IJBPAS, October, 2012, 1(9): 1258-1269

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

# STRUCTURE PREDICTION AND ANALYSIS OF G-PROTEIN COUPLED RECEPTORS

**BASTIKAR V[1], ANJUM AA [2], GUPTA PP[2], GUPTE A[1], SHAH R P[2], DESAI D M [3] AND GANGAWANE AK[3*]**

**1:** Department of Biotechnology & Pharmacy, University of Rome "Tor Vergata", Rome, Italy

**2:** Department of Biotechnology and Bioinformatics, Dr Padmashree D Y Patil University, Sector 15, Plot 50 CBD Belapur, Navi Mumbai, 400614, Maharashtra, India

**3:** Department of Biotechnology, MGM-CET & MGM Institute of Health Sciences, Kamothe, Navi Mumbai

***Corresponding Author: Email: ajganga@yahoo.com**

## ABSTRACT

A relevant and accurate description of three-dimensional (3D) protein structures can be achieved by characterizing recurrent local structures. Computational prediction techniques provide an attractive and alternative method towards a better accurate result directly from the amino-acids sequence data to a 3-dimensional structure. As GPCRs are targets for 50% of all the existing medications. GPCR's are thought to be in equilibrium between their active and inactive state and are affected by the binding of ligands. A number of algorithms and methodologies have been implemented to model the secondary and tertiary structure of the GPCR. The targeted sequence is modeled into a secondary and tertiary putative structure, implementing multiple protein structure prediction tools and algorithms. The predicted 3-dimensional structure has been optimized for a better conformation and stability.

**Keywords: G-Protein, Computational, Secondary and Tertiary Structure**

## INTRODUCTION

To understand basic biological processes such as cell signalling and protein interaction required a, detailed knowledge of the three-dimensional structures of the active participants is necessary. A large number of protein structures have been experimentally determined using primarily the X ray crystallography and NMR

spectroscopy technique. Currently more than 50,000 entries in the Protein Data Bank (PDB) **[1]**, which archives experimentally, determined structures of proteins and protein complexes, as well as nucleic acids and other biological macromolecules. However, the list of protein sequences in the PDB such that no two sequences are more than 20% identical to each other, contains only about 4000 sequences **[2]**. Many of these proteins can be grouped further into about 1000 folds in 2000 super families **[3]**. Since it was first recognized that proteins can share similar structures **[4]**, computational methods have been developed to build models of proteins of unknown structure based on related proteins of known structure **[5]**. Most such modelling efforts, referred to as homology modelling or comparative modelling, follow a basic protocol: 1) for a target sequence of unknown structure, identify a template structure with sequence related to the target and align the target sequence to the template sequence and structure; 2) for core secondary structures and all well-conserved **[6]**.The Combined Oral Contraceptive Pill (COCP), often referred to as the birth-control pill, or simply "the pill", is a birth control method. Combined oral contraceptive pills were developed to prevent ovulation by suppressing the release of gonadotropins. Combined hormonal contraceptives, including COCPs, inhibit follicular development and prevent ovulation as their primary mechanism of action **[7-11]**.

Melatonin, also known chemically as N-acetyl-5-methoxytryptamine, is a naturally occurring hormone found in animals and in some other living organisms, including algae **[12]**. Its Circulating levels vary in a daily cycle, and melatonin is important in the regulation of the circadian rhythms of several biological functions **[13]**. The melatonin signal forms part of the system that regulates the sleep-wake cycle by chemically causing drowsiness and lowering the body temperature. Melatonin also lowers FSH levels. It is believed that these hormonal changes could in some women impair fertility. In birds and mammals it was found that melatonin switches on a recently discovered hormone called gonadotropin-inhibitory hormone(GnIH),which has been found to have the opposite effect to the key hormone priming the body for sex-gonadotropin-releasing hormone(GnRH) **(Figure 1)**.Switching off GnRH causes gonads(testes and ovary) to shrink and be inactive. A melatonin receptor is a G protein-coupled receptor (GPCR) which binds melatonin **[14]**. Three types of melatonin receptor have been cloned. The $MT_1$ (or $Mel_{1A}$ or MTNR1A) and $MT_2$ (or $Mel_{1B}$ or MTNR1B) receptor subtypes are

present in humans and other mammals, **[15]** while an additional melatonin receptor subtype $MT_3$ (or $Mel_{1C}$ or MTNR1C) has been identified in amphibia and birds **[16]**.

