

# Molecular characterization, modeling, in silico analysis of equine pituitary gonadotropin alpha subunit and docking interaction studies with ganirelix

Anuradha Bhardwaj<sup>1</sup>  · Varij Nayan<sup>2</sup>  · Parvati Sharma<sup>1</sup> · Sanjay Kumar<sup>3</sup> · Yash Pal<sup>4</sup> · Jitender Singh<sup>4</sup>

Received: 10 May 2017 / Accepted: 6 July 2017  
© Springer-Verlag GmbH Germany 2017

**Abstract** Equine pituitary gonadotropins (eLH, eFSH, eCG) are heterodimeric glycoprotein hormones with alpha ( $\alpha$ ) and beta ( $\beta$ ) subunits. It is responsible for maintenance of pregnancy in mares during early gestation and fairly valuable for inducing superovulation in animals other than equines. The alpha subunit is common, while beta subunit is species-specific in all glycoprotein hormones. In the present investigation, molecular cloning and in silico characterization including homology modeling and molecular docking analysis of the equine chorionic gonadotropin (eCG) alpha subunit was carried out for gaining structural and functional insights into the eCG alpha subunit and its possible interaction with ganirelix, a gonadotropin-releasing hormone (GnRH) antagonist. The equine

chorionic gonadotropin (eCG) alpha subunit expressed in pituitary gland was selected, amplified from total RNA, cloned and sequenced. The in silico analyses were made for homology modelling, structural details, epitope identification and chromosomal localization. Molecular docking studies of eCG alpha were undertaken with a drug ganirelix which is used to control ovulation and has antagonistic activity against GnRH. The protein sequence corresponding to selected open reading frame (ORF) was 99–100% similar with domesticated horse, Przewalski's horse, and 92–93% with Burchell's zebra and donkey. Molecular docking studies revealed the possible interaction of eCG alpha with ganirelix. The possible drug-macromolecule interactions were visualized between eCG alpha and ganirelix. The study will provide structural insight into unique sites and an alternate route of gonadotropin suppression applicable to assisted reproductive technologies.

Anuradha Bhardwaj and Varij Nayan contributed equally to this work.

**Electronic supplementary material** The online version of this article (doi:[10.1007/s40203-017-0025-1](https://doi.org/10.1007/s40203-017-0025-1)) contains supplementary material, which is available to authorized users.

✉ Anuradha Bhardwaj  
dranu.biotech@gmail.com;  
Anuradha.Bhardwaj@icar.gov.in

Varij Nayan  
varij.biochem@gmail.com

Parvati Sharma  
parvati.hsr@gmail.com

Sanjay Kumar  
kumarsanjay66@yahoo.com

Yash Pal  
yashpal1888@gmail.com

Jitender Singh  
jschhokar@yahoo.co.in

**Keywords** Equine · Gonadotropin · *In silico* · Ganirelix · Docking

<sup>1</sup> Basic and Supporting Discipline Unit (BSDU), ICAR-National Research Centre on Equines, Hisar, Haryana 125001, India

<sup>2</sup> Animal Physiology and Reproduction (APR) Division, ICAR-Central Institute for Research on Buffaloes, Hisar, Haryana 125001, India

<sup>3</sup> Equine Health Unit (EHU), ICAR-National Research Centre on Equines, Hisar, Haryana 125001, India

<sup>4</sup> Equine Production Centre, ICAR-National Research Centre on Equines, Bikaner, Rajasthan 334 001, India

## Abbreviations

eCG	Equine chorionic gonadotropin
PMSG	Pregnant mare serum gonadotropin
GnRH	Gonadotropin-releasing hormone
LH	Luteinizing hormone
FSH	Follicle stimulating hormone
ORF	Open reading frame
RNA	Ribonucleic acid
cDNA	Complementary deoxyribonucleic acid
TAE	Tris-acetate EDTA buffer
PCR	Polymerase chain reaction
LB	Luria broth
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
NCBI	National center for biotechnology information
PDB	Protein data bank
IEDB	Immune epitope database
SEM	Standard error of the mean
IVF	<i>In vitro</i> fertilization

## Background

Reproduction is a complex but extensively studied process. Akin to other mammalian species, in equines, several physiological events coupled with reproduction are extremely coordinated, synchronized and regulated by the endocrine system, i.e. a series of communications of the hypothalamic–pituitary–gonadal factors. The pituitary gland is truly dubbed as the “master player of endocrine orchestra” of the living mammalian body and act as the principal endocrine regulator of reproduction, growth and metabolism as well as other environmental conditions including response to stress. Distinctive cell types in the anterior pituitary gland including somatotrophs and gonadotrophs, secrete polypeptide hormones such as prolactin, growth hormone and the gonadotropins viz. luteinizing (LH) and follicle stimulating (FSH) hormones (Kalra et al. 1997; Gregory et al. 2000). Pregnant mare serum gonadotropin (PMSG), also known as equine chorionic gonadotropin (eCG) is responsible for maintenance of pregnancy in mares during early gestation (Christakos and Bahl 1979). Besides, this hormone is fairly valuable for inducing superovulation in animals other than equines. eCG is composed of two non-similar and non-covalently associated  $\alpha$  and  $\beta$  subunits (Pierce and Parsons 1981). Within the same species, the  $\alpha$  subunit is encoded by a single gene while different genes encode  $\beta$  subunits providing receptor binding specificity to the glycoprotein hormone heterodimers thereby offering the biological specificity to each hormone (Combarnous 1992). However in contrast to primates and other farm animals, in horses, both LH and CG  $\beta$  subunits are encoded by the same gene (Sherman et al. 1992). Both the pituitary eLH and placental

eCG display the dual LH and FSH like activity in non-equine species with identical FSH/LH activity ratios (Chopineau et al. 1993). Although eCG and eLH exhibit identical  $\alpha$  and  $\beta$  polypeptide chains, their carbohydrate contents are different. Undeniably, among all the glycoprotein hormones, eCG is the most heavily glycosylated with 45% carbohydrate by weight versus 30% for eLH (Bousfield et al. 1996; Hokke et al. 1994). The nature of oligosaccharide side chains plays an imperative role in the processing and biological activity of the respective hormones (Ulloa-Aguirre et al. 2001). The chorionic gonadotropin, expressed in the placenta of higher primates (Maston and Ruvolo 2002), is indispensable for positive advancement of pregnancy. CG mediates its action through the LH/CG receptor, and is responsible for preventing regression of the corpus luteum. This stimulates sustained progesterone production that upholds the uterine lining in a physiological stage receptive for implantation. A 621 base pair fragment of the cDNA for the  $\alpha$ -subunit of human chorionic gonadotropin was isolated by cloning in a plasmid vector, and the complete nucleotide sequence was determined (Fiddes and Goodman 1979). They also reported that the four human glycoprotein hormones are encoded by common alpha subunit gene (Fiddes and Goodman 1981). Min et al. (1994) gave first report on nucleotide sequence and placental expression of eCG alpha subunit. They determined the full length of the equine  $\alpha$  subunit cDNA sequence and investigated the expression ratio of  $\alpha$  and  $\beta$  subunit mRNA in the equine placenta on day 70 of gestation. Not only in equines, but other primates and non-primates pituitary glycoprotein alpha subunits were studied, compared with other reference sequences and characterized (Abel et al. 2014; Weltzien et al. 2003; Chi et al. 2015; Chaube et al. 2015). From time to time, various reports on eCG alpha subunit have been published. However, to our knowledge, this is the first report on three dimension structure prediction of eCG alpha protein, its antigenic propensity score and its interaction with potential drug molecule which may modulate the *in vivo* activity of the protein. Ganirelix is a synthetic decapeptide with high antagonistic activity against naturally occurring gonadotropin-releasing hormone (GnRH). It is mainly used in assisted reproductive technology to control ovulation. Ganirelix prevents ovulation until it is triggered by injecting human chorionic gonadotropin (hCG). Ganirelix is also known to interact with hCG. Treatment with ganirelix effectively prevents premature LH rises (Lambalk et al. 2006). Therefore, it was interesting to assess the possible drug-macromolecule interactions between ganirelix and eCG alpha. The present study will pave ways and provide insight into unique sites in alpha subunit of pituitary gonadotropin that will be useful in future studies

based on equine gonadotropins as well as in assisted reproductive technologies.

