Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Structural Analysis and Functional Prediction of the Target Protein AR of Medicagol, an Active Component of Millettia reticulata

Yong-Fei Liu

1570184163@qq.com

Guizhou Light Industry Vocational University

Jing-Wen Li

Guizhou Light Industry Vocational University

Xue-Xu Yang

Guizhou Light Industry Vocational University

Mao-Zhao Wu

Guizhou Light Industry Vocational University

Tao Wang

Guizhou Light Industry Vocational University

Article

Keywords: Homology of medicine and food, Millettia reticulata, Medicagol, Androgen receptor, Structural analysis, Functional prediction, Bioinformatics

Posted Date: October 20th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7481346/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

To investigate the interaction mechanism between the steroidal active component Medicagol from Millettia reticulata and the androgen receptor (AR) using bioinformatics, and to explore its potential therapeutic effects. Active components of Millettia reticulata were screened using the TCMSP database. Multiple sequence alignment and phylogenetic tree construction of AR were performed with MEGA X (10.0.5). The primary structure was analyzed using ProtParam, hydrophobicity predicted by Protscale, and conserved domains identified via Prosite. Secondary structure was analyzed using NPS@SOPMA, while tertiary structure was predicted by homology modeling from the RCSB PDB database. Subcellular localization was predicted using DeepLoc-2.0, and molecular docking was performed with SwissDock to verify the interaction between Medicagol and AR. The AR protein consists of 920 amino acids with a molecular weight of 99.19 kDa and a theoretical isoelectric point of 6.00, classifying it as a hydrophilic and unstable protein. Its secondary structure is predominantly random coil (74.13%), with the presence of a nuclear receptor ligand-binding domain (NR LBD) and C4-type zinc finger motifs. Medicagol anchors into the AR ligand-binding pocket through hydrophobic interactions and hydrogen bonds, with key residues including Pro, Glu, Lys, and Ala. AR is highly conserved among primates, clustering closely with Pongo abelii and Gorilla gorilla gorilla. Molecular docking confirmed that Medicagol forms a stable complex with AR. This study elucidates the molecular binding mechanism of Medicagol with AR, providing theoretical support for the development of AR-targeted immunoregulatory drugs and the precise utilization of medicinal-food homologous resources.

Introduction

The concept of homology of medicine and food refers to natural resources that possess both medicinal and dietary values, and through dietary means, regulate bodily functions and prevent diseases. This embodies the traditional wisdom of Chinese medicine, which advocates that medicine and food share the same origin[1]. Millettia reticulata (chicken blood vine), a perennial woody vine from the Fabaceae family, is commonly used in traditional Chinese medicine and medicinal diets, possessing high medicinal value and health benefits. In China, M. reticulata is predominantly distributed in the southern and southwestern regions, particularly in Guangxi, Yunnan, Guizhou, Sichuan, and Guangdong provinces[2].In traditional Chinese medicinal diets, M. reticulata is often used in preparations such as stewed chicken with chicken blood vine or chicken blood vine soup, which are prescribed to invigorate blood and qi.

These formulations are primarily utilized for regulating menstrual disorders, postpartum weakness, and rheumatic pains in women3. The main active components of M. reticulata include isoflavones, triterpenes, and sterols, which exhibit pharmacological properties such as blood toning, blood circulation enhancement, antioxidant effects, anti-inflammatory activities, and immunomodulation[4]. Modern pharmacological studies have demonstrated that M. reticulata extracts promote hematopoiesis, vasodilation, and improved blood circulation, showing beneficial therapeutic effects for conditions such as anemia, menstrual irregularities, and rheumatoid arthritis[5].

Androgen receptor (AR), a ligand-dependent transcription factor belonging to the nuclear receptor superfamily, primarily mediates the biological effects of androgens such as testosterone and dihydrotestosterone (DHT)[6]. AR is widely expressed in tissues such as the prostate, muscles, bones, testes, and central nervous system, playing a pivotal role in male sexual differentiation, maintenance of sexual function, skeletal homeostasis, and muscle synthesis[7].

The mechanism of androgen receptor (AR) action includes both classical and non-classical pathways. In the classical pathway, androgens bind to cytoplasmic AR, inducing a conformational change, dimerization, and nuclear translocation. Once in the nucleus, AR binds to androgen response elements (AREs), thereby regulating the expression of target genes[8,9]. In contrast, the non-classical pathway involves cross-talk with signaling pathways such as MAPK and PI3K/Akt, leading to rapid modulation of cellular functions. This pathway is particularly relevant in the context of cancer initiation and progression[10]. Dysregulation of AR has been closely linked to diseases such as prostate cancer, androgen insensitivity syndrome, and polycystic ovary syndrome, making AR a critical target for various therapeutic strategies[11].

A comprehensive application of bioinformatics analysis[12], evolutionary analysis[13], and molecular docking[14] techniques was employed to systematically investigate the interaction between Medicagol, a sterol component of Millettia reticulata, and the androgen receptor (AR). Compared to traditional experimental methods, bioinformatics approaches can efficiently and accurately predict the potential targets of natural products and analyze their binding mechanisms, thus providing a theoretical basis for experimental validation and overcoming the limitations of high costs and long timelines. Building upon this, molecular docking was further utilized to validate the mode of receptor-ligand binding. In comparison with similar studies, this study reveals the mechanisms of natural product-nuclear receptor interactions through multi-dimensional data analysis, offering a new perspective for research on sterol components from Millettia reticulata and providing a theoretical foundation for drug development based on natural products. It also lays the groundwork for the development of novel immunoregulatory drugs targeting AR proteins.

Materials and Methods

Data Source

The data sources are shown in Table 1.

Table 1
Data sources

Database	Website
TCMSP	https://old.tcmsp-e.com/tcmsp.php
DeepLoc-2.0	https://services.healthtech.dtu.dk/services/DeepLoc-2.0/
protparam	https://web.expasy.org/protparam
protscale	https://www.expasy.org/protscale
Prosite	https://prosite.expasy.org
NPS@SOPMA	https://npsa.lyon.inserm.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html
RCSB PDB:Homepage	https://www.rcsb.org

Methods

Prediction of Active Ingredients and Targets of Millettia reticulata

TCMSP is a systems pharmacology platform specifically designed for the research and development of traditional Chinese medicine. It helps researchers explore the mechanisms of drug action within specific signaling pathways, different cell types, and under personalized conditions, thereby facilitating the screening of new drugs and the development of combination therapy strategies. Additionally, the database can be used to construct networks of drug-target-disease associations, revealing potential therapeutic mechanisms[15]. The screening criteria were set as oral bioavailability (OB) ≥ 30% [16] and drug-likeness (DL) ≥ 0.18[17].