Computational modeling of G-Protein Coupled Receptors (GPCRs) has recently become an interesting area of research, as GPCRs are targets for 50% of all existing medications **[17]**. The G-Protein Coupled Receptors have a "mean pair wise amino acid identity" of only 17%, making it challenging to both classify and create homology models **[18, 19]**.

The melatonin receptors are closely related to bovine Rhodopsin, the first protein crystal structure deposited in the Protein Data Bank. Experimental methods (e.g., X-ray crystallography and NMR spectroscopy) often fail to determine three-dimensional (3D) structure of GPCRs. The only exception is the bovine visual rhodopsin: its 3D X-ray structure has been determined with atomic resolution. **Figure 2** represents the TM folding of rhodopsin.

It is commonly accepted now that all GPCRs share similar fold of TM domain. This reveals an opportunity to build their 3D models by means of molecular modelling **[20]**.

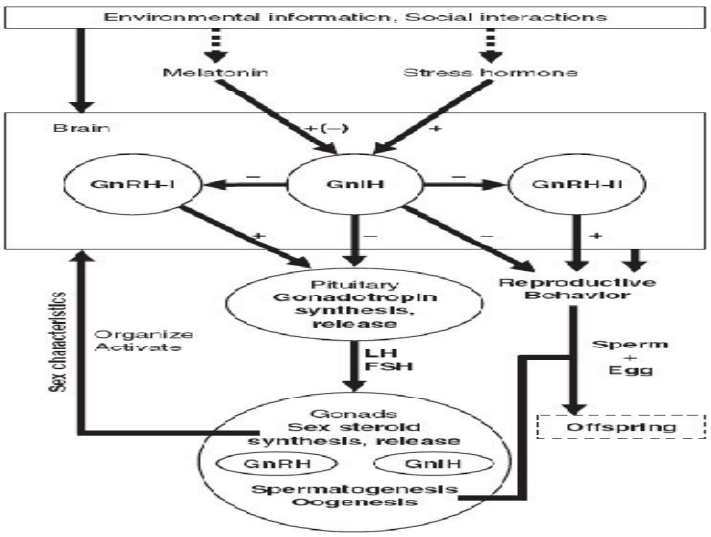


**Figure 1: Control Mechanisms of Reproductive Physiology and Behavior by GnIH (RFRP) and GnRH in Birds and Mammals**

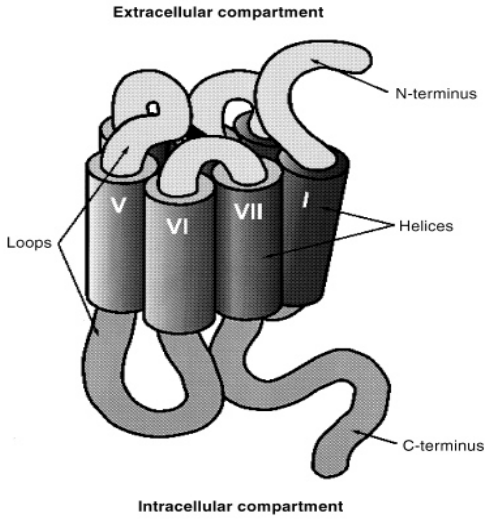


**Figure 2: Schematic representation of G - protein coupled receptor folding**

## MATERIALS AND METHODS

The protein sequence of Melatonin receptor 1a and melatonin receptor 1b was retrieved from NCBI, the sequences are NP_005949.1| melatonin receptor 1A [Homo sapiens] sequence length 350 **[21]** and NP_005950.1| melatonin receptor 1B [Homo sapiens] 362 amino acid long respectively **[22]**.