## Methods

### Animals, sampling and ethics statement

Pituitary glands were collected after natural death of mares at the organized equine farm of NRCE, Bikaner. All the procedures for sample collection were performed according to the approval by the Institute Animal Ethics Committee (IAEC) of ICAR-National Research Centre on Equines vide NRCE/CPCSEA/2012-13/dated the 18th July 2012.

### RNA isolation and cDNA synthesis

RNA extraction was done with RNAqueous Kit (Ambion Inc., Austin, Texas) as per manufacturer's instructions. Total RNA was quantified and absorbance was measured at 260 and 280 nm with a UV-spectrophotometer (Nanodrop ND-1000, USA). The integrity of RNA preparation was evaluated using agarose (1%) gel electrophoresis (Tarson, India) after its denaturation with glyoxal and dimethyl sulfoxide (DMSO) (McMaster and Carmichael 1977) in 1xTAE buffer at 60 V for 30 min (EPS-Pharmacia-500/400). The High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) was used for cDNA synthesis following the manufacturer's instructions.

### PCR amplification of alpha subunit of equine gonadotropin

For amplification of  $\alpha$  subunit of eCG, reported primers were selected (Galet et al. 2000) as forward- 5'-CGGAATTTCGCTCCTGATGTGCAGGATTGCC-3' (30 bp) and reverse- 5' GCTCTAGACCTTTAAATCT TGTGGTGAT-3' (28 bp). The PCR conditions were: initial denaturation at 94 °C for 4 min, cycling denaturation at 94 °C for 30 s, annealing at 57 °C for 45 s, extension at 72 °C for 45 s, followed by 35 cycles with a final extension step at 72 °C for 10 min. The PCR product (~300 bp) was checked in agarose gel electrophoresis.

### Cloning of alpha eCG gene in TOPO-TA cloning vector

The TOPO-TA vector (Invitrogen, CA) was used for cloning the PCR product according to the manufacturer's instructions. The PCR product to be ligated was purified using Qiagen gel purification kit. For ligation of  $\alpha$ -eCG

gene into the TOPO-TA vector, the concentration of both the product and the vector was kept in the molar Ratio of 4:1. The well isolated blue and white colonies were observed. The white colonies were picked from the master plate and sub-cultured on LB agar plates with antibiotics such as ampicillin (100 µg/ml) and tetracycline (50 µg/ml) as well as X-gal (40 µl of 20 mg/ml concentration) and IPTG. The colonies, which were able to grow on selective antibiotics, were considered as transformed. Recombinant plasmids were isolated from *E. coli DH-5 alpha* cells (harbouring the plasmid  $\alpha$ -eCG-TOPO-TA) by alkaline lysis method. For amplification of  $\alpha$ -eCG gene colony PCR was carried out. The PCR products were analyzed in agarose gel electrophoresis along with DNA molecular weight marker.

### Custom sequencing of cloned alpha-eCG gene and analysis

The purified PCR product and recombinant plasmid with cloned eCG alpha gene was sent for custom sequencing to Xcleris genomics lab, Ahmadabad, India. The sequence obtained was analysed with CLC sequence viewer v 7.6.1 and similarity search was done with the reported sequences retrieved from the NCBI.

### In silico analysis of alpha-eCG

#### Sequence retrieval and analysis

The analysis of the translated amino acid sequences of the ORF corresponding to the cDNA coding for alpha eCG (partial) was done by ORF Finder, coding potential calculator (CPC), NCBI BLAST as well as other bioinformatics tools such as CLC sequence viewer, DNASTAR Lasergene software suite version 7.1 and PSIPRED bioinformatics softwares and other tools described in the following section. The deduced amino acid sequence (79 amino acid) of alpha subunit of eCG in FASTA format was used as the query sequence in blastp (Protein-Protein Blast). The other sequences were retrieved from NCBI (<http://ncbi.nlm.nih.gov/protein/>) database in FASTA format. Protein domain families were analyzed by NCBI Conserved Domain search (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and Pfam (<http://pfam.xfam.org/>) server. Secondary structural analyses were performed by using PSIPREDv3.0 (McGuffin et al. 2000) of the PSIPRED protein structure prediction web server at UCL department of computer science. Further, several mammalian sequences for the alpha eCG and precursors were retrieved from the NCBI data base as follows: *Equus caballus* (modern domesticated horse; XP\_005596847.1, NP\_001093233.1, AAA57252.1), *Equus przewalskii* (the

endangered Przewalski's horse, XP\_008513373.1), *Equus burchellii* (Burchell's zebra, O46642.1), *Equus acinus* (donkey, XP\_014703937.1), cat (NP\_001041623.1), lesser hedgehog tenrec (XP\_004696323.1), bat (XP\_008144704.1, XP\_006773539.1), Coquerel's sifaka lemur (XP\_012493635.1), rabbit (1110225A), walrus (XP\_004411807), polar bear (XP\_008685589), dog (NP\_001002988), amur tiger (NP\_001277528), pig (NP\_999611.1), mice (NP\_034019.1), brushtail possum (AAC63900.1), African elephant (XP\_003404412), galago (XP\_003803482), tarsier (XP\_008066056), marmoset (NP\_001254689.1), Arabian camel (XP\_010986025.1), sheep (NP\_001009464.1), cattle (NP\_776326.1) and goat (NP\_001272613.1). Sequences were aligned with the help of CLC Sequence Viewer v7.7 (QIAGEN Aarhus A/S) and the sequence alignment was saved. The sequence alignment file was again used in the MEGA 7.0.14 program. Molecular phylogenetic analysis was performed from this data following the Maximum-Likelihood (ML) method for 1000 bootstrap replications. The evolutionary history was inferred by using the Maximum Likelihood method based on the Jones-Taylor-Thornton (JTT) matrix-based model (Kumar et al. 1994). The tree with the highest log likelihood (−530.7622) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 27 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 78 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Caspermeyer 2016). To understand the genome organization of the eCG alpha gene, the chromosomal location was also determined based on the *Equus caballus* genomic sequence annotation from NCBI (*Equus caballus* isolate Twilight breed thoroughbred chromosome 10, EquCab2.0; NC\_009153).

### 3D modeling

3D models were predicted by using Geno3D (automatic molecular modeling tool) server of PBIL-IBCP Lyon-Gerland (<https://geno3d-prabi.ibcp.fr/>) (Combet et al. 2002). The model generated was evaluated in terms of Ramachandran plot, main and side chain parameters and residue properties by PROCHECK (Laskowski et al. 1993), Z-score and QMEAN score. The 3D homology model was checked by Procheck and assessed by web servers such as

ProSA-Web (Wiederstein and Sippl 2007) and QMEAN (Benkert et al. 2009) for model quality estimation. To assess the quality of model, QMEAN Z-score overall quality factor was used and QMEAN score for density plot was used to assess the organization of various atoms. SWISS-MODEL was also used to assess the quality and structural features of the developed model and local model quality estimation was done by ANOLEA (atomic non-local environment assessment) (Melo et al. 1997), GROMOS (Christen et al. 2005) and QMEAN6.

Evaluation of the structure as well as other stereochemical analyses were accomplished by online available Geno3D (Combet et al. 2002) and VADAR version 1.8 (<http://vadar.wishartlab.com/>) (Willard et al. 2003) servers. The model was visualized by using 3-D Molecule Viewer of Vector NTI Advance, Version 11.5.2 (Life Technologies) and PyMol molecular visualization software (Schrodinger 2010). The probable binding sites on the developed eCG alpha model are very important and therefore, the active Sites of the protein were analysed by MetaPocket 2.0 (<http://projects.biotec.tu-dresden.de/metapocket/index.php>) (Huang 2009).

### Virtual screening and molecular docking

Molecular docking was executed to reveal the interactions of ganirelix towards its target protein namely the eCG alpha. PyRx-Virtual screening tool version 0.8 (pyrx.sourceforge.net, Scripps Institute) (Dallakyan and Olson 2015) was used for virtual molecular screening and docking with the lead compound ganirelix (Additional file 2: Table S1). As mentioned earlier, it is a synthetic decapeptide with high antagonistic activity against naturally occurring gonadotropin-releasing hormone (GnRH). The PDB file of the eCG alpha 3D homology model earlier developed, was used as the macromolecule. The 2D structure file was downloaded from the Pubchem open chemistry database available at NCBI in sdf/xml format. OpenBabelGUI (O'Boyle et al. 2011) was used to create all the input files in.pdbqt format and saved. Before docking, PDB file of eCG alpha 3D model was energy-minimized and prepared by removal of water molecules, addition of polar hydrogens and Gasteiger charges. To conclude, corrected PDB files were saved in.pdbqt format. AutoDock wizard was used and the ligand and macromolecule were selected. A grid box with a size of 63.93 Å (X), 42.69 Å (Y) and 25.0 Å (Z) was used. The centre of the grid box was allocated using x, y, z coordinates of −2.72236290175, 1.41345873493 and 0.6341, respectively. Molecular docking was performed using AutoDock Vina, which is a part of the PyRx 0.8 software. Docking poses of ganirelix–eCG alpha docked structures were visualized using the PyMOL v 1.8.2.0 software 32 and the



protein–ligand interactions were generated using the LigPlot<sup>+</sup> v.1.4 software (Wallace et al. 1995).

