

Sudeep Roy, Suaib Luqman, Ashok Sharma

Biotechnology Division, CSIR- Central Institute of Medicinal and Aromatic Plants, P.O: CIMAP, Lucknow-226015, U.P, India

Email: roysudeep28@gmail.com, s.luqman@cimap.res.in, ashoksharma@cimap.res.in

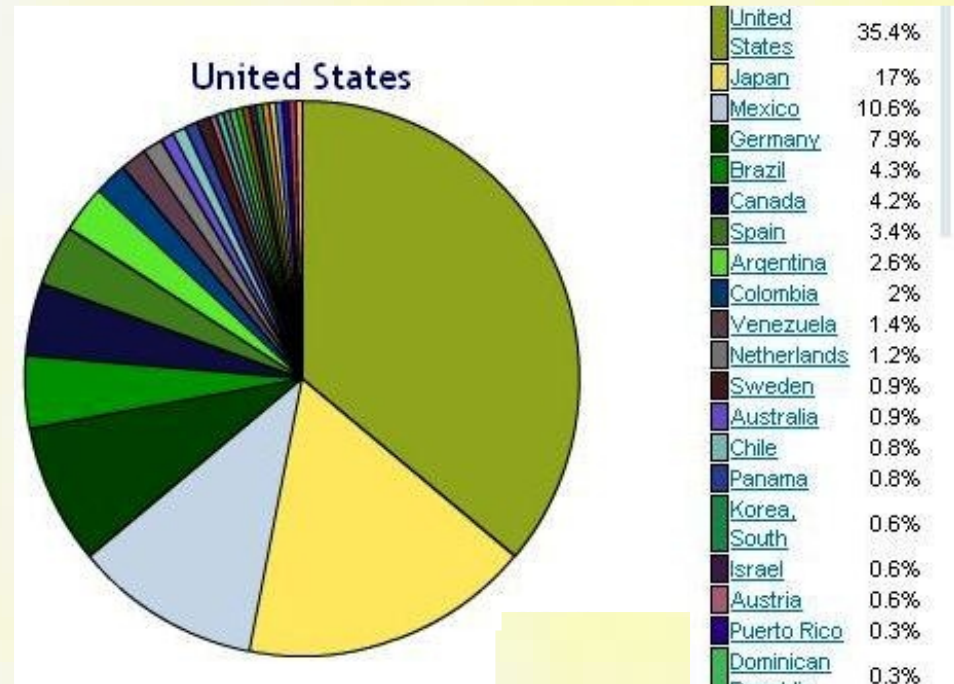
Introduction

- Candida albicans*, a fungus normally present on the skin and in mucous membranes such as the vagina, mouth, or rectum.
- Affects throat, intestines, and heart valves via travelling through blood stream.
- Commensal and a constituent of the normal gut flora comprising microorganisms that live in the human mouth and gastrointestinal tract.
- Lives in 80% of the human population without causing harmful effects, although overgrowth of the fungus results in candidiasis.

Cause

- Worms and Parasites
- Antibiotics and Stress
- Alcohol and Drugs
- Birth Control Pills