Protein Sequence Alignment and Evolutionary Analysis

Protein sequences of the androgen receptor (AR) from Homo sapiens, Pan troglodytes, Theropithecus gelada, Pan paniscus, Gorilla gorilla gorilla, Pongo abelii, Piliocolobus tephrosceles, Papio anubis, Rhinopithecus roxellana, and Symphalangus syndactylus were subjected to multiple sequence alignment and phylogenetic tree construction using MEGA X (version 10.0.5). The analysis aimed to examine the conservation of AR family proteins. The sequences, including HsaAR, PtrAR-1, PpaAR, PabAR-3, GgoAR-X1, PteAR-X1, TgeAR-X1, PanAR-X1, and RroAR, were retrieved from the UniProt database. A phylogenetic tree was constructed using the Neighbor-Joining (NJ) method[18], with branch values representing the percentage of overlapping trees for related taxa (1000 bootstrap replicates). Larger branch values indicate closer evolutionary relationships, whereas smaller values suggest more distant relationships.

Physicochemical Property Analysis and Structure Prediction of Proteins

To predict the physicochemical properties of the target protein, the ProtParam tool was used for analysis, providing information such as molecular weight, theoretical isoelectric point, half-life, extinction coefficient, charge characteristics, and amino acid composition[19]. Additionally, the Protscale platform was employed to assess the hydrophobicity of the protein, outputting a graphical representation of the hydrophobic index for each amino acid residue as well as the overall hydrophobic trend of the protein. For functional structure prediction, the Prosite database was used to scan the protein's amino acid sequence for conserved domains, allowing for the identification of functionally relevant structural domains. The NPS@SOPMA website was utilized for secondary structure analysis, with its core algorithm, SOPMA, achieving a prediction accuracy rate of over 80%. This tool outputs the proportions of various secondary structures such as α-helices and β-sheets. The tertiary structure model was constructed based on the RCSB PDB database, providing a visual reference of the protein's three-dimensional spatial conformation.

Protein Subcellular Localization

The biological function of a protein within a cell depends on its subcellular environment, making its localization information crucial for understanding the protein's function and structural characteristics[20]. In this study, DeepLoc-2.0[21] was employed, a next-generation protein subcellular localization prediction platform that significantly enhances prediction accuracy compared to traditional tools like PSORT II.

Molecular Docking

Molecular docking is a key technique in computational drug design, which involves constructing a three-dimensional model of the interaction between a small molecule and its target (such as proteins or DNA) to predict their binding mode and affinity[22]. This method is widely used in drug discovery, target validation, and protein structure research[23]. In this study, SwissDock was utilized to perform docking calculations, and PyMOL software was used for preprocessing structural data and visualizing the results. The selected protein structures required relatively low resolution and were determined through X-ray crystallography.

Results and Discussion

Potential Active Compounds in Millettia reticulata

A total of 68 compounds were retrieved from the TCMSP database for Millettia reticulata. Among them, 24 compounds met the criteria of oral bioavailability (OB) ≥ 30% and drug-likeness (DL) ≥ 0.18, as shown in Table 2. Since the present study focused on the structural and functional characteristics of proteins,

DL was used as the primary screening index. Based on this criterion, Medicagol (DL = 0.6, OB = 57.49%) was selected as the compound of interest. Notably, Medicagol has so far been identified exclusively in Millettia reticulata. The TCMSP database was further queried for potential targets of Medicagol, which included the androgen receptor (AR), heat shock protein HSP90, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma isoform, serine/threonine-protein kinase Chk1, and the mRNA of cAMP-dependent protein kinase catalytic subunit alpha (PKA C-α). Among these, the androgen receptor was chosen for subsequent analysis.

Table 2
Effective components of Poria cocos

ID	Molecule Name	Molecular Weight	OB (%)	DL
MOL000469	3-Hydroxystigmast-5-en-7-one	428.77	40.93	0.78
MOL000033	(3S,8S,9S,10R,13R,14S,17R)-10,13-dimethyl-17-[(2R,5S)-5-propan-2-yloctan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-3-ol	428.82	36.23	0.78
MOL000449	Stigmasterol	412.77	43.83	0.76
MOL000296	hederagenin	414.79	36.91	0.75
MOL000358	beta-sitosterol	414.79	36.91	0.75
MOL000493	campesterol	400.76	37.58	0.71
MOL000491	Augelicin	426.5	37.5	0.66
MOL000470	8-C-α-L-arabinosylluteolin	418.38	35.54	0.66
MOL000503	Medicagol	296.24	57.49	0.6
MOL000507	Psi-Baptigenin	282.26	70.12	0.31
MOL000490	petunidin	317.29	30.05	0.31
MOL000497	licochalcone a	338.43	40.79	0.29
MOL000468	8-o-Methylreyusi	298.31	70.32	0.27
MOL000502	Cajinin	300.28	68.8	0.27
MOL000501	Consume close grain	302.3	68.12	0.27
MOL000483	(Z)-3-(4-hydroxy-3-methoxy-phenyl)-N-[2-(4-hydroxyphenyl)ethyl]acrylamide	313.38	118.35	0.26
MOL000461	3,7-dihydroxy-6-methoxy-dihydroflavonol	302.3	43.8	0.26
MOL000006	luteolin	286.25	36.16	0.25
MOL000471	aloe-emodin	270.25	83.38	0.24
MOL000492	(+)-catechin	290.29	54.83	0.24
MOL000417	Calycosin	284.28	47.75	0.24
MOL000500	Vestitol	272.32	74.66	0.21
MOL000392	formononetin	268.28	69.67	0.21
MOL000506	Lupinidine	234.43	61.89	0.21