### Structure-based prediction of antibody epitopes

Though, the reliable prediction of epitopes is very exciting and challenging, it provides great opportunity for the design and fabrication of immunodiagnostics. Therefore, Linear and discontinuous B cell epitopes were predicted using the ElliPro: Antibody Epitope Prediction server of the IEDB analysis resource (Ponomarenko et al. 2008). The SCRATCH Protein Predictor (<http://scratch.proteomics.ics.uci.edu/>) was used for ascertaining the antigenicity propensity score (Cheng et al. 2005). The eCG alpha protein sequence was also analysed from a diverse selection of methods for secondary structure, hydropathy, antigenicity, amphiphilicity, surface probability and flexibility by using the Protean Software (for Protein Structure Analysis and Prediction) of the DNASTAR Lasergene software suite version 7.1 (Burland 2000).

## Results

### Amplification and cloning of alpha subunit of eCG

The mare's pituitary was successfully processed for total RNA isolation and cDNA synthesis. The purity of RNA preparation was determined by finding the  $A_{260}/A_{280}$  ratio which was found to be in the range of 1.8–2.0. The cDNA was synthesized and used for amplification of alpha-eCG of approximately 300 bp by using reported primers corresponding to equine chorionic gonadotropin hormone alpha subunit (Fig. 1a).

### The sequence alignment, phylogenetic and chromosomal localization analyses of alpha subunit of eCG

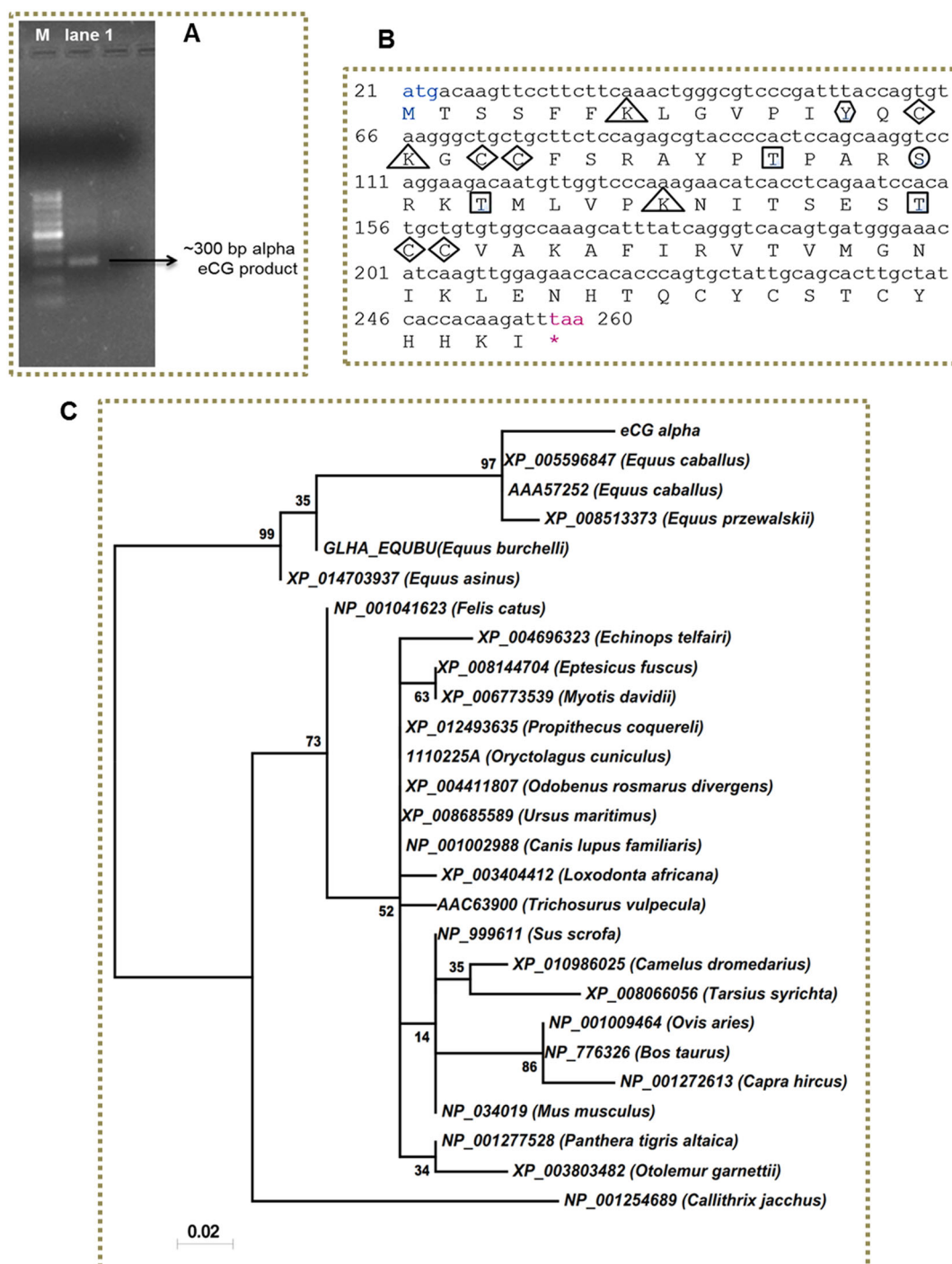
The sequencing of the PCR product and recombinant plasmid with cloned eCG alpha gene revealed a nucleotide sequence of 306 bp including the sequences corresponding to the primers as well as those of sites for restriction enzymes (*EcoRI* and *XbaI*) (Additional file 1: Figure S1). The ORF Finder (Open Reading Frame Finder) of NCBI was used to find the ORF corresponding to the nucleotide sequence. Three ORFs were obtained with 79, 67 and 48 amino acids excluding the primer sequences. These ORFs were subjected to protein blast at NCBI and preferred based on the maximum sequence identity. Thereafter, the corresponding amino acid sequence (79 amino acid) of alpha subunit of eCG was selected (Fig. 1b) and subjected for further in silico analysis.

The nucleotide sequences corresponding to the selected ORF showed sequence similarity of 99, 97 and 96 percent with the reported glycoprotein hormone or chorionic gonadotropin alpha sequences of *Equus caballus* (horse) as well as *Equus przewalskii* (Przewalski's Horse or Dzungarian Horse), *Equus burchelli* (Plains Zebra) and *Equus asinus* (Donkey) respectively. The translated ORF sequence was: **MTSSFFKLGVPYQCKGCCFSRAYPTPARSRKTM LVPKNITSESTCCVAKAFIRVTVM GNIKLENHTQCYCSTCYHHKI\***.

Putative conserved domains characteristics of glycoprotein hormone, i.e. Hormone\_6 super family (actual alignment was observed with pfam00236 member of the superfamily cl02448) were detected in the NCBI Conserved Domain search. In Pfam searches also significant matches were observed with Hormone\_6 family (PF00236) and CL0079 clan. The sequence obtained was compared to the reference sequences of other species available at GenBank. The high similarity was observed between alpha eCG and other reported sequences (Additional file 1: Figure S2).

The alpha-subunit of eCG was found to be highly conserved among equines and other species as suggested by earlier reports (Cahoreau et al. 2015). The alpha-subunit of eCG sequence was also analyzed for putative ORF sequences and detailed RE map was generated for observing putative sites for restriction enzymes which may interfere while cloning the gene in vector. The alpha subunit of eCG is formed of 24 amino acid signal sequence and a mature protein of 96 amino acids. In the present study only the mature polypeptide was selected for amplification and based on the primers used, the amplification product coding for an ORF of 79 amino acids was used and was devoid of the signal peptide. However, further the amino acid sequences, MTSS were also left based on the conservation among sequences in various species.