Accessing Secondary and Tertiary Structure Prediction Tools:

### GOR

The GOR method is based on information theory and was developed by J.Garnier, D.Osguthorpe and B.Robson. Sequence submitted to GOR IV, uses all possible pair frequencies within a window of 17 amino acid residues and is reported by J.Garnier, J.F.Gibrat and B.Robson in Methods in Enzymology, vol 266, p 540-553 (1996). After cross validation on a database of 267 proteins, the version IV of GOR has a mean accuracy of 64.4% for a three state prediction. The predicted secondary structure is one of the highest probabilities compatible with a predicted helix segment of at least four residues and a predicted extended segment of at least two residues.

### ESyPred3D

ESyPred3D is a new automated homology modelling program. This method gets benefit of the increased alignment performances of a new alignment strategy using neural networks. Alignments are obtained by combining, weighting and screening the results of several multiple alignment programs. The final three dimensional structures are building using the modelling package MODELLER.

### 3DJigSaw

3D-JigSaw is an automated system to build three-dimensional models for proteins based on homologues of known structure. There are two ways in which 3DJigSaw builds the

models namely Automatic mode and Interactive mode.

In automatic mode the program looks for homologous templates in the sequence databases (PFAM, PDB) and splits the query sequence into domains. If good templates are found, the best covered domain is then modelled. In the interactive mode the program looks for homologous templates in the sequence databases (PFAM, PDB) and splits the query sequence into domains. Identified templates are ranked according to the coverage of the query, their sequence identity and their crystallographic resolution. Information from each template is easily accessed, including its alignment to the query sequence.

The method of homology modelling is based on the observation that protein tertiary structure is better conserved than amino acid sequence. Thus, even proteins that have diverged appreciably in sequence but still share detectable similarity will also share common structural properties, particularly the overall fold. Because it is difficult and time-consuming to obtain experimental structures from methods such as X-ray crystallography and protein NMR for every protein of interest, homology modelling can provide useful structural models for generating hypotheses about a protein's function and directing further experimental work.

The homology modelling procedure can be broken down into four sequential steps:

1. Finding homologous PDB files.
2. Creation of the alignment, using single or multiple alignments; analysis of alignments; gap deletions and additions; secondary structure weighting.
3. Structure calculation and,
4. Model refinement.

## RESULTS AND DISCUSSION

### 1. NCBI

The protein sequence of Melatonin receptor 1a and melatonin receptor 1b was retrieved from NCBI; the sequences are 350 and 362 amino acid long respectively.

>gi|5174593|ref|NP_005949.1| melatonin receptor 1A [Homo sapiens]

MQGNGSALPNASQPVLRGDGARPSWLASALACVLIFTIVVDILGNLLVILSVYRNKK
LRNAGNIFVVSLAVADLVVAIYPYPLVLMSIFNNGWNLGYLHCQVSGFLMGLSVIGS
IFNITGIAINRYCYICHSLKYDKLYSSKNSLCYVLLIWLLTLAAVLPNLRAGTLQYDPR
IYSCTFAQSVSSAYTIAVVVFHFLVPMIIVIFCYLRIWILVLQVRQRVKPDRKPKLKPQ

DFRNFVTMFVVFVLFAICWAPLNFIGLAVASDPASMVPRIPEWLFVASYYMAYFNSC
LNAIIYGLLNQNFRKEYRRIIVSLCTARVFFVDSSNDVADRVKWKPSPLMTNNNVVK
VDSV

>gi|5174595|ref|NP_005950.1| melatonin receptor 1B [Homo sapiens]

MSENGSFANCCEAGGWAVRPGWSGAGSARPSRTPRPPWVAPALSAVLIVTTAVDV
VGNLLVILSVLRNRKLRNAGNLFLVSLALADLVVAFYPYPLILVAIFYDGWALGEEH
CKASAFVMGLSVIGSVFNITAIAINRYCYICHSMAYHRIYRRWHTPLHICLIWLLTVV
ALLPNFFVGSLEYDPRIYSCTFIQTASTQYTAAVVVIHFLLPIAVVSFCYLRIWVLVLQ
ARRKAKPESRLCLKPSDLRSFLTMFVVFVIFAICWAPLNCIGLAVAINPQEMAPQIPEG
LFVTSYLLAYFNSCLNAIVYGLLNQNFRREYKRILLALWNPRHCIQDASKGSHAEGL
QSPAPPIIGVQHQADAL