Protein Sequence Alignment and Evolutionary Analysis

To investigate the evolutionary characteristics of the androgen receptor (AR) protein, multiple sequence alignments of AR proteins from Homo sapiens (human), Pan troglodytes (chimpanzee), Theropithecus gelada (gelada baboon), Pan paniscus (bonobo), Gorilla gorilla gorilla (Western lowland gorilla), Pongo abelii (Sumatran orangutan), Piliocolobus tephrosceles (gray-cheeked mangabey), Papio anubis (Anubis baboon), Rhinopithecus roxellana (Golden snub-nosed monkey), and Symphalangus syndactylus (siamangs) were performed using MEGA X (version 10.0.5) to analyze the conservation of the AR protein family (Fig. 2). The phylogenetic tree was constructed using the Neighbor-Joining (NJ) method, where the branch values represent the similarity between taxonomic groups. Higher branch values indicate closer evolutionary relationships, whereas lower values suggest more distant connections. The alignment results show that AR proteins across different species are highly conserved, with only a few amino acid differences, indicating that AR is highly conserved among primates. The phylogenetic tree analysis revealed that Homo sapiens (HsaAR) shares a close evolutionary relationship with Pongo abelii (PabAR-3) and Gorilla gorilla gorilla (GgoAR-X1), supporting the hypothesis of a common origin for humans and great apes[24].

Note

Hsa - Homo sapiens, Ptr - Pan troglodytes, Ppa - Pan paniscus, Ggo - Gorilla gorilla gorilla, Pab - Pongo abelii, Pte - Piliocolobus tephrosceles, Tge - Theropithecus gelada, Pan - Papio anubis, Rro - Rhinopithecus roxellana, Ssy - Symphalangus syndactylus.**a**: Multiple sequence comparison of AR gene-encoded proteins across different species.: Phylogenetic tree construction of AR gene-encoded proteins across different species.

Figure 2. Multiple comparisons and buildup of AR gene encoded proteins with proteins from different species

AR Gene-Encoded Protein Primary Structure Prediction

Prediction of the Physicochemical Properties of AR Gene-Encoded Protein

The primary structure of the AR protein was analyzed using ProtParam. The results indicated that the molecular weight of the protein is 99.19 kDa, and the theoretical isoelectric point (pI) is 6.00. The protein is composed of 920 amino acids, representing 20 different amino acid types, as detailed in Table 3. Among these, glycine (10.40%) is the most abundant, while tryptophan is the least abundant at 0.9%. The total number of negatively charged residues (Asp + Glu) in the AR protein is 92, accounting for 10%, while positively charged residues (Arg + Lys) total 81, representing 8.8%. These characteristics directly influence the hydrophilicity of the protein. In mammals, the AR protein has a half-life of 30 hours and an instability index (II) of 56.78, classifying it as an unstable protein. Moreover, the aliphatic index[25] is

68.04, indicating a high thermal stability of the protein. The AR protein's GRAVY value is -0.439, suggesting it is hydrophilic.

Table 3
Number and percentage of amino acids in each category

Name	Quantity	Percentage
Gly	96	10.40%
Leu	87	9.50%
Ala	81	8.80%
Ser	81	8.80%
Pro	74	8.00%
Gln	71	7.70%
Glu	55	6.00%
Arg	41	4.50%
Lys	40	4.30%
Val	40	4.30%
Asp	37	4.00%
Thr	37	4.00%
Tyr	33	3.60%
Phe	28	3.00%
Cys	27	2.90%
Ile	23	2.50%
Met	22	2.40%
Asn	20	2.20%
His	19	2.10%
Trp	8	0.9%

Hydrophobicity Prediction of the AR Gene-Encoded Protein

The hydrophobicity of the AR protein was predicted using the Protscale tool, employing the Kyte & Doolittle algorithm[26]. The results are shown in Fig. 3. A score greater than 0 indicates hydrophobicity, while a score less than 0 indicates hydrophilicity. The most hydrophobic amino acid, located at position 814, has a score of 2.778, whereas the lowest score of -3.500 occurs between positions 61 and 76. The

average hydrophobicity of the AR protein was calculated to be -0.441, which is less than zero, indicating that it is a hydrophilic protein. This result is consistent with the previous ProtParam analysis. Based on the combined analysis of its physicochemical properties and hydrophobicity prediction, AR is identified as a hydrophilic and unstable protein.

Prediction of Conserved Domains in the AR Gene-Encoded Protein

The amino acid sequence of the AR protein was analyzed using Prosite, and the results are shown in Fig. 4. The analysis revealed that the C-terminal region of the AR protein (positions 669–900) contains the Nuclear Receptor Ligand Binding Domain (NR LBD). The NR LBD belongs to the nuclear receptor superfamily, which can activate transcription factors and is widely involved in regulating various physiological functions, including individual development, reproductive processes, metabolic balance, and homeostasis[27]. Therefore, the activity of AR is highly dependent on the mechanism of action of the nuclear receptor-related structural domains. Additionally, AR contains two typical C4-type zinc finger domains within the 557–632 region. These domains typically consist of three or more cysteine residues and stabilize their conformation by binding to zinc ions. However, there is no high conservation between their sequences[28]. Zinc finger structures play a key regulatory role in DNA or RNA binding and protein-protein interactions, playing a central role in AR-mediated transcriptional and translational regulation. The presence of NR LBD and C4-type zinc finger structures suggests that AR may play a central role in nuclear receptor signaling pathways.

Prediction of the Secondary Structure of the AR Gene-Encoded Protein

The secondary structure of the AR protein was analyzed using the NPS @ SOPMA website (https://npsa.lyon.inserm.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html). The higher the content of α-helix and β-sheet, the more ordered the protein structure. As shown in Fig. 5, the AR protein contains three types of secondary structures, with random coils accounting for the highest proportion (74.13%), followed by α-helices (21.30%) and β-sheets (4.57%). Figure 6 displays the relationship curve of the various secondary structures.

Note

Blue: α-helix, Purple: β-sheet, Green: β-turn, Yellow: Random coil

Figure 6. Structural profiles of α-helix, β-folding, irregular curl and β-turns

Prediction of the Tertiary Structure of the AR Gene-encoded Protein

In this study, the tertiary structure of the AR protein was predicted using the RCSB PDB database (Fig. 7). This method employs homology modeling, searching for known proteins with similar structures in the PDB database to serve as templates for constructing and optimizing the structure model of the target protein. As shown in Fig. 7, the tertiary structure of AR consists of α-helix, β-sheet, and random coil, which is consistent with the results from the secondary structure prediction.