As shown in the Additional file 1: Figure S2, the alpha subunit sequence was found to be about 99–100% similar with the gonadotropin alpha subunit sequences or precursors of modern domesticated horse and the endangered Przewalski's horse, and 92–93% with Burchell's zebra and the donkey. Similarly, a sequence identity of 84% (cat and lesser hedgehog tenrec), 82% (bat, Coquerel's sifaka lemur, rabbit, walrus, polar bear and dog), 81% (amur tiger, pig, mice, brushtail possum and African elephant), 80% (galago and tarsier), 79% (marmoset), 78% (Arabian camel, sheep and cattle) and 77% (goat) respectively, was observed with the reported gonadotropin alpha and precursor sequences. To visualize the phylogenetic relationship of the deduced ORF amino acid sequences of alpha subunit of eCG gene with those of other sequences known in other mammalian species, we analyzed it with other 26 sequences mentioned earlier, i.e. 5 of equid family



**Fig. 1** PCR amplification, sequence elucidation and phylogenetic analyses of alpha subunit of eCG (**a**). The PCR amplification of alpha-eCG showing band at ~300 bp in agarose gel electrophoresis. Lane M 100 bp Marker; Lane 1 amplified  $\alpha$ -eCG. **b** The nucleotide sequences corresponding to the selected ORF with deduced amino acids. The putative phosphorylation sites are Serine (position 30), Threonine (position 26, 33 and 41), and Tyrosine (position 13). The Serine residues are shown as circles, Threonine as solid rectangles, and tyrosine as hexagon. The position of triangles indicates putative glycation sites, and diamonds indicates predicted palmitoylation sites. The position of asterisk indicates the end of open reading frame. **c** Phylogenetic relationship of the equine chorionic gonadotropin

alpha (eCG alpha) protein sequence with other 26 reported protein sequences of different mammals including other equines. The data provided by the multiple sequence alignment patterns of amino acids were used for the molecular phylogenetic analysis. In the phylogenetic tree, the sources of all the sequence data are provided as accession numbers and species names. The phylogeny is an un-rooted tree recovered using the maximum-likelihood (ML) method based on 1000 bootstrap replications and Jones–Taylor–Thornton (JTT) matrix-based model by the MEGA 7.0.14 program. The tree with the highest log likelihood (−530.7622) is shown here. Bootstrap values greater than >0% is shown on the branch

members and those of cat, lesser hedgehog tenrec, bat, Sifaka lemur, rabbit, walrus, polar bear, dog, tiger, pig, mouse, brushtail possum, marmoset, camel, elephant, galago, tarsier, sheep, cattle and goat. Based on multiple sequence alignment of the alpha subunit of eCG with other sequences of these selected species, the phylogenetic tree was constructed. The result is shown in Fig. 1c.

*Equus caballus* genomic sequence annotation from NCBI (*Equus caballus* isolate Twilight breed thoroughbred chromosome 10, EquCab2.0; NC\_009153) was used for determining the genome organization of the eCG alpha gene and the chromosomal location was determined. As shown in the Additional file 1: Figure S3, the eCG alpha genes were found to be distributed on chromosome 10 at 2 places based on the matches (aligned regions spanning positions from 39938737–39938896 and 39937910–39937996). Similar results were evident while using Ensemble BLAST. Searches for these regions in Ensemble displayed the CGA gene (ENSECAG00000018841) for eCG alpha at chromosome 10 and a synteny between horse chromosome 10 (position 39937900–39940063) and Human chromosome 6 (position 87085498–87095406) and dog chromosome 12 (position 46610973–46613039) (Additional file 1: Figure S4).

### Generation, validation and visualization of the 3D model of the eCG alpha

Geno3D (Automatic molecular modeling tool) of PBIL-IBCP Lyon-Gerland was used for construction of the homology model of eCG alpha subunit. NCBI BLAST search of the sequence of its translated protein sequence chain was also performed against the PDB database to find a suitable template. The sequence chain to be modelled was selected based on similarity search as “SSFFKLGVPIYQCKGCCFSRAYPTPARSRKTMLVPK NITSESTCCVAKAFIRVTVMGNIKLENHTQCYCSTCY HHK” and modelled using the four selected templates (pdb1dz7A\_0, pdb1e9jA\_0, pdb1hd4A\_0 and pdb1xwdA\_0) with an identity of 73.3 percent. The mean deviation between the four templates on this chain was 6.123312 Å. A Ramachandran plot of the predicted eCG alpha model (Additional file 1: Figure S5) indicated the region of possible angle formations by  $\phi$  (phi) and  $\psi$  (psi) angles. Ramachandran plot of the eCG alpha 3D model showed 95.4% residues in favoured and allowed regions, whereas 3 residues were in disallowed regions.

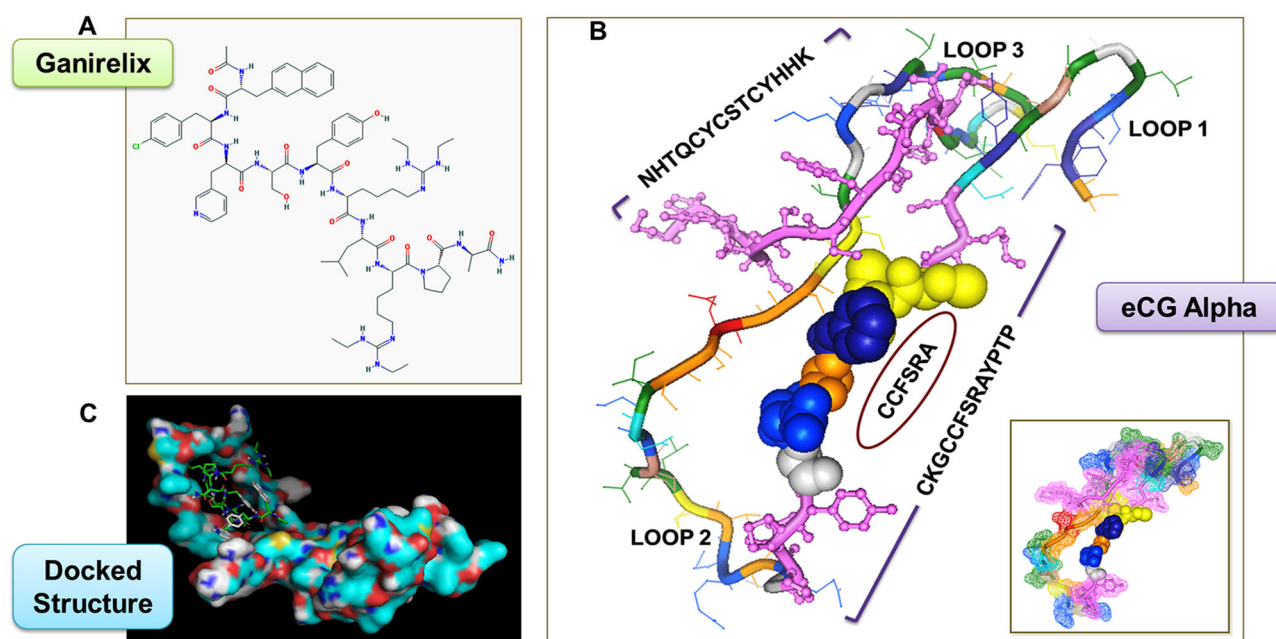
The ProSA was used to check the 3D homology model of eCG alpha for potential errors. As illustrated in the Additional file 1: Figure S6, the ProSA presented its two characteristics, i.e. the Z-score and a plot of its residue scores which shows local model quality by plotting energies as a function of amino acid sequence

position. The ProSA Z-Score of  $-2.29$  (highlighted as black dot) and the better residue scores plot were indicative of the modal quality. ProSA-web z-score plot showed that the z-score of the model was within the range of scores for similar protein structures evident by NMR spectroscopy and X-Ray crystallography. ProSA-web plot of residue scores of the modelled structure showed the local model quality by plotting energies as a function of amino acid sequence position. No positive values corresponding to problematic parts of the input structure in window size 40 which indicates average energy over each 40 residue fragment (thick line). The second line with a smaller window size of 10 residues as background (thin line) also exhibited a better local model quality.

The eCG alpha model was also assessed for the quality and structural features by SWISS-MODEL (Additional file 1: Figure S7). The ANOLEA (atomic non local environment assessment) is used to assess packing quality of the models. The program performs energy calculations on a protein chain, evaluating the “Non-Local Environment” (NLE) of each heavy atom in the molecule. QMEAN is a composite scoring function for both the estimation of the global quality of the entire model as well as for the local per-residue analysis of different regions within a model. GROMOS is a general-purpose molecular dynamics computer simulation package for the study of biomolecular systems and can be applied to the analysis of conformations obtained by experiment or by computer simulation. The y-axis of the plot represents the energy for each amino acid of the protein chain. Negative energy values (in green) represent favourable energy environment whereas positive values (in red) unfavourable energy environment for a given amino acid.

