

Research Article

Evaluation of SAR for Amphotericin B Derivatives by Artificial Neural Network

S Sardari*¹ and M Dezfulian²

¹Dept of Biotechnology, Pasteur Institute, Tehran, Iran ² Dept of Biology, Ahvaz University, Ahvaz, Iran

Abstract

Aim: This study was designed to investigate the role of several descriptive structure-activity features in the antifungal drug, amphotericin B and analyze them by artificial neural networks.

Method: Artificial neural networks (ANN) based on the back-propagation algorithm were applied to a structure-activity relationship (SAR) study for 17 amphotericin B derivatives with antifungal and membrane directed activity. A series of modified ANN architectures was made and the best result provided the ANN model for prediction of antifungal activity using the structural and biologic property descriptors.

Results: The best architecture, in terms of cycles of calculation was 12-15-2. Among the most important factors were biological descriptors that correlated best with the model produced by ANN. Among the chemical and structural descriptors, positive charge on Y substitution was found to be the most important, followed by lack of availability of free carboxyl and parachor.

Conclusion: This model is found to be useful to elucidate the structural requirements for the antifungal activity and can be applied in the design and activity prediction of the new amphotericin B derivatives.

Key words: Amphotericin B, SAR, Artificial Neural Network.

*Corresponding Author, ssardari@hotmail.com

Introduction

Amphotericin B (AmB) is a polyene discovered in 1956 and produced by *Streptomyces nodosus*. This agent is the first effective compound available for systemic mycosis, used against serious mycotic infections. AmB is regarded by many as the drug of choice for the seriously ill patients, in spite of a number of major drawbacks. It is nephrotoxic, poorly tolerated, and has such low solubility in biological fluids that it has to be formulated as a colloidal suspension in bile salts and administered as a slow infusion (Sarba, 1990). In addition, to reduce the toxicity, lipid-based and liposomal formulations of AmB have been either marketed or are undergoing further studies (Hiemenz, *et al.*, 1996; Wasan, *et al.*, 1997). Nevertheless, AmB is still claimed by many to be the best standard, against which other agents should be compared. It has a broad spectrum of activity and is fungicidal against most species, at least *in vitro*. The newer azole antifungals have still not replaced it for many indications, particularly in the immunocompromised hosts.

The polyenes exert their effect by associating with the eukaryotic membrane sterols and disrupting membrane integrity, as can be shown by a rapid efflux of potassium ions from treated cells (Hunter, 1995). It is believed that AmB inserts into membrane by having a high affinity for ergosterol in the fungal membrane. This affinity is lower for cholesterol, therefore, AmB shows a greater effect on fungi in low concentration than on host cells.

A better understanding of the structural features can lead to making derivatives with higher affinity to ergosterol and increased fungitoxicity compared to mammalian cell. Adverse effects would be important in rational design of new compounds with improved profile. A number of substituted amphotericin B molecule have been synthesized and their activity studied by principle component analysis methodology (Cheron, *et al.*, 1988). The QSAR of amphotericin B and its 16 semi-synthetic derivatives (Table 1), obtained by modification at carboxyl and amino groups,

were lead to certain assumptions on the role of net molecular charge on the results of biological tests (antibiotic action, erythrocyte K^+ release, haemolysis) by numerical methods (Cheron, *et al.*, 1988).

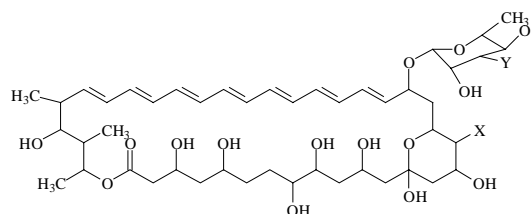


Figure 1. General structure of amphotericin B derivatives studied.