Prediction of Subcellular Localization of the AR Gene-encoded Protein

Protein localization is closely related to its functional mechanisms[29]. The subcellular localization of the AR protein was predicted using the DeepLoc-2.0 website, with results shown in Fig. 8. The most likely subcellular localization of AR is the nucleus (probability = 0.8104), followed by the cytoplasm (probability = 0.4178). Experimental validation of its precise cellular localization can further provide a theoretical foundation for drug design.

Molecular Docking

Molecular docking was used to analyze the interaction between the effective compound of *Radix Sanguisorbae* (Mei Dou Chun) and its potential target, the androgen receptor (AR). The structure of the protein has a resolution of 1.8 Å and has been validated by X-ray diffraction according to the database. Figure 9 shows the molecular docking visualization results. The optimal binding energy of Mei Dou Chun with the androgen receptor was found to be -5.570 kcal/mol, which is lower than − 5 kcal/mol, indicating a stable binding between them[30].

Conclusion

This study systematically analyzed the structural and functional properties of the androgen receptor (AR) protein, a potential target of the active compound Mei Dou Chun from Radix Sanguisorbae, using bioinformatics methods. Additionally, molecular docking was employed to validate the bioinformatics predictions. The study provides a comprehensive analysis of AR's structure and function and elucidates the molecular mechanism of action between Mei Dou Chun and AR. This work offers new insights into the target research of active ingredients from food-medicine homologous resources.By integrating phylogenetic analysis, protein structure modeling, and molecular docking, this study establishes a complete analytical process that encompasses "target identification—structural-functional analysis—mechanism validation." In contrast to traditional approaches relying on single methods (e.g., molecular docking or experimental validation), this strategy enables more efficient identification of key molecular players and the elucidation of their potential interaction pathways. This, in turn, enhances research efficiency and the accuracy of target screening. However, this study primarily relies on databases and computational models, making it a theoretical exploration. The predicted results still need to be validated through experimental evidence, such as cell-based experiments, animal models, or receptor activation

studies. Further experimental validation of these predictions would not only improve the reliability of the research but also lay the foundation for developing AR-targeted immunomodulatory drugs.

Declarations

Data availability

Data generated during the current study are available from the corresponding author upon reasonable request

Funding

This work was supported by the Guizhou Provincial Major Science and Technology Support Program (Qiankehe Zhicheng [2025] Major 001)

Author contributions

Y-F L: Conceptualization, Methodology, Writing. X-X Y, J-W L: Writing-review & editing. M-Z W,T W: Conceptualization, Writing-original draft, Writing-review & editing, Funding acquisition, Resources, Supervision

Competing interests

The authors declare no competing interests.

Additional informationCorrespondence and requests for materials should be addressed to Y-F L

References

1. Wu, C., Liu, J., Cheng, S., et al. Research progress on the medicinal and edible plants in Guangxi and their utilization. Guangxi Plants, 45(03), 450-465 (2025).
2. Xiang, Y., Liu, Y., Feng, J., et al. A study on the traditional textual research of Radix Sanguisorbae in classical prescriptions. Chinese Journal of Experimental Traditional Chinese Medicine, 31(06), 238-248.DOI:10.13422/j.cnki.syfjx.20241166 (2025).
3. Shi, L., & Ying, S. Practical studies on Chinese medicinal food therapy. Beijing: China Medical Science and Technology Press (2019).
4. Zhang, Y., Xiao, F., Zhao, X., et al. Research progress on the chemical components and pharmacological activities of Radix Sanguisorbae. Guangzhou Chemical Engineering, 52(24), 13-16 (2024).
5. Liao, J., Jin, C., Chen, Z., et al. Research progress on the chemical components, pharmacological effects of Radix Sanguisorbae and prediction of its quality markers (Q-Marker). Chinese Herbal Medicines, 54(20), 6866-6877 (2023).

6. Gao, W., Bohl, C. E., & Dalton, J. T. Chemistry and structural biology of androgen receptor. Chemical Reviews, 105(9), 3352-3370 https://doi.org/10.1021/cr020456u (2005).
7. Barsky, S. T., & Monks, D. A. The role of androgens and global and tissue-specific androgen receptor expression on body composition, exercise adaptation, and performance. Biology of Sex Differences, 16(1), 28 https://link.springer.com/article/10.1186/s13293-025-00707-6 (2025).
8. Asim, M. Decoding androgen receptor signaling: Genomic vs. non-genomic roles in prostate cancer. Neoplasia, 58, 101066. https://doi.org/10.1016/j.neo.2024.101066 (2024).
9. Safı, R., Wardell, S. E., Watkinson, P., et al. Androgen receptor monomers and dimers regulate opposing biological processes in prostate cancer cells. Nature Communications, 15(1), 75-76 https://doi.org/10.1038/s41467-024-52032-y (2024).
10. Dotto, G. P., Buckinx, A., Özdemir, B. C., et al. Androgen receptor signaling in non-prostatic malignancies: Challenges and opportunities. Nature Reviews Cancer, 25(2), 93-108 https://doi.org/10.1038/s41568-024-00772-w (2025).
11. Obinata, D., Takayama, K., Inoue, S., et al. Exploring androgen receptor signaling pathway in prostate cancer: A path to new discoveries. International Journal of Urology, 31(6), 590-597 https://doi.org/10.1111/iju.15424 (2024).
12. Huang, J., Mao, L., Lei, Q., & Guo, A. Y. Bioinformatics tools and resources for cancer and their application. Chinese Medical Journal (English), 137(17), 2052-2064 (2024).
13. Wang, Z., Chen, J., Chen, J., et al. Identifying interdisciplinary topics and their evolution based on BERTopic. Scientometrics, 129(11), 7359-7384 https://doi.org/10.1007/s11192-023-04776-5 (2024).
14. HT, D. U., Wang, L., Ding, J., et al. Application status and challenges of molecular docking in the development of traditional Chinese medicine. China Journal of Chinese Materia Medica, 49(3), 671-680 https://doi.org/10.19540/j.cnki.cjcmm.20231013.703 (2024).
15. Ru, J. Construction and application of systematic pharmacology database and analysis platform of traditional Chinese medicine (Master's thesis). Northwest Agriculture and Forestry University, Yangling, China (2015).
16. Chen, B., Zhou, S., Yi, J., et al. Effects of tonifying Yang and returning five soup on the expression of PI3K, Akt, and PTEN in Cav-1-/- mice cerebral ischemia model. China Journal of Traditional Chinese Medicine Information, 27(04), 35-40 (2020).
17. Hussein, Y. T., & Azeez, Y. H. DFT analysis and in silico exploration of drug-likeness, toxicity prediction, bioactivity score, and chemical reactivity properties of the urolithins. Journal of Biomolecular Structure and Dynamics, 41(4), 1168-1177 https://doi.org/10.1080/07391102.2021.2017350 (2023).
18. Ge, T., Luo, X., Wang, Y., Sedlmair, M., Cheng, Z., Zhao, Y., Liu, X., Deussen, O., & Chen, B. Optimally ordered orthogonal neighbor joining trees for hierarchical cluster analysis. IEEE Transactions on Visualization and Computer Graphics, 30(8), 5034-5046 https://doi.org/10.1109/TVCG.2023.3284499 (2024).

