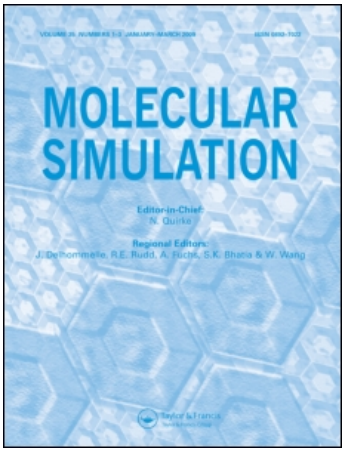

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Backpropagation neural network analysis applied to β-sheet breakers used against Alzheimer's amyloid aggregation

Peter Mager

Online publication date: 03 June 2010

To cite this Article Mager, Peter(2002) 'Backpropagation neural network analysis applied to β-sheet breakers used against Alzheimer's amyloid aggregation', Molecular Simulation, 28: 3, 239 — 247
To link to this Article: DOI: 10.1080/08927020290014358
URL: http://dx.doi.org/10.1080/08927020290014358

Molecular Simulation, 2002, Vol. 28(3), pp. 239–247

BACKPROPAGATION NEURAL NETWORK ANALYSIS APPLIED TO β-SHEET BREAKERS USED AGAINST ALZHEIMER'S AMYLOID AGGREGATION

PETER P. MAGER*

Research Group of Pharmacochemistry, Institute of Pharmacology and Toxicology of the University, D-04107 Leipzig, Härtelstr. 16-18, Saxony, Germany

(Received December 2000; accepted January 2001)

Structure-activity relationships of β-sheet inhibitors against Alzheimer's Aβ(1–42) amyloid aggregation were studied by backpropagation neural network analysis. It was found that the total and electrostatic energies of geometry-optimized conformations of the ligands, and the hydration energy, simulate the biological potency.

Keywords: Morbus Alzheimer; Aβ peptide; β-sheet breakers; Neural networks

1. INTRODUCTION

Alzheimer's disease (AD) is a chronic, neurodegenerative disorder which is characterized by pathological brain lesions composed of amyloid deposition. The major protein constituent of the deposits is the so-called amyloid β-peptide (Aβ peptide). Several variants of the naturally occuring Aβ's differing only at the C-terminus (amino acid residues 1–40, 1–42, 1–43) exist. It was demonstrated that a freshly prepared random-coil or an α-helix conformation of the Aβ peptides is nontoxic or less toxic, while an enhanced neurotoxicity is observed by transformation to a β-sheet conformation [1].

*e-mail: magp@server3.medizin.uni-leipzig.de

ISSN: 0892-7022 DOI: 10.1080/08927020290014358

Quite recently, it was found [2] that the short α-helical folding part in the lipophilic central area of monomeric Aβ peptides is one of the potential molecular targets of the binding of short anti-β-sheet ligands preventing thus polymerization to neurotoxic aggregates. The present study deals with structure-activity relationships of small peptide ligands against Alzheimer's Aβ peptide aggregation.

2. METHOD

2.1. Synthesis and Pharmacological Data

The synthesis routes of various peptide ligands that inhibit Alzheimer's amyloidosis were described elsewhere, together with all data that verify the structures [3–8]. The percentages of amyloid formation of a certain series of peptide ligands were determined by fluorescence spectroscopy and taken from a previous study [3].

2.2. Molecular Modelling

Geometry optimization was used to improve the geometry data and to get the most stable conformer [2]. The conformations were initially energy minimized using the MM+ force field without an electrostatic term. The MM+ empirical potential (force field) is an improved MM2/MM3 version [9, 10]. The whole MM+ procedure was repeated with electrostatic parameters of the connectivity-based iterative partial equalization of orbital electronegativity [11]. Correlation-gradient geometry optimization [12] was then achieved. The structures were refined using a conjugate gradient minimizer (Fletcher-Reeves modification of the Polak-Ribière method). Convergence was obtained when the gradient root mean square RMS was $RMS < 0.05$.

2.3. Chemical Parameters of Ligands

The following energies of the peptide ligands were calculated: total, bond, angle, dihedral, van-der-Waals, stretch-bend, and electrostatic energies. Also, the distribution coefficient P (octanol/water), the pK_a value (at $[K^+]+[Na^+] = 0.05\,M$, $[Cl^-] = 0.05\,M$; 25°C), mass, surface area (grid), molar volume, hydration energy, molar refractivity, polarizability, and hydrophobicity index were determined [13–17].

2.4. Theoretical Approaches

2.4.1. Neural Network

The algorithm of the generalized backpropagation neural network was described elsewhere [18]. The sigmoidal backpropagation functions were solved by the nonlinear Levenberg-Marquardt algorithm. The learning rate momentum was equal to 0.8, the learning rate minimum and maximum were equal to 0.001 and 0.3. The squared multiple correlation coefficients (R^2) were statistically tested using the maximum-likelihood criterion [19]. Optimization was achieved by (i) estimating the global error vector prior to adjusting weights, and (ii) updating successively the weights until convergence was reached.