A number of computational methods have been found useful in the area of structure-activity relationship (SAR) and structure-property relationship (SPR) (Hansch, *et al.*, 1995, 2001; Mize, *et al.*, 1995; Estrada, *et al.*, 2001; Langgard, *et al.*, 2000). In this study, we report the application of neural network methodology to correlate SAR data and find new correlations between the biological and physicochemical descriptor values such as parachor, lipophilicity, SMILES, log P and the membrane activity for amphotericin B and its derivatives. The following is the list of such descriptors:

- 1- *S. cerevisiae* YK50, the dosage that led to 50% potassium release from yeast cells compared to 100% value for the boiled cell suspension;
- 2- Log P, logarithm of n-octanol/water partition coefficient;
- 3- Erythrocyte H50, the dosage that led to 50% haemolysis;
- 4- Erythrocyte EK50, the dosage that led to 50% potassium release from erythrocytes compared to 100% value for the lysed cells;
- 5- Positive charge on Y functionality at pH= 7.4;
- 6- Negative charge on X functionality at pH= 7.4;
- 7- Positive charge on X functionality at pH= 7.4;
- 8- Lack of free $-COO^-$;
- 9- Ergosterol 410 UV band position in the presence of the compound tested;

Table 1. The neural network architecture produced with the descriptors for AmB derivatives with the error limit of <0.05

Compound	X	Y	IC ₅₀ (µg/ml)	IC ₅₀ (µg/ml)
			<i>S. cerevisiae</i>	<i>C.</i>
1 (AmB)	COOH	NH ₂	0.05	0.03
2	COOCH ₃	NH ₂	0.07	0.05
3	CONH(CH ₂) ₃ N(CH ₃) ₂	NH ₂	0.08	0.04
4	COOH	NHCOCH ₃	0.35	0.28
5	COOCH ₃	NHCOCH ₃	0.55	0.25
6	CONH(CH ₂) ₃ N(CH ₃) ₂	NHCOCH ₃	0.25	0.3
7	COOCH ₃	NHCOCH ₂ NH ₂	0.07	0.06
8	COOH	NHCOCH ₂ NH(CH ₃) ₂	0.08	0.08
9	COOCH ₃	NHCOCH ₂ NH(CH ₃) ₂	0.12	0.1
10	CONH(CH ₂) ₃ N(CH ₃) ₂	NHCOCH ₂ NH(CH ₃) ₂	0.13	0.12
11	COOCH ₃	NHCO(CH ₂) ₂ NH(CH ₃) ₂	0.1	0.12
12	COOH	NH-(CHCO)(CH ₂ CO)-NC ₂ H ₅	0.2	0.1
13	COOCH ₃	NH-(CHCO)(CH ₂ CO)-NC ₂ H ₅	1	1
14	CONH(CH ₂) ₃ N(CH ₃) ₂	NH-[(CHCO)(CH ₂ CO)]-NC ₂ H ₅	0.2	0.2
15	COOH	NH-CH ₂ -[C(OH)OCH ₂ (CH(OH)) ₃]	0.05	0.06
16	COOH	*N(CH ₃) ₃	0.08	0.06
17	COOCH ₃	*N(CH ₃) ₃	0.07	0.06

- 10- Cholesterol 410 UV band position in the presence of the compound tested;
- 11- SMILES code, Simplified Molecular Input Line Entry System, a chemical notation system based on the principles of molecular graph theory;
- 12- Parachor, an additive physical property of a substance related to its molar volume, determined by the kind, number of atoms and their manner of arrangement.

An artificial neural network (ANN) is a biologically inspired algorithm designed to simulate the way in which the human brain processes information. ANN, are composed of many processing elements (PE). Each processing element has inputs, transfer functions and output. Processing elements are connected with coefficients and are arranged in layers, *i.e.*, input layer, output layer and hidden layers in between (Sardari, and Sardari 2002). Application of ANN in pharmaceutical research is a new field with novel potentials to be discovered. A variety of areas have been described to benefit from such algorithm predictions that range from industrial design systems to optimal formulation prediction and

SAR evaluations (Rowe, et al., 1998; Agatonovic-Kustrin, et al., 2000). Here, the design and application of ANN for antifungal drug AmB and its synthetic derivatives is reported.