19. Azimi, R., Ozgul, M., Kenney, M. C., et al. Bioinformatic analysis of small humanin-like peptides using AlfaFold-2 and Expasy ProtParam. Investigative Ophthalmology & Visual Science, 65(7), 1320-1320 (2024).
20. Villanueva, E., Smith, T., Pizzinga, M., et al. System-wide analysis of RNA and protein subcellular localization dynamics. Nature Methods, 21(1), 60-71 https://doi.org/10.1038/s41592-023-02101-9 (2024).
21. Thumuluri, V., Almagro Armenteros, J. J., Johansen, A. R., et al. DeepLoc 2.0: Multi-label subcellular localization prediction using protein language models. Nucleic Acids Research, 50(W1), W228-W234 https://doi.org/10.1093/nar/gkac278 (2022).
22. Muhammed, M. T., & Aki-Yalcin, E. Molecular docking: Principles, advances, and its applications in drug discovery. Letters in Drug Design & Discovery, 21(3), 480-495 https://doi.org/10.2174/1570180819666220922103109 (2024).
23. Raval, K., & Ganatra, T. Basics, types, and applications of molecular docking: A review. IP International Journal of Comprehensive and Advanced Pharmacology, 7(1), 12-16 http://dx.doi.org/10.18231/j.ijcaap.2022.003 (2022).
24. Spear, J. K., Grabowski, M., Sekhavati, Y., et al. Evolution of vertebral numbers in primates, with a focus on hominoids and the last common ancestor of hominins and panins. Journal of Human Evolution, 179, 33-59 https://doi.org/10.1016/j.jhevol.2023.103359 (2023).
25. Wei, F., Kouro, T., Nakamura, Y., et al. Enhancing mass spectrometry-based tumor immunopeptide identification: Machine learning filter leveraging HLA binding affinity, aliphatic index, and retention time deviation. Computational and Structural Biotechnology Journal, 23, 859-869 https://doi.org/10.1016/j.csbj.2024.01.023 (2024).
26. Ali, S., Chourasia, P., & Patterson, M. From PDB files to protein features: A comparative analysis of PDB bind and STCRDAB datasets. Medical & Biological Engineering & Computing, 62(8), 2449-2483 https://doi.org/10.1007/s11517-024-03074-3 (2024).
27. Rajan, S., & Yoon, H. S. Covalent ligands of nuclear receptors. European Journal of Medicinal Chemistry, 261(11), 58-69 https://doi.org/10.1016/j.ejmech.2023.115869 (2023).
28. Brown, R. S. Zinc finger proteins: Getting a grip on RNA. Current Opinion in Structural Biology, 15(1), 94-98 https://doi.org/10.1016/j.sbi.2005.01.006 (2005).
29. Lees, M. B., & Bizzozero, O. A. Structure and acylation of proteolipid protein. Myelin. Routledge, 237-255 (2023).
30. Maahury, M. F., & Allo, V. L. The computational calculation and molecular docking of aeroplysinin-1 as antibacterial. Indonesian Journal of Chemical Research, 9(2), 124-128 https://doi.org/10.30598//IJCR.2020.9-MIR (2021).

Figures

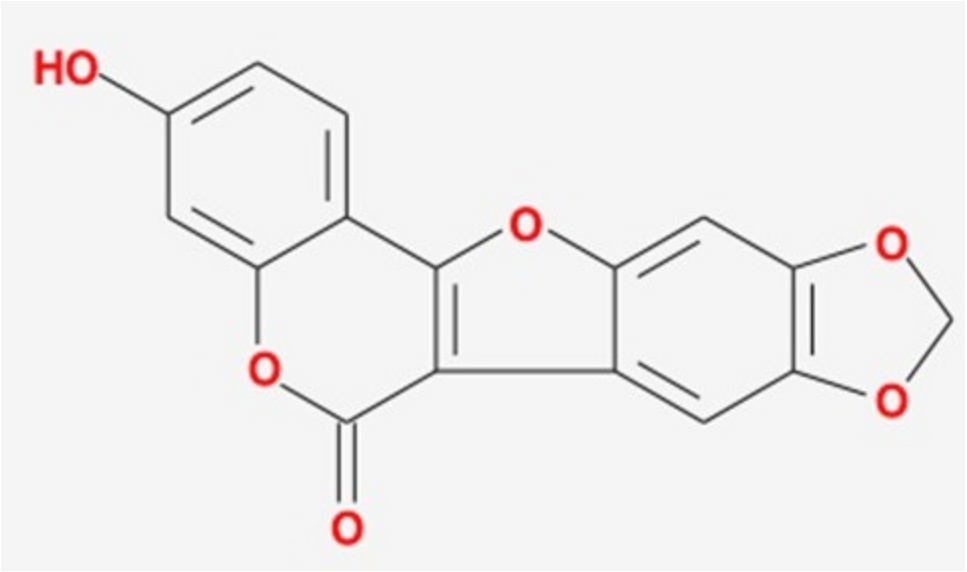
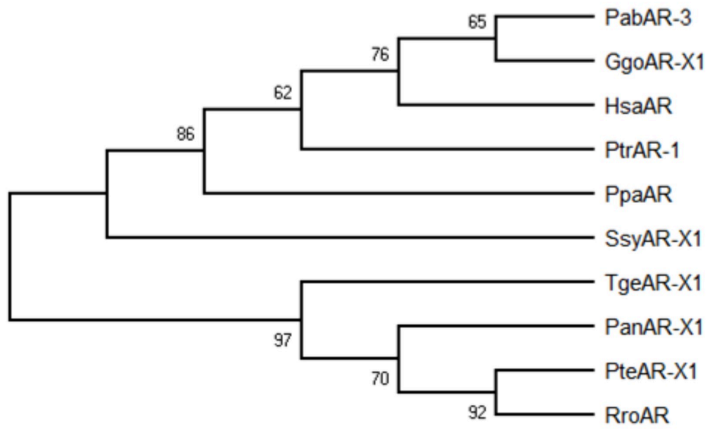