The QMEAN score of the eCG alpha homology model was matched to scores of X-ray crystallography resolved reference structures. A highly resolved structure has the zero average Z-score. A total QMEAN Score of 0.438 (Z-score:  $-2.00$ ) for estimated absolute quality was found for the eCG alpha indicating a fairly good model was also evident (Additional file 1: Figure S8). The QMEAN score along with residue error plot data estimated per-residue inaccuracies along the sequence; support the validity of the predicted model.

The homology model was thereafter visualized by 3-D Molecule Viewer (a component of Vector NTI Advance 11.5.2). The 3D Homology Model (Fig. 2b) of the eCG alpha subunit illustrates the location of the three alpha subunit loops: Loop 1, Loop 2 and Loop 3. This structure also illustrates the location of the region with sequences “CCFSRA” that probably may act as a binding antagonist. The eCG alpha contains the two conserved motifs with “CKGCCFSRAYPTP” & “NHTQCYCSTCYHHK”



**Fig. 2** Molecular docking of Ganirelix and eCG alpha. **a** Two dimensional structure of Ganirelix. The hydroxyl groups (–OH) oxygen and carbonyl oxygen are shown in red colour and the –Cl group in green colour. **b** The 3D Homology Model of the eCG  $\alpha$ -subunit. **a** This figure illustrates the location of the two conserved motifs with “CKGCCFSRAYPTP” & “NHTQCYCSTCYHHK”

sequences respectively. **b** Surface dimension of the eCG  $\alpha$ -subunit. For calculating the surface, Connolly method was employed. **c** The docking results of eCG alpha-ganirelix system as visualized by PyMOL. The eCG alpha protein is represented as surface diagram. Ganirelix is shown as stick around the binding pocket

sequences respectively. 3D Mol tool was used to calculate and display molecular surface dimensions via the Connolly method.

The binding site analysis of eCG alpha was done by MetaPocket 2.0. The MetaPocket uses the combined analyses by eight methods viz. LIGSITE, PASS, Q-SiteFinder, SURFNET, Fpocket, GHECOM, ConCavity and POCASA and have revealed the predicted binding sites in the eCG alpha (Additional file 1: Figure S9).

### Virtual screening and molecular docking

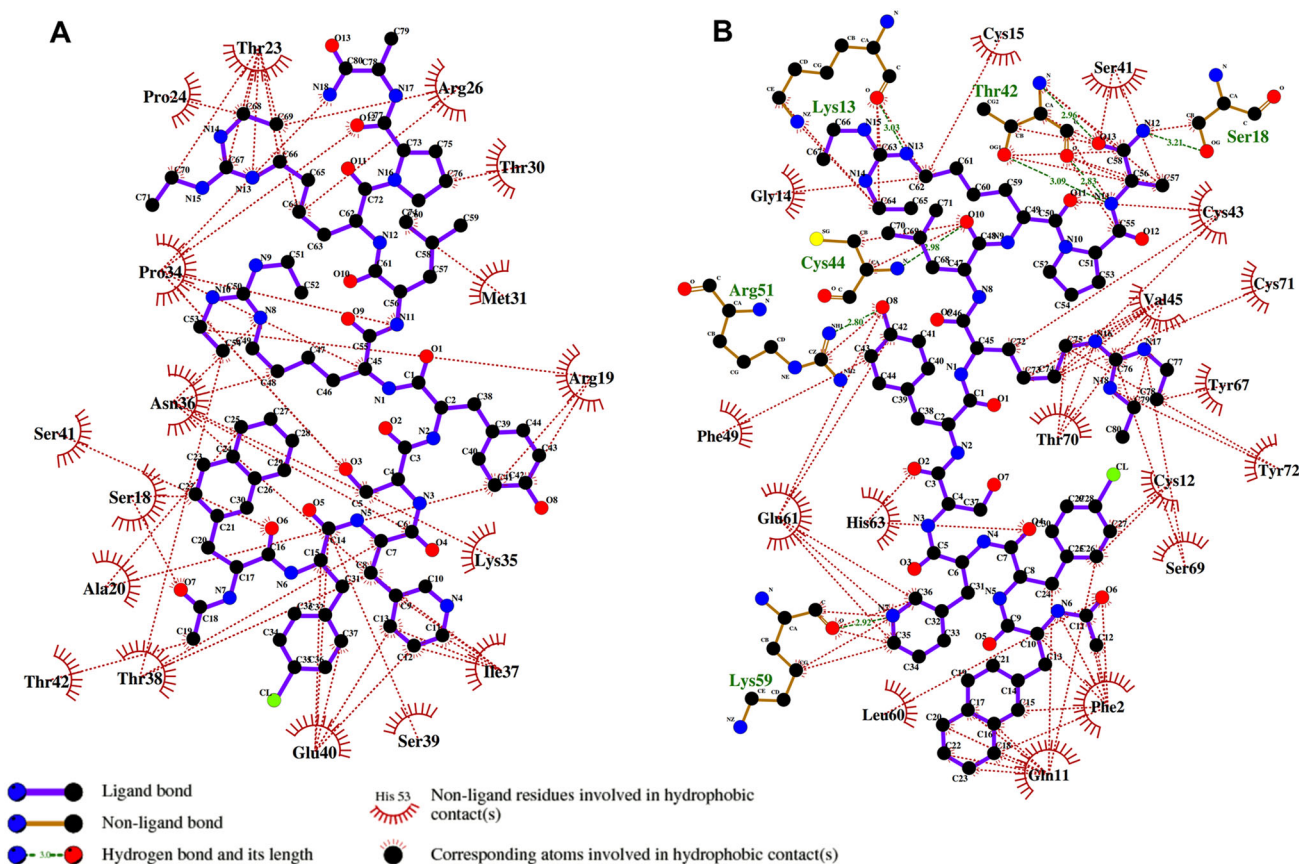
To investigate the molecular binding interactions of ganirelix (Fig. 2a) and eCG alpha molecules and to elucidate the possible molecular mechanism, PyRx Autodock 4 docking analysis was applied. About nine different poses or binding modes were observed. The molecular binding affinity of these structures ranged from  $-5.0$  to  $-4.5$  kcal/mol. The average molecular binding affinity (docking energy) scores of the ganirelix on the eCG alpha were  $-4.733$  kcal/mol (SEM =  $\pm 3.99$ ). When all these docked structures were analysed using PyMOL (Fig. 2c and Additional file 1: Figure S10), Dassault Systèmes BIOVIA Discovery Studio Visualizer (v16.1.0.15350) and LIGPLOT<sup>+</sup> (Wallace et al. 1995), it was observed that

some residues located in the different subsites of eCG alpha are involved in making hydrophobic stacking interactions.

While, in some of the poses, hydrogen bonding interactions were also evident. In the binding mode with the lowest binding affinity of  $-5.0$  kcal/mol, we could observe hydrophobic interactions with several amino acid residues such as Ser18, Arg19, Ala20, Thr23, Pro24, Arg26, Thr30, Met31, Pro34, Lys35, Asn36, Ile37, Thr38, Ser39, Glu40, Ser41, Thr42 (Fig. 3a). It seems that these 17 amino acid residues stabilized the eCG alpha-ganirelix complex by exerting the hydrophobic interactions only. The amino acid residues Arg19, Thr23, Pro34, Asn36, Ile37, Thr38 and Glu40 presented maximum hydrophobic contacts with the ganirelix molecule and therefore, could be the crucial interacting residues.