The sequence of the template structure, crystal structure of bovine rhodopsin has two chains **(Table 1)**:

**Table 1: Sequence of the Template structure, Crystal Structure of Bovine Rhodopsin with two Chains**

| Query sequence | NP_005949.1| melatonin receptor 1A [Homo sapiens] | NP_005950.1| melatonin receptor 1B [Homo sapiens] |
|---|---|---|
| Pdb template | 1F88:A | 1F88:B |
| Score | 92.0 bits | 86.3 bits |
| Identities | 79/339 (23%). | 73/319 (22%). |
| Positives | 148/339 (43%). | 136/319 (42%). |
| Gaps | 40/339 (11%). | 17/319 (5%). |
| Expect value | 3e-19 | 2e-17. |
| Length | 348 | 348 |

>1F88:B|PDBID|CHAIN|SEQUENCE

MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFL
TLYVTQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGF
FATLGGEIALWSLVVLAIERYVVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPP
LVGWSRYIPEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTV
KEAAAQQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFM

TIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQV
APA

>1F88:A|PDBID|CHAIN|SEQUENCE

MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFL
TLYVTQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGF
FATLGGEIALWSLVVLAIERYVVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPP
LVGWSRYIPEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTV
KEAAAQQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFM
TIPAFFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQV
APA.

**2.GOR**

The secondary structure prediction using GOR shows 63 residues forming alpha helix, 127 residues forming extended beta strand and the remaining 160 residues forming coils in Mel1a receptor. The secondary structure prediction using GOR shows 73 residues forming alpha helix, 102 residues forming extended beta strand and the remaining 187 residues forming coils in Mel1b receptor.

GOR4 Result for MEL1a:

```
      10       20       30       40       50       60       70
      |        |        |        |        |        |        |
MQGNGSALPNASQPVLRGDGARPSWLASALACVLIFTIVVDILGNLLVILSVYRNKK
LRNAGNIFVVSLA
ccccccccccccceeeccccchhhhhhhhhhhhhhhhhhhhhhcccceeehhhhhhhhhhcccccchhhhhhh
VADLVVAIYPYPLVLMSIFNNGWNLGYLHCQVSGFLMGLSVIGSIFNITGIAINRYCYI
CHSLKYDKLYS
hhhhhhhhcccccceeeeeccccccceeeeeeeccceeeccceccccccccceeeeeeeeeeccccccceec
SKNSLCYVLLIWLLTLAAVLPNLRAGTLQYDPRIYSCTFAQSVSSAYTIAVVVFHFLV
PMIIVIFCYLRI
cccceeeeeeehhhhhhhhhhcccccccceeccceeeeeeeccccccceeeeeeeccccccceeeeeeeee
WILVLQVRQRVKPDRKPKLKPQDFRNFVTMFVVFVLFAICWAPLNFIGLAVASDPAS
MVPRIPEWLFVAS
eeeeeeeeccccccccccccccccccccceeeeeeeeeccccccceeeeeeccccccccccccceeeeee
```

YYMAYFNSCLNAIIYGLLNQNFRKEYRRIIVSLCTARVFFVDSSNDVADRVKWKPSP
LMTNNNVVKVDSV

Eeeeeccccceeeecchhhhhhhhceeeeeeccccceeeeccccccccceeeccccceeccceeeeeec

Sequence length: 350

GOR4 Result for MEL1b:

10 20 30 40 50 60 70
| | | | | | |

MSENGSFANCCEAGGWAVRPGWSGAGSARPSRTPRPPWVAPALSAVLIVTTAVDV
VGNLLVILSVLRNRK

ccccccceeceeccceeeccccccccccccccccccccccccccceeeeeeeeccccccccchhhhhhhhhhh

LRNAGNLFLVSLALADLVVAFYPYPLILVAIFYDGWALGEEHCKASAFVMGLSVIGS
VFNITAIAINRYC

hhcccchhhhhhhhhhhhhhcccccceeeeeeeccccccccccccceeeeceeeeccccccceeeeeccee

YICHSMAYHRIYRRWHTPLHICLIWLLTVVALLPNFFVGSLEYDPRIYSCTFIQTASTQ
YTAAVVVIHFL

eeeccceeceeeeeeccceeeeeeeeeeeccccccccccccccccceeeeeeeeccccccchhhhhhhhcc

LPIAVVSFCYLRIWVLVLQARRKAKPESRLCLKPSDLRSFLTMFVVFVIFAICWAPLN
CIGLAVAINPQE

cccceeeeeeechhhhhhhhhhhhccccccceeccccccccchhhhheeeeeeeeccccccccceeeeccccc

MAPQIPEGLFVTSYLLAYFNSCLNAIVYGLLNQNFRREYKRILLALWNPRHCIQDASK
GSHAEGLQSPAP

ccccccccchhhhhhhhhhccccccceeeecchhhhhhhhhhhhhheeeccccccccccccccccccccccc

PIIGVQHQADAL

Ceeeeccceeec

Sequence length: 362

**Table 2: GOR Method**

| GOR 4 Method | MEL1a (350). | MEL1b (362). |
|---|---|---|
| Alpha helix (Hh) | 63 is 18.00% | 73 is 20.17% |
| $3_{10}$ helix (Gg) | 0 is 0.00% | 0 is 0.00% |
| Pi helix (Ii) | 0 is 0.00% | 0 is 0.00% |
| Beta bridge (Bb) | 0 is 0.00% | 0 is 0.00% |
| Extended strand (Ee) | 127 is 36.29% | 102 is 28.18% |
| Beta turn (Tt) | 0 is 0.00% | 0 is 0.00% |
| Bend region (Ss) | 0 is 0.00% | 0 is 0.00% |
| Random coil (Cc) | 160 is 45.71% | 187 is 51.66% |
| Ambigous states (?) | 0 is 0.00% | 0 is 0.00% |
| Other states | 0 is 0.00% | 0 is 0.00% |

3. **EsyPred3D**
   **Melatonin receptors modelled based on the template 1F88, chain A:**