Figure 1

Molecular structure of Medicagol



Note: Hsa - Homo sapiens, Ptr - Pan troglodytes, Ppa - Pan paniscus, Ggo - Gorilla gorilla gorilla, Pab - Pongo abelii, Pte - Piliocolobus tephrosceles, Tge - Theropithecus gelada, Pan - Papio anubis, Rro - Rhinopithecus roxellana, Ssy - Symphalangus syndactylus.**a**: Multiple sequence comparison of AR gene-encoded proteins across different species.



b: Phylogenetic tree construction of AR gene-encoded proteins across different species.

Figure 2. Multiple comparisons and buildup of AR gene encoded proteins with proteins from different species

Figure 2

See image above for figure legend

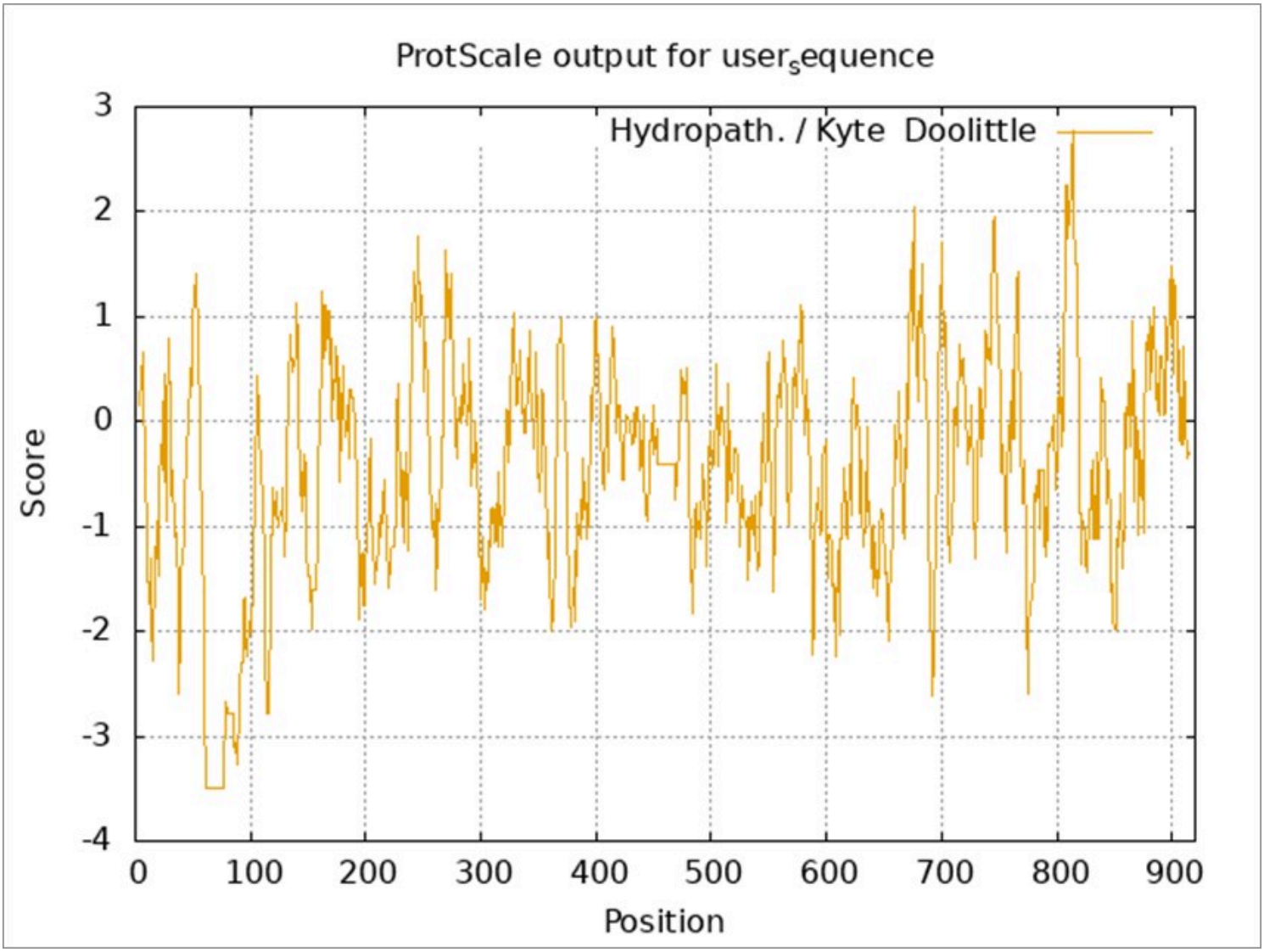


Figure 3

Predicted hydrophobicity of AR gene-encoded proteins

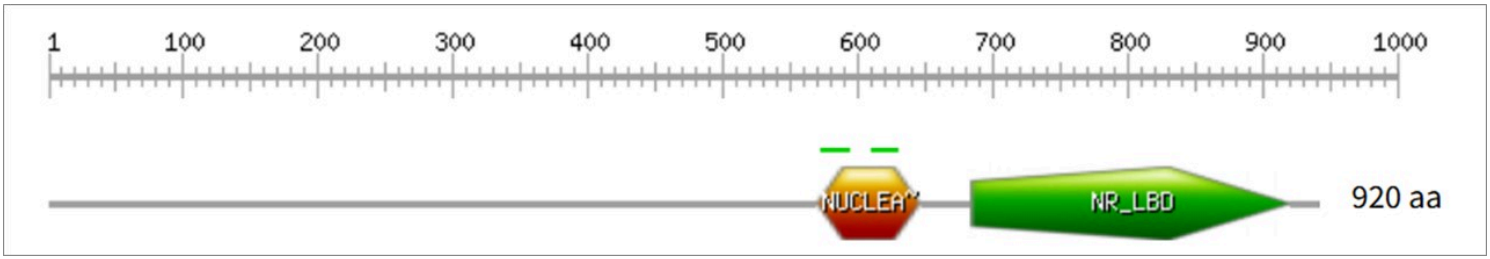


Figure 4

Conserved domains of protein encoded by AR gene



Figure 5

Predicted secondary structure of the protein encoded by the AR gene