A set of 17 structurally related AmB derivatives was selected and their experimentally derived membrane activity on erythrocyte cells and antifungal values against *Candida albicans* and *Sacharomyces cerevisiae* were collected (Cheron, et al., 1988). A standard feed-forward network, with back propagation rule and with single hidden layer architecture was chosen applying the EasyNN, 8.01 (1999-2001). The number of neurons was kept minimum to avert an over-fitting problem, which is usually produced by more weights due to higher numbers of neurons in input and hidden layers. However, to produce the optimum architecture, powerful enough to model the functions and not create errors more than 0.05%, the number of hidden layer neurons were varied from 1-48. The following architectures were produced that met the error limit condition using least number of calculation cycles (Table 2). Higher numbers of hidden layer did not improve the performance,

yet can decrease the speed of calculation. This finding is in accordance with previous reports (Ripley, et al., 1996). The descriptor parameters, including Log P and parachor, and SMILES

best with the model produced by ANN. Among the chemical and structural descriptors, positive charge on Y substitution was found to be the most important, followed by lack of availability of

Table 2. The neural network architecture produced with the descriptors for AmB derivatives with the error limit of <0.05.

Architecture	Cycles	Average Error	Relative normalized network error for compounds		Order of importance of Molecular/Biological Descriptor
			Lowest	Highest	
12-1-2	57	0.047	9	5, 13, 4, 6, 1	1, 4, 3, 7, 5, 9, 10, 6, 8, 11, 2, 12
12-2-2	14	0.049	8	13, 6, 5, 4	1, 4, 3, 5, 8, 9, 6, 7, 10, 11, 2, 12
12-3-2	15	0.049	2	13, 5, 6, 4	1, 4, 3, 5, 8, 7, 6, 10, 9, 2, 11, 12
12-4-2	11	0.048	16	13, 5, 4, 6	1, 4, 3, 5, 8, 12, 6, 9, 11, 7, 10, 2
12-5-2	12	0.049	16	13, 5, 4, 6	1, 4, 3, 5, 12, 7, 8, 10, 9, 2, 6, 11
12-6-2	10	0.049	16	13, 5, 6, 4	1, 3, 4, 5, 12, 9, 8, 10, 2, 6, 7, 11
12-7-2	10	0.048	2	13, 6, 4, 5	1, 3, 4, 9, 5, 8, 6, 12, 11, 7, 10, 2
12-8-2	9	0.048	2	13, 4, 6, 5	1, 4, 3, 5, 12, 8, 9, 2, 6, 7, 10, 11
12-9-2	9	0.048	16	13, 4, 6, 5	1, 3, 4, 12, 5, 8, 2, 11, 10, 9, 6, 7
12-10-2	8	0.048	2	13, 4, 6, 5, 14	1, 4, 3, 5, 8, 2, 12, 9, 11, 6, 10, 7
12-11-2	8	0.047	16	13, 4, 6, 5	1, 3, 4, 12, 9, 8, 5, 11, 6, 10, 7, 2
12-12-2	7	0.048	16	13, 4, 6, 5	1, 3, 4, 5, 8, 12, 9, 6, 7, 11, 10, 2
12-13-2	8	0.049	16	13, 4, 6, 5	1, 3, 4, 5, 12, 8, 9, 11, 10, 6, 2, 7
12-14-2	9	0.049	16	13, 4, 6, 5	1, 3, 4, 12, 5, 8, 10, 11, 9, 2, 7, 6
12-15-2	6	0.049	16	13, 4, 6	1, 3, 5, 4, 12, 8, 9, 11, 6, 10, 2, 7
12-24-2	7	0.047	16	13, 4, 6	1, 3, 4, 12, 8, 10, 9, 11, 5, 2, 6, 7
12-48-2	18	0.047	2	13, 4, 6, 17	1, 4, 3, 5, 10, 11, 12, 9, 8, 6, 2, 7

coding were calculated using the Toolkit for Estimating the Physicochemical Properties of Organic Compounds, v. 1.0, John Wiley & Sons, Inc.