Candida infection statistics



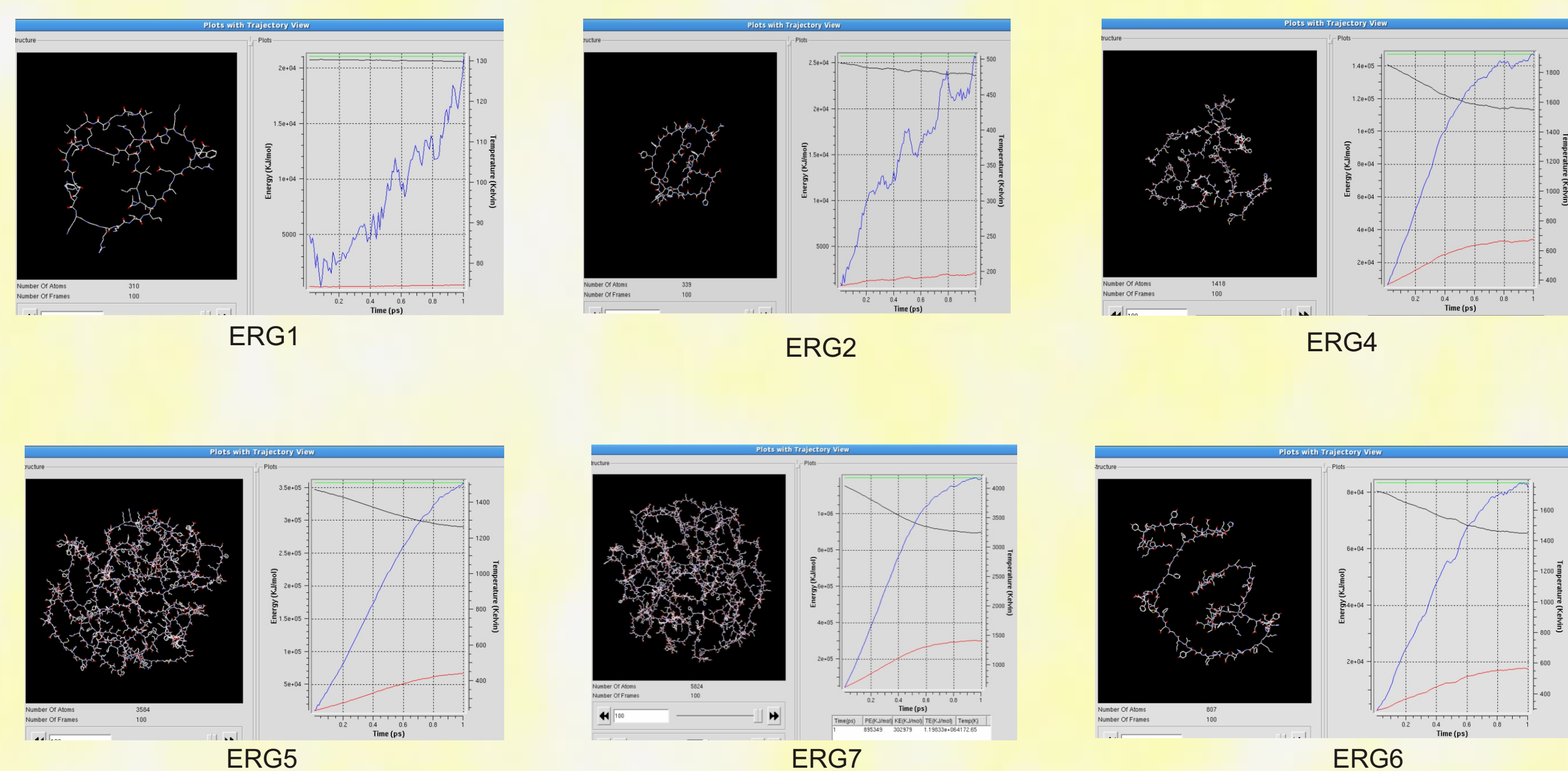
ERG PROTEINS

- Transcriptional regulator ERG is a protein, encoded by ERG gene in humans.
- Binds purine-rich sequences and is expressed at higher levels in early myelocytes than in mature lymphocytes.
- Act as a regulator of differentiation of early hematopoietic cells.
- Is vitally important to blood stem cells' unique ability to self-renew could give scientists new opportunities to use blood stem cells for tissue repair, transplantation and other therapeutic applications.

Objective and methodology of the proposed work

- Little information about the experimental structure (X-ray and NMR) of proteins from ergosterol biosynthetic pathway is available in RCSB Protein Databank (PDB).
- ERG proteins play a key role in metabolic pathway of ergosterol, their 3D structures are essential to determine most of their functions
- A program meant for comparative modeling, Modeller 9v7 and I-Tasser was utilized to serve our purpose.
- The modeled proteins were further validated by Procheck, Verify 3D, ERRAT and PROVE servers.
- ExPasy's Prot-param server was used for physico-chemical and functional characterization of these proteins.