However, role of H-bonds could also be evident. In one of the binding conformation mode (binding affinity =  $-4.7$  kcal/mol), ganirelix was found to be interacting with several amino acid residues (conventional hydrogen bond: Lys13, Ser18, Thr42, Cys44, Arg51, Lys59; hydrophobic interactions: Phe2, Gln11, Cys12, Gly14, Cys15, Ser41, Cys43, Val45, Phe49, Leu60, Glu61, His63, Tyr67, Ser69, Thr70 and Cys71). The shortest H-bond length between the ganirelix and the eCG alpha was  $2.80\text{\AA}$  and the





**Fig. 3** Ligplot diagrams (a, b) elucidating the protein–ligand interactions during molecular docking of Ganirelix and eCG alpha. Hydrophobic interactions are represented by red dotted lines radiating from the amino acids towards the ganirelix contact atoms. The ganirelix ligand structure is represented in purple. Cl, O, C and N

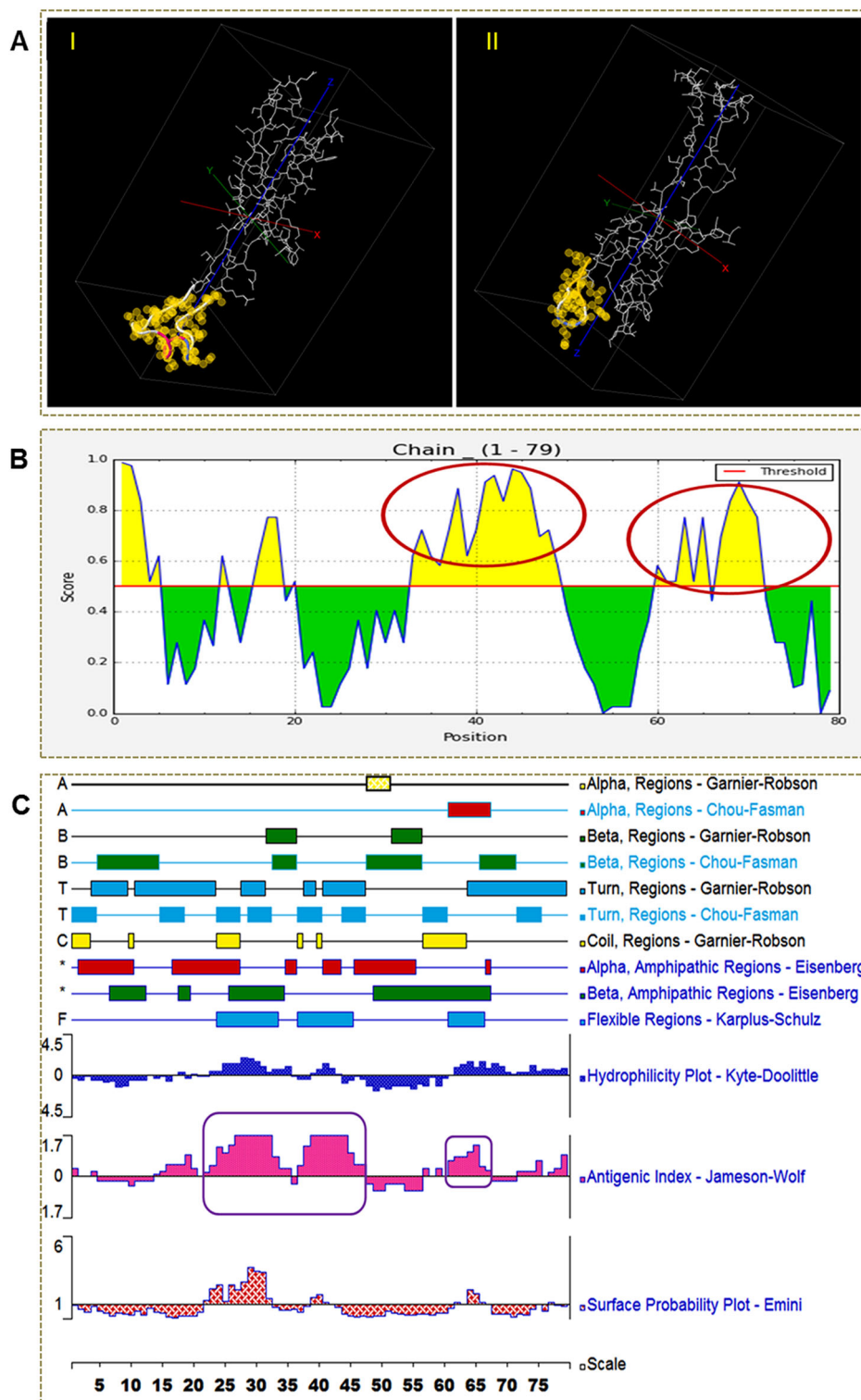
interacting amino acid was Arg51. The longest bond length between ganirelix and the eCG alpha was 3.21 Å with Ser18 (Fig. 3B). The residue Lys13 formed an H bond (3.03 Å) through its carbonyl oxygen with the N13 atom of ganirelix. The residue Thr42 of eCG alpha formed three hydrogen bonds (3.09, 2.83, and 2.96 Å) with the ganirelix. Altogether these 22 residues exerted eight hydrogen bonding and more than 70 hydrophobic interactions with ganirelix and stabilized the eCG alpha-ganirelix complex.

### **In silico analysis of eCG alpha for the identification of potential B cell epitope**

The eCG alpha protein was subjected to bioinformatics analyses including hydrophilicity/hydrophobicity, identification of potential B cell epitope or antigenicity/antigenicity propensity score. Using ElliPro (Ponomarenko et al. 2008), 4 structural templates with PDB IDs viz. 1HCN, 1HRP, 4AY9 and 4MQW were used for the homology modelling

atoms are denoted by green, red, black and blue spheres respectively. The H-bond interactions between different interacting atoms are represented by green dotted lines along with bond lengths. The possible amino acids for H-bond interactions are shown in green

and identification of epitopes. A 19 amino acid peptide sequence (KTMLVPKNITSESTCCVAK) and a thirteen amino acid peptide sequence (GNIKLENHTQCYC) seems promising linear epitopes (Fig. 4a). Four clusters were predicted to form conformational B cell epitope or discontinuous epitope. Out of these four, most promising residues clusters are shown in the Additional file 2: Table S2. 2D Score Chart or protrusion index chart for the eCG alpha protein sequence is exhibited in the Fig. 4b. The SCRATCH Protein Predictor (accessible online through <http://scratch.proteomics.ics.uci.edu/>) (Cheng et al. 2005) was also used for ascertaining the antigenicity propensity score. We also analyzed the eCG alpha protein sequence using various methods for helical wheel diagram presenting the distribution of amino acids spatially in a 2D graph (Additional file 1: Figure S11), secondary structure, hydropathy, antigenicity, amphiphilicity, surface probability and flexibility (Fig. 4c) by using the Protean Software (for Protein Structure Analysis and Prediction) of the DNASTAR Lasergene software suite version 7.1.



**Fig. 4** *In silico* analysis of eCG alpha for the identification of potential B cell epitope. **a** Predicted epitopic regions in the eCG alpha. **(I)** Most potent antigenic linear B cell epitope sequence, KTMLVPKNITSESTCCVAK is exhibited here. This epitope 3D structure is shown here as trace and highlighted in yellow. **(II)** Predicted discontinuous epitope region with residue cluster (L63, E64, N65, H66, T67, Q68, C69, Y70, C71) is highlighted as trace and yellow colour. **b** 2D Score Chart or protrusion index chart for the eCG

alpha protein sequence. The prominent epitope regions with amino acid sequences “KTMLVPKNITSESTCCVAK” and “GNIKLENHTQCYC” respectively are shown as yellow shaded regions and circled in yellow. **c** eCG alpha protein structure analysis and prediction by Protean software. The analyses were made for the secondary structure, hydrophobicity, antigenicity, amphiphilicity, surface probability and flexibility

## Discussion

To strengthen the research primarily targeting the gonadotropin therapy for fertility augmentation and in assisted reproductive technologies, we hypothesize that the alpha subunit of equine chorionic gonadotropin hormone plays an imperative role. Many workers across the globe reported the studies on alpha-subunit of pituitary glycoprotein hormones. Chopineau et al. (2004), reported the identification of amino-acids in the alpha-subunit first and third loops, that are crucial for the hetero-specific follicle-stimulating hormone activity of equid luteinizing hormone/gonadotropin. It was precisely shown that which amino-acids of the alpha-subunit are implicated in the interaction with gonadotropin receptors. Some earlier studies based on photo affinity labelling (Hong et al. 1999; Sohn et al. 2002) immunological studies (Moyle et al. 1982) and mutational studies (Yoo et al. 1991; Zeng et al. 1995; Arnold et al. 1998) demonstrated that the alpha subunit also interacts with the receptor. However, we report for the first time, the predicted three dimensional structures of the mature alpha subunit of eCG and its molecular docking with ganirelix which may play role in elaborating the responsibility played by alpha subunit of eCG in the process of reproduction and maintenance of pregnancy in equines as well as other species.