The best architecture, in terms of cycles of calculation was 12-15-2. The importance is relative to the greatest sum of absolute weights connected to the next layer of the architecture. Therefore, importance of an input descriptor is determined by the sum of the absolute values of the weights of all the outgoing architecture connections from the input node to the next layer. Among the most important factors were biological descriptors 1, 3, and 4, that correlated

free carboxyl and parachor. The least important chemical and structural descriptors are 2, 7, and 11. The correlation coefficients between the experimental and the predicted IC₅₀ value pertaining to all the compounds for *S. cerevisiae* and *C. albicans* obtained by ANN methodology are 0.94 and 0.82 respectively. This indicates the applicability of this method in prediction studies. Compounds 4, 5, 6, and 13 corresponded to the highest error that was generated during the training cycles. The least error values were attributable to compounds 2 and 16.

The results of this study show the ability of ANN algorithm in assigning the biological and physicochemical descriptors to the activity prediction of large antibiotic molecules such as AmB. The findings of this study reveals the significance of various factors participating in the desired antifungal activity; therefore such method can be applied in drug design and development of different derivatives with better profile of activity, less toxicity and more potency. In this regard, the produced architectures using ANN modeling can be applied to *in silico* study of the biological activity for new derivatives prior to synthetic steps.

References

1. Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J. Pharm. Biomed. Anal.* 2000; 22: 717-727.
2. Cheron M, Cybulska B, Mazerski J, Grzybowska J, Czerwinski A, Borowski E. Quantitative structure-activity relationships in amphotericin B derivatives. *Biochem. Pharmacol.* 1988; 37: 827-36.
3. Estrada E, Uriarte E. Recent advances on the role of topological indices in drug discovery research. *Curr. Med. Chem.* 2001; 8: 1573-1588.
4. Hansch C, Hoekman D, Leo A, Zhang L, Li P. The expanding role of quantitative structure-activity relationships (QSAR) in toxicology. *Toxicol. Lett.* 1995; 79: 45-53.
5. Hiemenz, JW, Walsh TJ. Lipid formulations of amphotericin B: recent progress and future directions. *Clin. Infect. Dis.* 1996; 22: S133-44.
6. Hunter, PA. New developments in non-azole antifungals for human disease. In: Fifty Years of Antimicrobials: Past Perspectives and Future Trends, Hunter, PA, Darby, G K, and Russell, NJ (eds.) Cambridge USA, Cambridge University Press, 1995, pp. 19-52.
7. Langgard M, Bjornholm B, Gundertofte K. Pharmacophore modeling by automated methods: Possibilities and limitations. *IUL Biotechnol. Series*, 2000; 2: 237-250.
8. Mize NK, Johnson JA, Hansch C, Cormier M. Quantitative structure-activity relationship and cytotoxicity. *Curr. Probl. Dermatol.* 1995; 23: 224-9.
9. Ripley BD. Pattern recognition and neural networks, Cambridge, USA, Cambridge University Press, 1996.
10. Rowe RC, Roberts RJ. Artificial intelligence in pharmaceutical product formulation: neural computing and emerging technologies. *PSTT*, 1998; 1: 200-205.
11. Sarba R, Branch, RA. Amphotericin B nephrotoxicity. *Drug Saf.* 1990; 5: 94-108.
12. Sardari S, Sardari D. Applications of Artificial Neural Network in AIDS Research and Therapy, *Curr. Pharm. Des.* 2002; 8: 659-670.
13. Wasan KM, Lopez-Berestein G. Diversity of lipid-based polyene formulations and their behavior in biological systems. *Eur. J. Clin. Microbiol. Infect. Dis.* 1997; 16: 81-92.