Molecular dynamics results



Molecular dynamics ensembles result

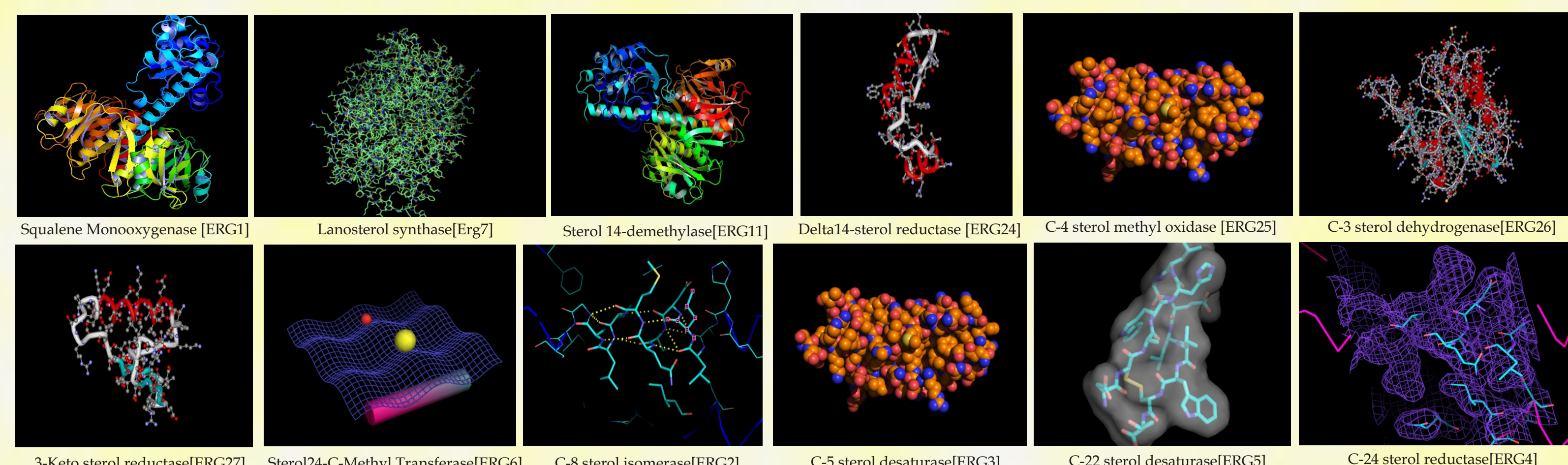
Protein Name	Ensembles	Time(ps)	PE(KJ/mol)	KE(KJ/mol)	TE(KJ/mol)	Temp(K)
ERG1	NVE	1(ps)	20545.2	503.736	21048.9	131.137
	NVT	1(ps)	20430.8	154.762	20585.6	40.289
ERG2	NVE	1(ps)	23630.2	2106.2	25736.4	501.121
	NVT	1(ps)	23687.4	1687.61	25375	401.527
ERG4	NVE	1(ps)	113091	33974	147065	1923.78
	NVT	1(ps)	112304	33741	146045	1910.59
ERG5	NVE	1(ps)	289549	67483.6	357033	1510.58
	NVT	1(ps)				

Suggested Readings

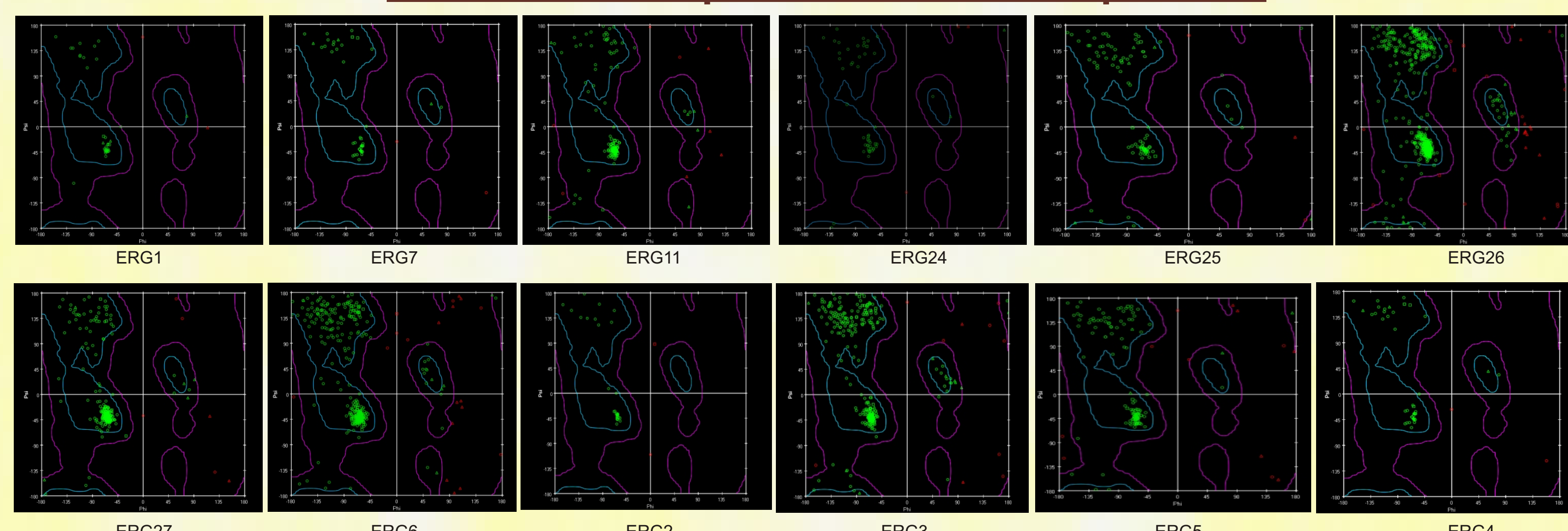
- Sali A. and Blundell T. L. 1993. *Journal of Molecular Biology*, 234, 779-815.A
- Guex, N. and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis*, 18: 2714-2723.
- RA Laskowski et al. 1993. *J Appl Crystallogr* 26: 283
- Ramachandran G.N., Ramakrishnan C. and Sasisekharan V. 1963. *Journal of Molecular Biology*, 7: 95-99.
- Eisenberg D., Luthy R. and Bowie J. U. 1997. *Methods in Enzymology*, 277, 396-404.
- Colovos C. and Yeates T.O. 1993. *Protein Science*, 2: 1511-1519.
- A. Rayan. 2009. *Bioinformatics* 3: 263
- Y. Zhang, J. Skolnick. 2005. TM-align: A protein structure alignment algorithm based on TM-score. *Nucleic Acids Research*, 33: 2302-2309
- Gill, S.C. and P.H. Von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, 182(2):319-326
- Ikeai, A., 1980. Thermostability and aliphatic index of globular proteins. *J. Biochem.* 88(6):1895-1898.
- Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157(1):105-132