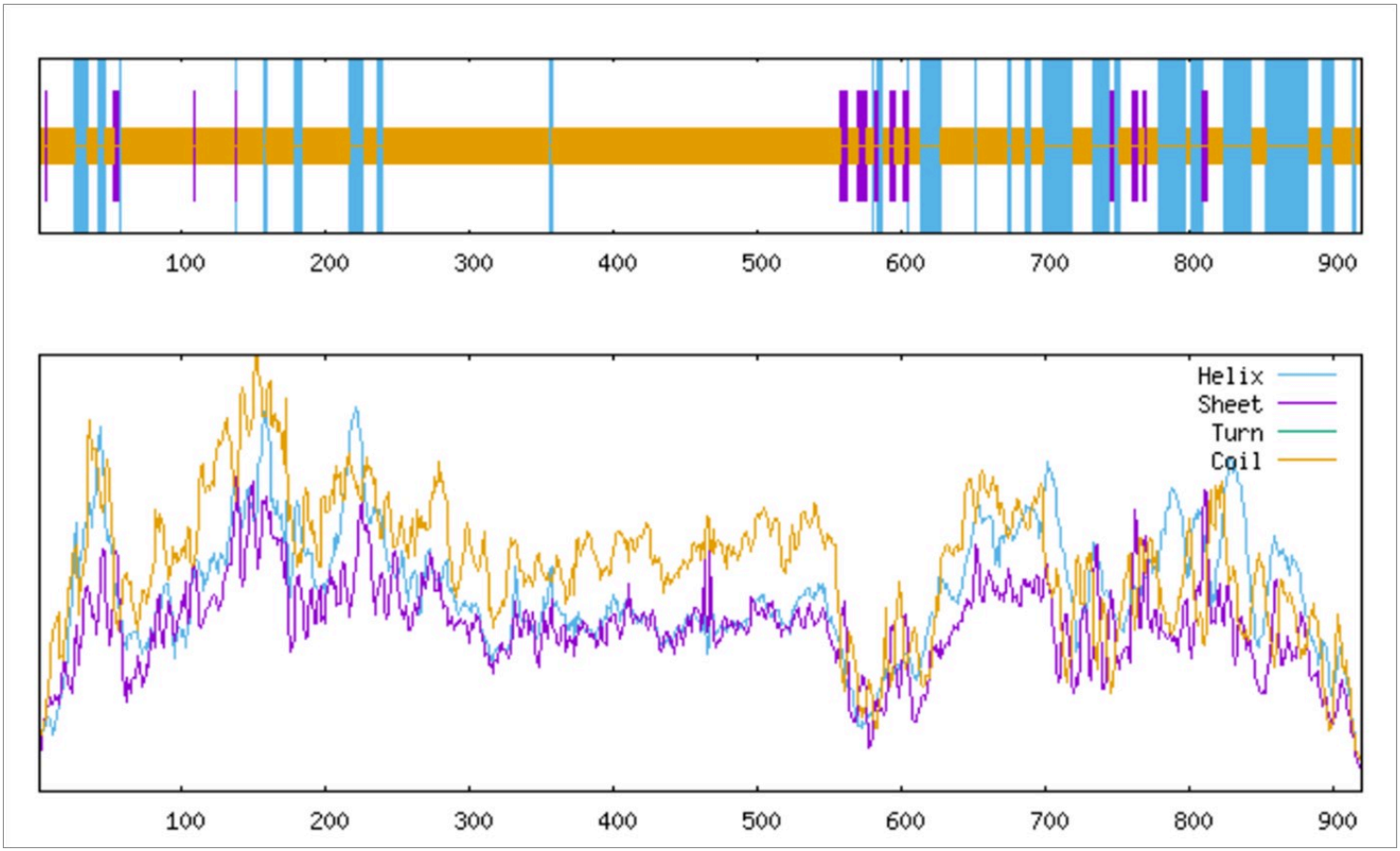


Figure 6

Structural profiles of α-helix, β-folding, irregular curl and β-turns

Note: Blue: α-helix, Purple: β-sheet, Green: β-turn, Yellow: Random coil



Figure 7

Tertiary structure of protein encoded by AR gene

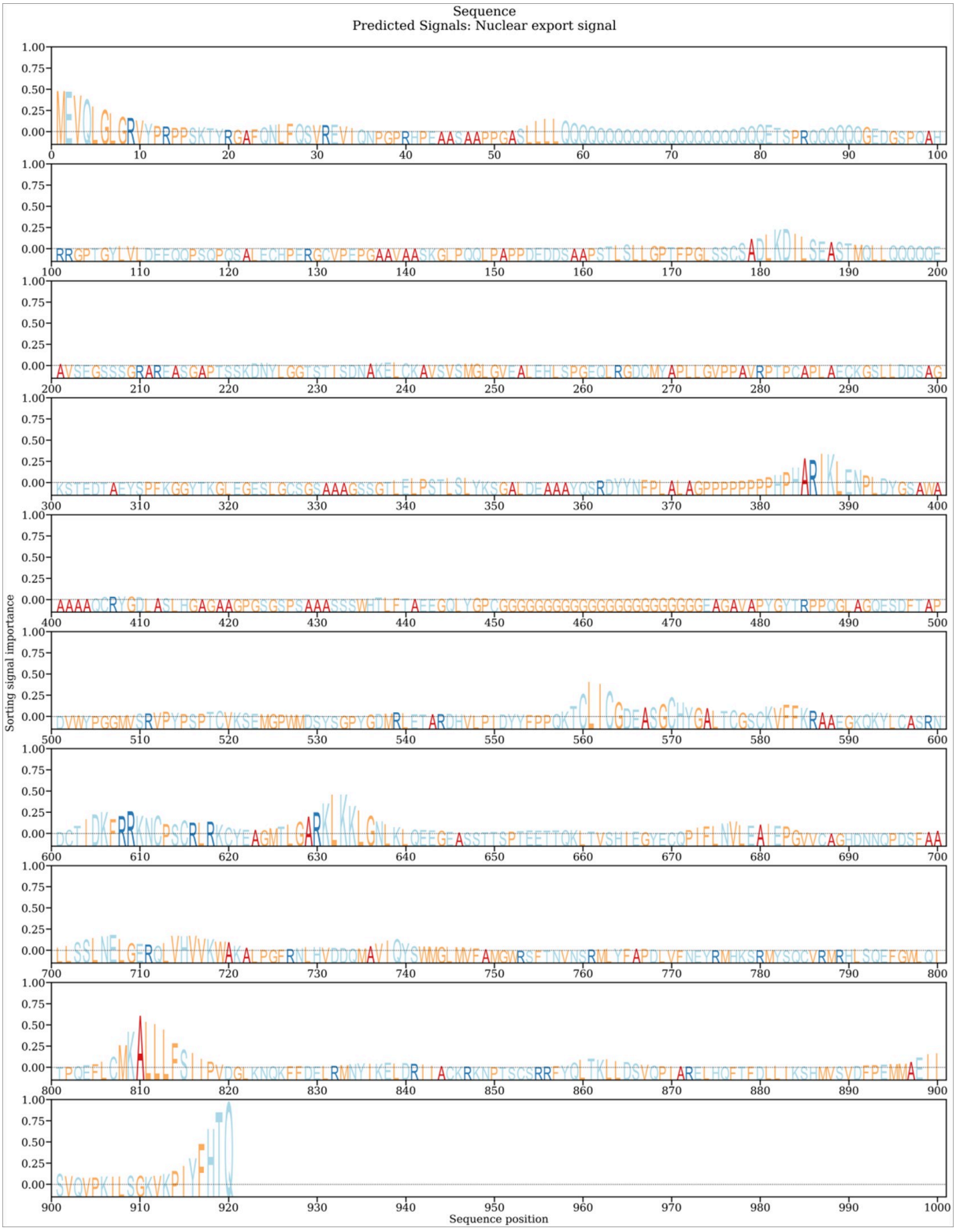


Figure 8

Subcellular Localization of AR Protein

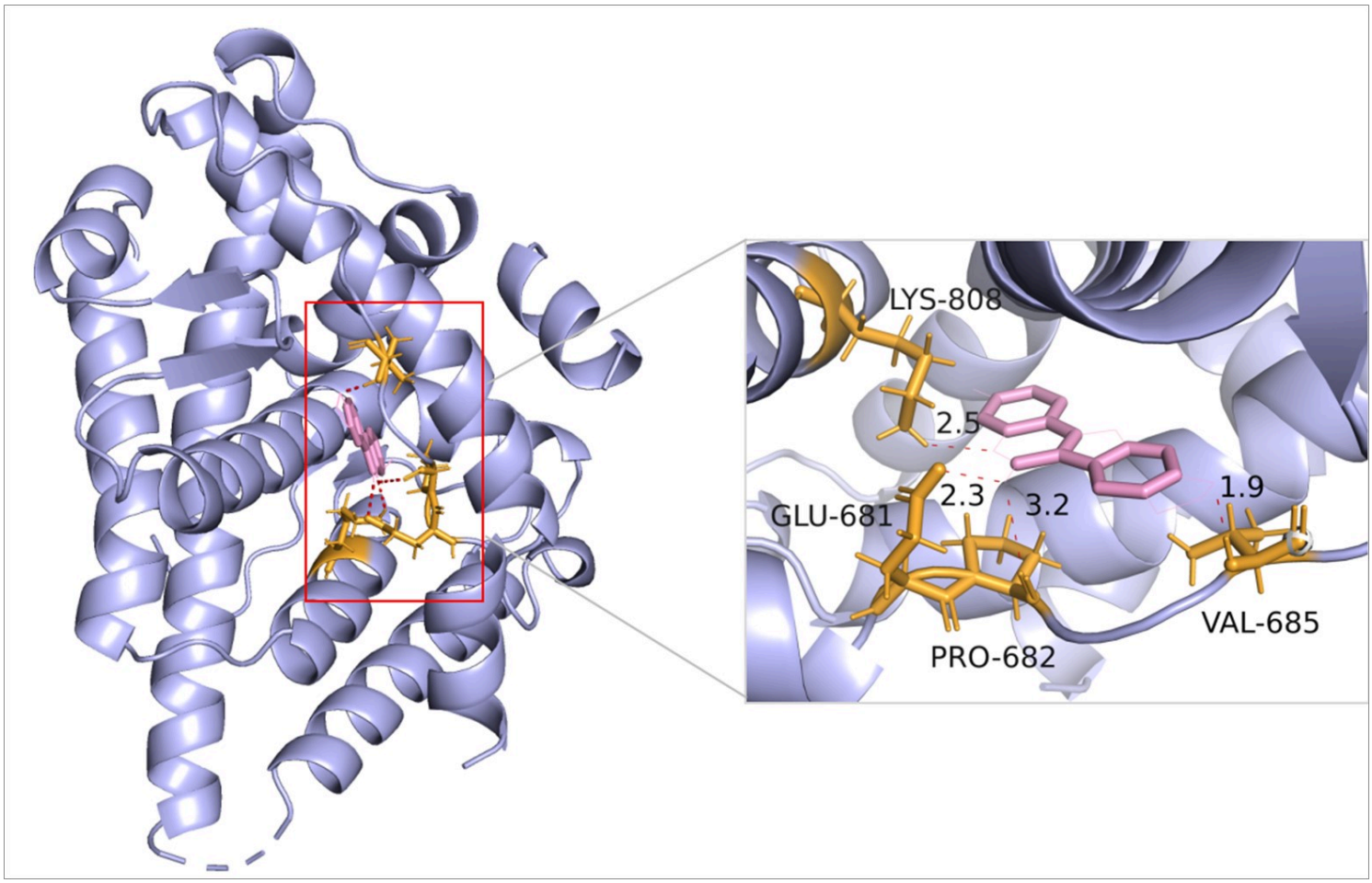


Figure 9

Molecular docking map of Medicagol with AR