In many of the preliminary studies while evaluating the application of the GnRH antagonists to clinical practice, GnRH antagonist's capability to block the LH surge in a reproductive cycle was also witnessed. The obstruction of the pituitary GnRH receptors caused the hindrance of the LH surge as well as the interruption of FSH production and further arrest of follicular development (Ditkoff et al. 1991; Kettel et al. 1991). It also became evident that on application of GnRH antagonists to natural cycle IVF, addition of exogenous gonadotropins was needed to maintain follicular growth. It is worthwhile to mention that GnRH and eCG have been used in timed artificial insemination (TAI) protocols to improve the pregnancy rate (PR) in cattle (Gaievski et al. 2015). Therefore, it was interesting to investigate the role of GnRH antagonist such as ganirelix here. Ganirelix/Ganirelix acetate (or diacetate) (Orgalutran® or Antagon™) is a synthetic decapeptide and derived from GnRH with amino acid substitutions at positions 1, 2, 3, 6, 8, and 10. Deemed as an injectable competitive gonadotropin-releasing hormone antagonist (GnRH antagonist), it is primarily used in assisted reproduction to control ovulation. Ganirelix acetate (Antagon™) was launched in Europe in 2001 by Organon for the treatment of fertility disorder. In 1999 Organon received US FDA approval for the drug for inhibiting premature LH surges in women undergoing multifollicular ovarian stimulation (Papanikolaou et al.

2005). Ganirelix has been used as a pharmacological tool to prevent ovulation, until it is prompted by injecting human chorionic gonadotrophin (hCG) with which it is known to interact. It has been the basis of treatment regimen for IVF patients who are undergoing ovarian stimulation and might be restricted to days spanning LH premature surge (Borm and Mannaerts 2000). All the earlier works have reiterated that use of GnRH antagonists such as ganirelix suppress the gonadotropins by blocking the GnRH receptor.

Based in the in silico computational methods, including molecular docking, we pursued to gain an insight into the structural mechanism of eCG recognition by ganirelix. Employing the present molecular docking study, it was evident that ganirelix can also interact with eCG alpha (Figs. 2, 3) and the interacting numbers of amino acids and conventional hydrogen bonds and hydrophobic interactions might be critical factors for regulating target protein activity. Docking with ganirelix was predictive of a possible preference for the eCG alpha subunit which is common to gonadotropins. We observed nine different poses or binding modes for the docked structure. The molecular binding affinity of these structures ranged from  $-5.0$  to  $-4.5$  kcal/mol. The average molecular binding affinity (docking energy) scores of the ganirelix on the eCG alpha were  $-4.733$  kcal/mol (SEM =  $\pm 3.99$ ). These were indicative of the fact that the binding free energies and affinity can fluctuate in silico, signifying that ganirelix may access or fit the eCG alpha through several modes. The docking study has revealed the stabilization of eCG-ganirelix complex through several hydrophobic interactions and hydrogen bonds. These data also suggest that computer aided drug design process using PyRx, and Discovery Studio tools is highly reliable and can be a good example for in identifying the interaction between eCG alpha and ganirelix. Previous studies have shown that ganirelix suppress gonadotropins by blocking GnRH receptor and mentioned earlier in the manuscript. Also, the alpha subunit is common to all the gonadotropins. Complementary to these facts, our data of molecular docking study indicate for an alternate route of gonadotropin (in our case, that of eCG alpha) sequestration by interactions of alpha subunit with the ganirelix. This may elucidate an alternative possibility for ganirelix mediated gonadotropin suppression directly by binding apart from GnRH receptor blocking. This hypothesis may be subject to further studies. In our study, we also report the putative antigenic sites for alpha-eCG to act as immunogen or as an analog to generate antibodies required in various studies related to reproduction in farm animals. Putative chromosomal localization of the eCG alpha was also determined through in silico ways.



## Conclusions

In summary, the equid alpha subunit is an absolute requirement for expression of both the LH and FSH activities of equine chorionic gonadotropin hormone (Chopineau et al. 1997) and through various in silico analysis of eCG alpha sequence, we attempted for gaining structural insights as well as providing an alternate route of gonadotropin suppression. It has to be noted here that our results are largely based on the in silico analysis and data mining, and have not been proven experimentally. However, they still provide an alternate explanation for the gonadotropin suppression or sequestration through gonadotropin alpha subunit and ganirelix interactions. Future studies will be helpful in elucidation of the eCG alpha and ganirelix interactions.

**Acknowledgements** The authors are very grateful to Director ICAR-NRCE, Hisar for providing the necessary facilities for this study. They also acknowledge suggestions of Gupta, A. K.

**Authors' contributions** AB and VN conceived and designed the experiments; YP and JS collected the equine pituitary samples; AB, PS, SK and VN performed the experiments; AB and VN analyzed the data; AB, VN and SK contributed reagents/materials/analysis tools; AB and VN wrote the paper. All authors read and approved the final manuscript.

## Compliance with ethical standards

**Statement of ethics approval** Experimental protocol was reviewed and approved by Institute Animal Ethics Committee (IAEC) of ICAR-National Research Centre on Equines, Hisar. All the procedures for sample collection were performed according to the approval by the Institute Animal Ethics Committee (IAEC) of ICAR-National Research Centre on Equines vide NRCE/CPCSEA/2012-13/dated the 18th July 2012.

**Consent to participate** We give our consent to participate under the 'Ethics, consent and permissions' heading.

**Availability of data and materials** The data generated during this study and supporting this study are included in this article and also in additional files within the Additional file 1: Figure S1–S11 and Additional file 2: tables S1–S2. The dataset analyzed during the current study is available from the corresponding author on a reasonable request.

**Conflict of interest** None of the authors have any competing interests in the manuscript.

**Funding** This work was financially supported by ICAR-National Research Centre on Equines, Hisar (URL: [www.nrce.gov.in](http://www.nrce.gov.in)) under the institute project (code: IXX02769).

## References

Abel MH, Widen A, Wang X, Huhtaniemi I, Pakarinen P, Kumar TR, Christian HC (2014) Pituitary gonadotrophic hormone synthesis,

secretion, subunit gene expression and cell structure in normal and follicle-stimulating hormone  $\beta$  knockout, follicle-stimulating hormone receptor knockout, luteinizing hormone receptor knockout, hypogonadal and ovariectomised female mice. *J Neuroendocrinol* 26(11):785–795

Arnold CJ, Liu C, Lindau-Shepard B, Losavio ML, Patrascu MT, Dias JA (1998) The human follitropin alpha-subunit C terminus collaborates with a beta-subunit cystine noose and an alpha-subunit loop to assemble a receptor-binding domain competent for signal transduction. *Biochemistry* 37(7):1762–1768

Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res* 37:W510–W514

Borm G, Mannaerts B (2000) Treatment with the gonadotrophin-releasing hormone antagonist ganirelix in women undergoing ovarian stimulation with recombinant follicle stimulating hormone is effective, safe and convenient: results of a controlled, randomized, multicentre trial. The European Orgalutran Study Group. *Hum Reprod* 15(7):1490–1498

Bousfield GR, Butnev VY, Gotschall RR, Baker VL, Moore WT (1996) Structural features of mammalian gonadotropins. *Mol Cell Endocrinol* 125(1–2):3–19

Burland TG (2000) DNASTAR's Lasergene sequence analysis software. *Methods Mol Biol* 132:71–91

Cahoreau C, Klett D, Combarnous Y (2015) Structure-function relationships of glycoprotein hormones and their subunits' ancestors. *Front Endocrinol (Lausanne)* 6:26

Caspermeyer J (2016) MEGA evolutionary software re-engineered to handle today's big data demands. *Mol Biol Evol* 33(7):1887

Chaube R, Joy KP, Acharjee A (2015) Catfish gonadotrophins: cellular origin, structural properties and physiology. *J Neuroendocrinol* 27(6):536–543

Cheng J, Randall AZ, Sweredoski MJ, Baldi P (2005) SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res* 33:W72–W76

Chi ML, Ni M, Li JF, He F, Qian K, Zhang P, Chai SH, Wen HS (2015) Molecular cloning and characterization of gonadotropin subunits (GTH $\alpha$ , FSH $\beta$  and LH $\beta$ ) and their regulation by hCG and GnRH $\alpha$  in Japanese sea bass (*Lateolabrax japonicus*) in vivo. *Fish Physiol Biochem* 41(3):587–601