Results

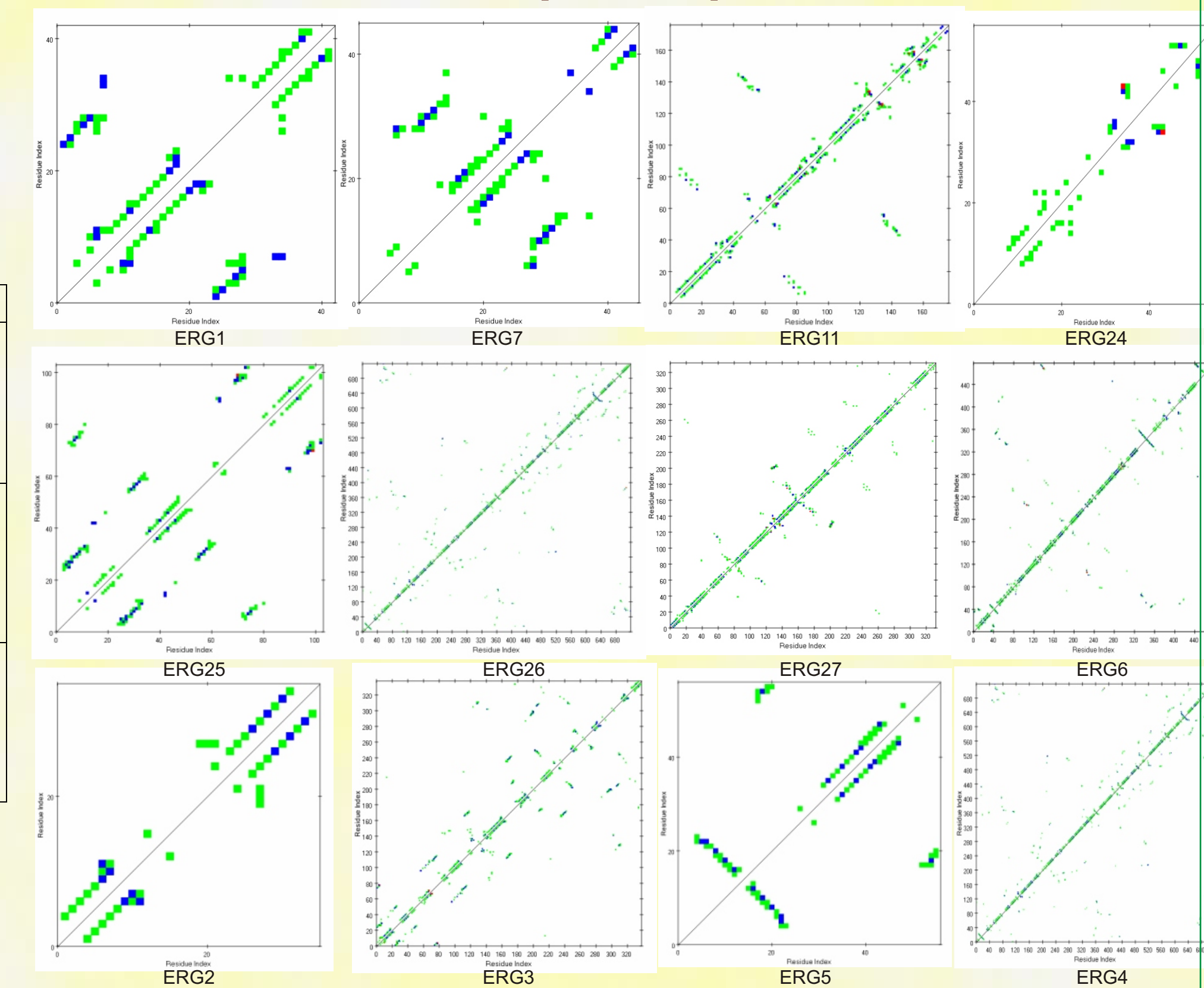
ERG modeled proteins



Ramachandran plot of the modeled proteins



Contact plot of proteins



Validation result

Protein Name	Modeling methodology	PMID details	Validation Procheck	Validation What-Check	Validation Verify 3D	Validation ERRAT	Validation PROVE
ERG9	Modeller	PMID7769 5	Ramachandran plot: 92.0% core, 6.0% allowed, 0.3% generously allowed, 1.7% disallowed	Stereochemical check: Ok	82.23% of the residues had an averaged 3D-1D score > 0.2	Overall quality factor: 82.26	Result: Satisfactory
ERG2	Modeller	PMID7766 1	Ramachandran plot: 95.9% core, 3.0% allowed, 0.0% generously allowed, 0.0% disallowed	Stereochemical check: Ok	73.55% of the residues had an averaged 3D-1D score > 0.2	Overall quality factor: 41.667	Result: Satisfactory
ERG6	Modeller	PMID7769 2	Ramachandran plot: 98.2% core, 10.1% allowed, 1.2% generously allowed, 0.6% disallowed	Stereochemical check: Ok	75.55% of the residues had an averaged 3D-1D score > 0.2	Overall quality factor: 45.989	Result: Satisfactory

Physicochemical characterization

Protein name	Sequence Length	Mol.wt.	pI	-R	-R	EC	Instability Index	Aliphatic Index	GRAVY
ERG1	496	55298.2	8.89	51	59	46675-46300	32.40	97.50	-0.033
ERG2	81	8773.0	5.75	7	3	7450	27.64	83.21	0.101
ERG3	386	45447.3	6.30	39	34	85510-85260	39.02	91.63	0.006
ERG4	469	54935.9	7.00	33	33	153725-153100	36.66	85.46	0.215
ERG5	517	59652.0	6.21	67	63	78770-78730	37.22	90.50	-0.171
ERG6	736	43085.5	5.74	58	47	60865-60740	31.52	71.57	-0.559
ERG7	730	83998.8	5.56	88	71	190540-189540	38.76	81.58	-0.301
