human TG2. Rat and mice also have very less distance so it can be concluded that both are closely related to each other, but in vitro study showed that human TG2 shows ~70% sequence homology with guinea pig [46]. Now TG2 has been considered as phosphor-protein in various cancer types, which would entail that TG2 may act as a target protein for upstream kinases. TG2 phosphorylation at Serine and other potentially important phosphorylation sites will provide information for the development of novel drug molecules to reduce the enhanced cancer growth associated matter [47]. Even in current study Netphos tool revealed that phosphorylation site of the human TG2 present at serine, theronine. The human TG2 has four catalytic domains and structure of this domain is similar to that of other members of TG family. The only difference seems to be in catalytic triad which is located at different sites (277,335,358 residues) on TG2 [48]. The unique GTP binding site on TG2 is located in a cleft between the catalytic core and the first  $\beta$ -barrel. The activity of TG2 is also regulated by redox potential [49]. Under the reducing conditions, the activity of TG2 seems to be increased, while it is inhibited by oxidative conditions. This redox potential also depends on disulphide. The disulfide bonds formed by cysteine triads (Cys230, Cys370, and Cys371) under oxidative become conditions inactivate. The protein model proposed in this study may be used for further docking with possible cofactors or relevant protein interactions to understand the potential mechanism of inhibitory action and molecular mechanism [50].

#### CONCLUSION

This proposed work has practical significance as it provides a foundation to not only the structure but also the post-translational

modification of this protein. Post-translational modification analysis can further be expanded to obtain new insights into the foundation of conformational changes in the cellular environment and also down-regulation of TG2. The structure can be used for interaction study with co-factors or other proteins on the cell to throw light on interaction mechanism.

### AUTHORS CONTRIBUTION

Prachi Parvatikar performed research, analyzed data and contributed to writing the paper. Shivkumar Madagi conceived the project, designed and analyzed.

#### **CONFLICTS OF INTERESTS**

The author has no conflicts of interest

## REFERENCES

- 1. Eckert RL, Kaartinen MT, Nurminskaya M, Belkin AM, Colak G, Johnson GV, *et al.* Transglutaminase regulation of cell function. Physiol Rev 2014;94:383-417.
- 2. Achyuthan KE, Greenberg CS. Identification of a guanosine triphosphate-binding site on guinea pig liver transglutaminase. The role of GTP and calcium ions in modulating activity. J Biol Chem 1987 5;262:1901-6.
- Mangala LS, Mehta K. Tissue transglutaminase (TG2) in cancer biology. Prog Exp Tumor Res 2005;38:125–38.
- Csosz E, Bagossi P, Nagy Z, Dosztanyi Z, Simon I, Fesus L. Substrate preference of transglutaminase 2 revealed by logistic regression analysis and intrinsic disorder examination. J Med Bacteriol 2008;383:390-2.
- Chhabra A, Verma A, Mehta K. Tissue transglutaminase promotes or suppresses tumors depending on cell context. Anticancer Res 2009;29:1909-19.
- Iismaa SE, Mearns BM, Lorand L, Graham RM. Transglutaminases and disease: lessons from genetically engineered mouse models and inherited disorders. Phys Res 2009;89:991-1023.
- Antonyak MA, Li B, Regan AD, Feng Q, Dusaban SS, Cerione RA. Tissue transglutaminase is an essential participant in the epidermal growth factor-stimulated signalling pathway leading to cancer cell migration and invasion. J Biol Chem 2009;284:17914-25.
- Nanda N, Iismaa SE, Owens WA, Husain A, Mackay F, Graham RM. Targeted inactivation of Gh/tissue transglutaminase II. J Biol Chem 2001;276:20673-8.
- Hasegawa G, Motoi SU, Ichikawa Y, Ohtsuka T, Kumagai S, Kikuchi M, *et al.* A novel function of tissue-type transglutaminase: protein disulphide isomerase. Biochemical J 2003;373:793-3.
- Mishra S, Murphy LJ. The p53 oncoprotein is a substrate for tissue transglutaminase kinase activity. Biochem Biophys Res Commu 2006;339:726-30.
- 11. Wang Y, Ande SR, Mishra S. Phosphorylation of transglutaminase 2 (TG2) at serine-216 has a role in TG2 mediated activation of nuclear factor-kappa B and in the downregulation of PTEN. BMC Cancer 2012;12:277.
- 12. Mishra S, Saleh A, Espino PS, Davie JR, Murphy LJ. Phosphorylation of histones by tissue transglutaminase. J Biol Chem 2006;281:5532-8.
- Mishra S, Murphy LJ. Phosphorylation of transglutaminase 2 by PKA at Ser216 creates 14-3-3 binding sites. Biochem Biophys Res Commun 2006;347:1166-70.
- 14. Park H, Park ES, Lee HS, Yun HY, Kwon NS, Baek KJ. A distinct characteristics of  $G\alpha h$  (transglutaminase II) by compartment: GTPase and transglutaminase activities. Biochem Biophys Res Commun 2001;284:496-500.
- 15. Park D, Choi SS, Ha KS. Transglutaminase 2: a multi-functional protein in multiple subcellular compartments. Amino Acids 2010;39:619-31.
- Pinkas DM, Strop P, Brunger AT, Khosla C. Transglutaminase 2 undergoes a large conformational change upon activation. PLoS Biol 2007;5:327.
- 17. Kanchan K, Fuxreiter M, Fésüs L. Physiological, pathological, and structural implications of non-enzymatic protein-protein

interactions of the multifunctional human transglutaminase 2. Cell Mol Life Sci 2015;72:3009-35.