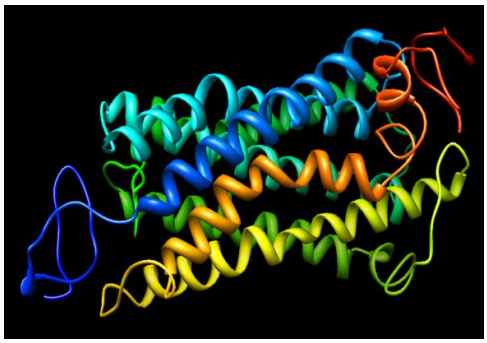

**Figure 3a: Mel 1a in Wireframe Form**

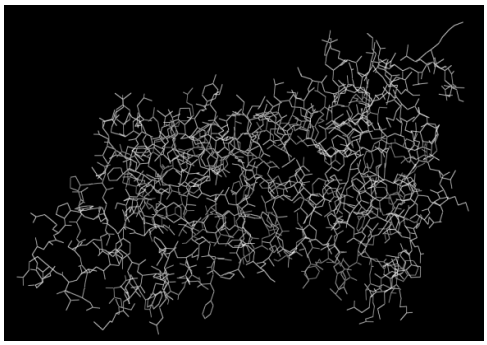

**Figure 3b: Mel1a in Ribbon Form**

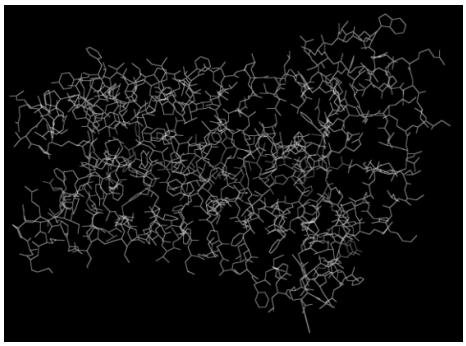

**Figure 4a: Mel1b in Wireframe Form**

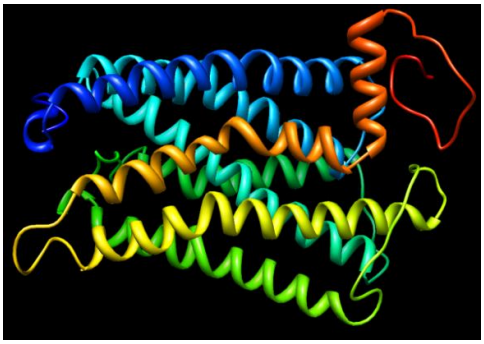

**Figure 4b: Mel1b in Ribbon Form**

**4. 3DJigSaw**

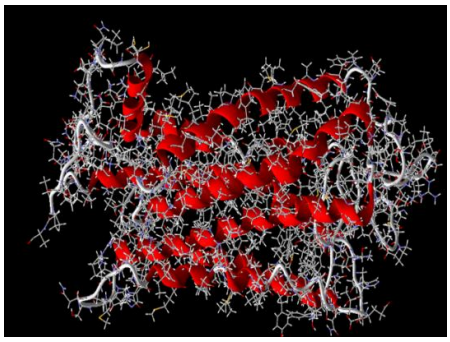

**Figure 5a: Mel 1a Initial energy: 12362.9**

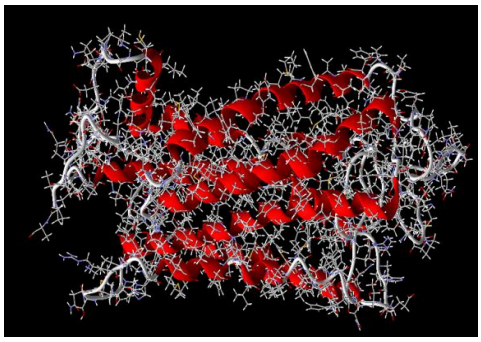

**Figure 5b: Mel 1a Final energy: 9527.58**

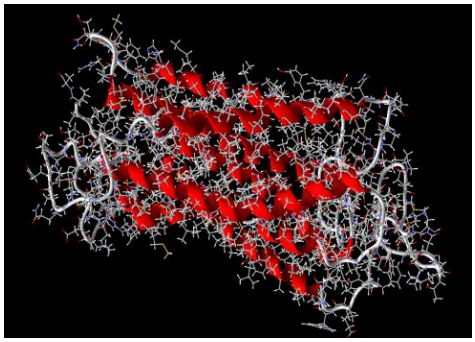

**Figure 6a: Mel1b Initial energy: 12307.00**

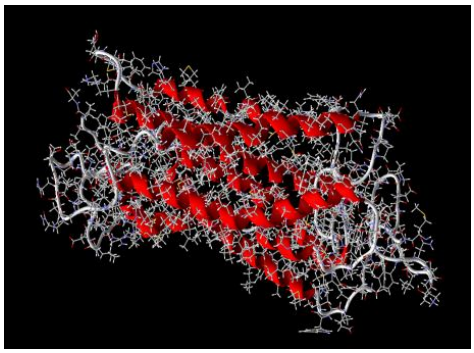

**Figure 6b: Mel1b Final energy: 9425.38**

## CONCLUSION

Three-dimensional (3D) structural information is very critical for understanding the functional properties of proteins. Therein, 3D structures are a valuable source of data for understanding their biological roles, their potential implication in diseases, and for progress in drug design.

## ACKNOWLEDGEMENT

The authors thanks the above three department for granting permission to carry out the research work.

## REFERENCES

**[1]** Berman H, Henrick K, Nakamura H and Markley JL, The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data, Nucleic Acids Res., 35, 2007, D301–3.

**[2]** Wang G and Dunbrack RL Jr., PISCES: recent improvements to a PDB sequence culling server, Nucleic Acids Res., 33, 2005, W94–8.