ERG9	448	51171.2	6.57	54	52	46800-46300	36.42	99.82	-0.102
ERG11	528	60998.5	6.72	62	60	87460-	40.79	82.86	-0.272
ERG24	166	18848.2	6.54	11	11	28445-28420	28.40	125.06	0.485
ERG25	308	36560.9	6.83	28	26	105560-105310	34.77	85.16	-0.097
ERG26	350	39183.7	6.25	40	37	53080-52830	36.03	90.51	-0.226

Conclusion

- ERG9, ERG2, ERG6, ERG7, ERG11, ERG25 structures were successfully modeled and were found more stable than other ERG proteins.
- Molecular weight was observed between the range of 8773.0 - 83998.8 KDa for all ERG proteins in *Candida albicans*.
- All proteins were acidic in nature as their pH were less than 7.
- Aliphatic index analysis reveals high value for all ERG proteins of *Candida*.
- Higher aliphatic index of ERG proteins indicates that their structure are more stable over a wide range of temperature.
- The GRAVY value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence. The ERG proteins which have large negative values means those proteins are relatively more hydrophobicity as compared to proteins which have less negative values.
- Molecular Dynamics studies for different ensembles [NVE and NVT] were calculated. RMSD and standard deviations were also determined.

We are thankful to

- Council of Scientific and Industrial Research (CSIR), INDIA <http://www.csir.res.in>
- Department of Biotechnology (DBT), New Delhi, INDIA <http://dbtindia.nic.in>
- Council of Science and Technology, Uttar Pradesh

SBC@CIMAP

80th Annual Meeting of the Society of Biological Chemists (India)
12-15 November, 2011, Lucknow