Chopineau M, Maurel MC, Combarnous Y, Durand P (1993) Topography of equine chorionic gonadotropin epitopes relative to the luteinizing hormone and follicle-stimulating hormone receptor interaction sites. *Mol Cell Endocrinol* 92(2):229–239

Chopineau M, Martinat N, Marichatou H, Troispoux C, Auge-Gouillou C, Stewar F, Combarnous Y, Guillou F (1997) Evidence that the alpha-subunit influences the specificity of receptor binding of the equine gonadotrophins. *J Endocrinol* 155:241–245

Chopineau M, Martinat N, Gibrat JF, Galet C, Lecompte F, Foulon-Gauze F, Pourchet C, Guillou F, Combarnous Y (2004) Identification of amino-acids in the alpha-subunit first and third loops that are crucial for the heterospecific follicle-stimulating hormone activity of equid luteinizing hormone/choriogonadotropin. *Eur J Endocrinol* 150(6):877–884

Christakos S, Bahl OP (1979) Pregnant mare serum gonadotropin. Purification and physicochemical, biological, and immunological characterization. *J Biol Chem* 254(10):4253–4261

Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, Heinz TN, Kastenholz MA, Kräutler V, Oostenbrink C, Peter C, Trzesniak D, van Gunsteren WF (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem* 26(16):1719–1751

Combarnous Y (1992) Molecular basis of the specificity of binding of glycoprotein hormones to their receptors. *Endocr Rev* 13(4):670–691



- Combet C, Jambon M, Deléage G, Geourjon C (2002) Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics* 18(1):213–214
- Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. In: Hempel JE, Williams CH, Hong CC (eds). *Chemical biology: methods and protocols* (Volume 1263 of the series *Methods in Molecular Biology*). New York, Springer, pp 243–250. doi:10.1007/978-1-4939-2269-7\_19
- Dassault Systèmes BIOVIA (2016) Discovery Studio Visualizer, version 16.1.0.15350, Dassault Systèmes, San Diego
- Ditkoff EC, Cassidenti DL, Paulson RJ, Sauer MV, Paul WL, Rivier J et al (1991) The gonadotropin-releasing hormone antagonist (Nal-Glu) acutely blocks the luteinizing hormone surge but allows for resumption of folliculogenesis in normal women. *Am J Obstet Gynecol* 165(6 Pt 1):1811–1817
- Fiddes JC, Goodman HM (1979) Isolation, cloning and sequence analysis of the cDNA for the alpha-subunit of human chorionic gonadotropin. *Nature* 281(5730):351–356
- Fiddes JC, Goodman HM (1981) The gene encoding the common alpha subunit of the four human glycoprotein hormones. *J Mol Appl Genet* 1(1):3–18
- Gaievski FR, Lamb GC, Weiss RR, Bertol MAF, Segui MS, Abreu ACMR (2015) Gonadotropin releasing hormone (GnRH) and equine chorionic gonadotropin (eCG) improve the pregnancy rate on protocols for timed-artificial insemination in beef cattle. *Veterinária E Zootecnia* 22(3):471–480
- Galet C, Chopineau M, Martinat N, Combarnous Y, Guillou F (2000) Expression of an in vitro biologically active equine LH/CG without C-terminal peptide (CTP) and/or beta26-110 disulphide bridge. *J Endocrinol* 167(1):117–124
- Gregory SJ, Brooks J, McNeilly AS, Ingleton PM, Tortonese DJ (2000) Gonadotroph-lactotroph associations and expression of prolactin receptors in the equine pituitary gland throughout the seasonal reproductive cycle. *J Reprod Fertil* 119(2):223–231
- Hokke CH, Roosenboom MJ, Thomas-Oates JE, Kamerling JP, Vliegenthart JF (1994) Structure determination of the disialylated poly-(N-acetylactosamine)-containing O-linked carbohydrate chains of equine chorionic gonadotropin. *Glycoconj J* 11(1):35–41
- Hong S, Ji I, Ji TH (1999) The alpha-subunit of human choriogonadotropin interacts with the exodomain of the luteinizing hormone/choriogonadotropin receptor. *Endocrinol* 140:2486–2493
- Huang B (2009) MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS* 13(4):325–330
- Kalra SP, Horvath T, Naftolin F, Xu B, Pu S, Kalra PS (1997) The interactive language of the hypothalamus for the gonadotropin releasing hormone (GNRH) system. *J Neuroendocrinol* 9(8):569–576
- Kettel LM, Roseff SJ, Chiu TC, Bangah ML, Vale W, Rivier J et al (1991) Follicular arrest during the midfollicular phase of the menstrual cycle: a gonadotropin-releasing hormone antagonist imposed follicular-follicular transition. *J Clin Endocrinol Metab* 73(3):644–649
- Kumar S, Tamura K, Nei M (1994) MEGA: molecular evolutionary genetics analysis software for microcomputers. *Comput Appl Biosci* 10(2):189–191
- Lambalk CB, Leader A, Olivennes F, Fluker MR, Andersen AN, Ingerslev J, Khalaf Y, Avril C, Belaisch-Allart J, Roulier R, Mannaerts B (2006) Treatment with the GnRH antagonist ganirelix prevents premature LH rises and luteinization in stimulated intrauterine insemination: results of a double-blind, placebo-controlled, multicentre trial. *Hum Reprod* 21(3):632–639
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK, a program to check the stereochemical quality of protein structures. *J Appl Crystal* 26(2):283–291
- Maston GA, Ruvolo M (2002) Chorionic gonadotropin has a recent origin within primates and an evolutionary history of selection. *Mol Biol Evol* 19(3):320–335
- McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16(4):404–405
- McMaster GK, Carmichael GG (1977) Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc Natl Acad Sci USA* 74(11):4835–4838
- Melo F, Devos D, Depiereux E, Feytmans E (1997) ANOLEA: a www server to assess protein structures. *Proc Int Conf Intell Syst Mol Biol* 5:187–190
- Min KS, Shinozaki M, Miyazawa K, Nishimura R, Sasaki N, Shiota K, Ogawa T (1994) Nucleotide sequence of eCG  $\alpha$ -subunit cDNA and its expression in the equine placenta. *J Reprod Dev* 40(4):301–305
- Moyle WR, Ehrlich PH, Canfield RE (1982) Use of monoclonal antibodies to subunits of human chorionic gonadotropin to examine the orientation of the hormone in its complex with receptor. *Proc Natl Acad Sci USA* 79(7):2245–2249
- O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open babel: an open chemical toolbox. *J Cheminform* 3:33
- Papanikolaou EG, Kolibianakis E, Devroey P (2005) Emerging drugs in assisted reproduction. *Exp Opin Emerg Drugs* 10(2):425–440
- Pierce JG, Parsons TF (1981) Glycoprotein hormones: structure and function. *Annu Rev Biochem* 50:465–495
- Ponomarenko J, Bui H-H, Li W, Füsseder N, Bourne PE, Sette A, Peters B (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform* 9:514
- Schrodinger L (2010) The PyMOL molecular graphics system, version1.3r1
- Sherman GB, Wolfe MW, Farmerie TA, Clay CM, Threadgill DS, Sharp DC, Nilson JH (1992) A single gene encodes the beta-subunits of equine luteinizing hormone and chorionic gonadotropin. *Mol Endocrinol* 6(6):951–959
- Sohn J, Ryu K, Sievert G, Jeoung M, Ji I, Ji TH (2002) Follicle-stimulating hormone interacts with exolooop 3 of the receptor. *J Biol Chem* 277(51):50165–50175
- Ulloa-Aguirre A, Maldonado A, Damián-Matsumura P, Timossi C (2001) Endocrine regulation of gonadotropin glycosylation. *Arch Med Res* 32(6):520–532
- Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8(2):127–134
- Weltzien FA, Kobayashi T, Andersson E, Norberg B, Andersen Ø (2003) Molecular characterization and expression of FSHbeta, LHbeta, and common alpha-subunit in male Atlantic halibut (*Hippoglossus hippoglossus*). *Gen Comp Endocrinol* 131(2):87–96
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
- Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* 31(13):3316–3319
- Yoo J, Ji I, Ji TH (1991) Conversion of lysine 91 to methionine or glutamic acid in human choriogonadotropin alpha results in the loss of cAMP inducibility. *J Biol Chem* 266(27):17741–17743
- Zeng H, Ji I, Ji TH (1995) Lys91 and His90 of the alpha-subunit are crucial for receptor binding and hormone action of follicle-stimulating hormone (FSH) and play hormone-specific roles in FSH and human chorionic gonadotropin. *Endocrinology* 136(7):2948–2953