**[3]** Lo Conte L, SCOP: a structural classification of proteins database, Nucleic Acids Res., 28, 2000, 257–9.

**[4]** Perutz MF, Kendrew JC and Watson HC, Structure and function of haemoglobin, Journal of Molecular Biology, 13, 1965, 669–678.

**[5]** Browne WJ, North AC and Phillips DC, A possible three-dimensional structure of bovine alphalactalbumin based on that of hen's egg-white lysozyme, Journal of Molecular Biology, 42, 1969, 65–86.

**[6]** Qiang Wang and Adrian A, Canutescu, and Roland L. Dunbrack Jr., SCWRL and MolIDE: Computer programs for side-chain conformation prediction and homology modelling Fox Chase Cancer Center Cottman Avenue Philadelphia PA 19111, Nat. Protoc., 3(12), 2008, 1832–1847, doi:10.1038/nprot.2008.184.

**[7]** Trussell J, Contraceptive Efficacy, In Hatcher and Robert A, Contraceptive Technology, 19th Revised Edition, New York: Ardent Media, 2007, ISBN 0-9664902-0-7.

**[8]** Speroff Leon, Darney and Philip D, Oral Contraception- A Clinical Guide for Contraception, 4th Edition, Philadelphia: Lippincott Williams & Wilkins, 2005, 21–138. ISBN 0-781-76488-2.

**[9]** Loose DS and Stancel GM, Estrogens and Progestins, In Brunton, Laurence L.; Lazo, John S.; Parker, Keith L, Goodman & Gilman's the Pharmacological Basis of Therapeutics, 11th Edition, New York: McGraw-Hill, 2006, 1541–1571. ISBN 0-07-142280-3.

**[10]** Glasier A, Contraception, In DeGroot Leslie J, Jameson J. Larry, Endocrinology, 5th Edition, Philadelphia: Elsevier Saunders, 2006, 2993–3003. ISBN 0-7216-0376-9.

**[11]** Rivera R, Yacobson I and Grimes D, The mechanism of action of hormonal contraceptives and intrauterine contraceptive devices, Am. J. Obstet. Gynecol, 181 (5 Pt 1), 1999, 1263–9.

**[12]** Caniato R, Filippini R, Piovan A, Puricelli L, Borsarini A and Cappelletti E, Melatonin in plants, Adv. Exp. Med. Biol., 527, 2003, 593–7.

**[13]** Altun A and Ugur-Altun B, Melatonin: therapeutic and clinical utilization, Int. J. Clin. Pract., 61 (5), 2007, 835–45.

**[14]** Reppert SM, Melatonin receptors: molecular biology of a new family of G protein-coupled receptors, J. Biol. Rhythms, 12 (6), 1997, 528–31.

**[15]** Reppert SM, Weaver DR and Godson C, Melatonin receptors step into the light: cloning and classification of subtypes, Trends Pharmacology. Sci., 17 (3), 1996, 100–2.

**[16]** Sugden D, Davidson K, Hough KA and Teh MT, Melatonin, melatonin receptors and melanophores: a moving story, Pigment Cell Res., 17 (5), 2004, 454–60.

**[17]** Gudermann T, Nurnberg B and Schultz G, J. Mol. Med., 73, 1995, 51-63.

**[18]** http://www.biochem.vt.edu/modeling/homology.html.

**[19]** Andrew Kohlway, UMBC and Jeffry D, Madura, Homology Modeling and Docking of Melatonin Receptors, Duquesne University.

**[20]** Chugunov Anton, Efremov Roman, Molecular modeling of melatonin receptors, Laboratory of Biomolecular Modelling.

**[21]** Reppert SM, Weaver DR and Ebisawa T, Cloning and characterization of a Mammalian melatonin receptor that mediates reproductive and circadian responses, 13(5), 1994, 1177-85.

**[22]** Reppert SM, Godson C, Mahle CD, Weaver DR, Slaugenhaupt SA and Gusella JF, Proc Natl Acad Sci U S A, Molecular characterization of a second melatonin receptor expressed in human retina and brain: the Mel1b melatonin receptor. Laboratory of Developmental Chronobiology, Massachusetts General Hospital, Boston, MA 02114, USA. 92(19), 1995, 8734-8.