- Liu S, Cerione RA, Clardy J. Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. Proceedings Nat Acad Sci 2002;99:2743-7.
- Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelialmesenchymal transitions in development and disease. Cell 2009;139:871-90.
- 20. Hoffner G, Djian P. Transglutaminase and diseases of the central nervous system. Front Biosci 2005;10:3078-92.
- 21. Lai TS, Liu Y, Tucker T, Daniel KR, Sane DC, Toone E, *et al.* Identification of chemical inhibitors to human tissue transglutaminase by screening existing drug libraries. Chem Biol 2008;15:969-78.
- Jang TH, Lee DS, Choi K, Jeong EM, Kim IG, Kim YW, et al. Crystal structure of transglutaminase 2 with GTP complex and amino acid sequence evidence of the evolution of GTP binding site. PLoS One 2014;9:7005.
- 23. Han BG, Cho JW, Cho YD, Jeong KC, Kim SY, Lee BI. Crystal structure of human transglutaminase 2 in complex with adenosine triphosphate. Int J Biomater 2010;47:190-5.
- 24. Bergamini CM, Dondi A, Lanzara V, Squerzanti M, Cervellati C, Montin K, *et al.* Thermodynamics of binding of regulatory ligands to tissue transglutaminase. Amino Acids 2010;39:297-304.
- McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, *et al*. Analysis tool web services from the EMBL-EBI. Nucleic Acids Res 2013;41:597-600.
- Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The pfam protein families' database: towards a more sustainable future. Nucleic Acids Res 2016;44:279-85.
- Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, et al. BLAST: a more efficient report with usability improvements. Nucleic Acids Res 2013;41:29-33.
- Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Pro Sci 2004;13:1402-6.
- 29. Petersen TN, Brunak S, von Heijne G, Nielsen H. Signal P 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 2011;8:785-6.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, *et al.* Protein identification and analysis tools on the ExPASy server. Humana Press; 2005.
- Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 1999;294:1351-62.
- Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Struct Biol 2009;9:51.
- Possner DD, Claesson M, Guy JE. Structure of the glycosyltransferase Ktr4p from saccharomyces cerevisiae. PloS One 2015;10:0136239. https://doi.org/10.1371/journal.pone. 0136239
- Hintze BJ, Lewis SM, Richardson JS, Richardson DC. Molprobity's ultimate rotamer-library distributions for model validation. Proteins: Struct Funct Bioinf 2016;84:1177-89.
- Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 2014;42:W252-8.
- 36. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Proteins Sci 1993;2:1511-9.
- Henderson KN, Tye-Din JA, Reid HH, Chen Z, Borg NA, Beissbarth T, *et al.* A structural and immunological basis for the role of human leukocyte antigen DQ8 in celiac disease. Immunity 2007;27:23-34.
- Mangala LS, Mehta K. Tissue transglutaminase (TG2) in cancer biology. In: Transglutaminases. Karger Pub; 2005. p. 125-38.
- Sohn J, Kim TI, Yoon YH, Kim JY, Kim SY. Novel transglutaminase inhibitors reverse the inflammation of allergic conjunctivitis. J Clin Inves 2003;111:121-8.
- 40. Siegel M, Khosla C. Transglutaminase 2 inhibitors and their therapeutic role in disease states. Pharmacol Ther 2007;115:232-45.

- 41. Mariani P, Carsughi F, Spinozzi F, Romanzetti S, Meier G, Casadio R, *et al.* Ligand-induced conformational changes in tissue transglutaminase: Monte Carlo analysis of small-angle scattering data. Biophys J 2000;78:3240–51.
- 42. Milakovic T, Tucholski J, McCoy E, Johnson GV. Intracellular localization and activity state of tissue transglutaminase differentially impacts cell death. J Biol Chem 2004;279: 8715-22.
- 43. Bergamini CM, Signorini M. Purification of testicular transglutaminase by hydrophobic chromatography on phenyl-sepharose. Biochem Int 1992;27:557-65.
- 44. Huang L, Haylor JL, Hau Z, Jones RA, Vickers ME, Wagner B, *et al.* Transglutaminase inhibition ameliorates experimental diabetic nephropathy. Kidney Int 2009;76:383-94.
- Johnson TS, El-Koraie AF, Skill NJ, Baddour NM, El Nahas AM, Njloma M, *et al.* Tissue transglutaminase and the progression of human renal scarring. J Am Soc Nephrol 2003;14:2052-62.
- 46. Chica RA, Gagnon P, Keillor JW, Pelletier JN. Tissue transglutaminase acylation: the proposed role of conserved active site Tyr and Trp residues revealed by molecular

modeling of peptide substrate binding. Proteins Sci 2004;13:979-91.

- 47. Wang Y, Ande SR, Mishra S. Overexpression of phospho mutant forms of transglutaminase 2 downregulates epidermal growth factor receptor. Biochem Biophys Res Commun 2012;417:251–5.
- Begg GE, Holman SR, Stokes PH, Matthews JM, Graham RM, *et al.* Mutation of a critical arginine in the GTP-binding site of transglutaminase 2 disinhibits intracellular cross-linking activity. J Biol Chem 2006;281:12603–9.
- Stamnaes J, Pinkas DM, Fleckenstein B, Khosla C, Sollid LM. Redox regulation of transglutaminase 2 activity. J Biol Chem 2010;285:25402–9.
- 50. Umashankar V, Gurunathan S. Drug discovery: an appraisal. Int J Pharm Pharm Sci 2015;7:59-66.

### How to cite this article

• Shivkumar B Madagi, Prachi P Parvatikar. Sequence analysis and structural characterization of tissue transglutaminase 2(TG2) by *In silico* approach. Int J Pharm Pharm Sci 2017;9(10):37-42.