2.4.2. Selection of the Physicochemical Descriptors

The neural network was combined with the statistical MASCA approach [19] to overcome the danger of overfitting in neural network analysis, that is, MASCA is used to select the physicochemical descriptors.

3. RESULTS AND DISCUSSION

It was demonstrated that freshly prepared random-coil and α-helix conformations of Aβ peptides are nontoxic or less toxic, while a neurotoxicity is observed after inducing a time-dependent aging of β-sheet conformations by a suitable chemical environment [21–27]. A possible target site for preventing aggregation by self-recognizing peptide ligands is the central moiety of Aβ peptides.

The problems of peptide ligand administration are known: poor oral absorption, rapid metabolic degradation due to proteolytic metabolism, transport through the blood-brain barrier producing central effects, binding to phospholipids, and generation of immune responses (in particular, if medium-sized peptides are administered). Nevertheless, at a first phase of research, pure peptide ligands are tested. Of particular interest is proline in Alzheimer's amyloid research. There are three characteristics of the proline residue: (i) the endocyclic nitrogen atom of proline is not available in β-sheet hydrogen-bonding networks, (ii) the ring of proline cannot sterically fit into the β-sheet network, and (iii) the conformation of a peptidyl–prolyl bond is incompatible with peptide bond geometries found in β-sheets. In other words, the proline ring cannot fit sterically within β-sheets, and does not

occur within the interior of antiparallel β-sheets. Therefore, Val17 of the Aβ(1–42) peptide was replaced by proline ("mutation") in designing short peptide ligands.

The ligand LPFFD, that is,

$$(NH_2)\text{LeuProPhePheAsp}(CO_2H)$$

is now considered as reference compound. The two phenylalanine residues of LPFFD are in α-helix conformation, proline has a parallel β-sheet conformation, and Leu and Asp are in random-coil conformation. The logarithms of the distribution coefficients ($\log P$, octanol/water) are: $\log P = -1.92$ at $p\text{H} = 2$ (stomach), $\log P = -2.82$ at pH 6.6 (inflammatory tissue), $\log P = -3.25$ at $p\text{H} = 7.4$ (blood plasma, brain), and $\log P = 0.29$ (neutral microspecies). Therefore, it is clear that LPFFD cannot orally be administered, it cannot permeate through the blood-brain barrier. However, an active transport and/or permeation by the *Plexus chorioideus* and/or liquor-brain barrier cannot be ruled out, because of the zwitter-ionic nature of LPFFD ($p\text{K}_a$ values of the two free acidic groups of Asp: 2.91, 4.51; $p\text{K}_a$ of the basic N^+H_3 group of Leu: 7.81). The isoelectric point is IP = 3.70, at $p\text{H} = 7.4$ the ligand is negatively charged (−1.25).

The algorithms of the MASCA approach [19] were used to select the relevant physicochemical descriptors of peptide ligands. The following parameters were obtained (Tab. I): total energy (E_{tot}, kcal/mole), electrostatic energy (E_{el}, kcal/mole), and hydration energy (E_{hy}, kcal/mole). The

TABLE I Percentage (%) of Aβ(1–42) aggregation [3] and physicochemical descriptors of Gasteiger-MM+ geometry-optimized conformations of ligands (L-configuration): total energy (E_{tot}, kcal/mole), electrostatic energy (E_{el}, kcal/mole), and hydration energy (E_{hy}, kcal/mole)

Compd.	*Ligand*	%*	E_{tot}	E_{el}	E_{hy}
1	RDLPFFDVPID	57.9	31.7382	4.1737	−20.7810
2	RDLPFFPVD	85.7	25.2474	0.9452	−18.2760
3	LPFFPVD	61.3	49.4182	26.7469	−8.0680
4	LPFFVD	87.1	33.6044	21.3307	−8.2490
5	LPFF	82.4	31.2694	19.4475	−5.1230
6	PFF	88.9	−3.7944	−10.5543	−5.7640
7	LVFFA	93.3	−26.6232	−29.5455	−6.9700
8	LPFFD	45.4**	44.0812	19.0908	−11.2260
9	VFFA	98.5	−19.1764	−21.0511	−7.6670
10	YEVHHQKLVFF	80.9	−78.0475	102.4910	−37.0220
11	VHVSEEGTEPA	101.3	−52.3364	−88.0704	−20.8600
12	GYLTVAAVFRG	125.3	−73.6470	−75.4149	−29.6350
13	PADVLPLA-PRAVD	94.4	−32.6885	−79.3677	−19.7180

* 100% = aggregation of the Aβ(1–42) monomer.
** Highest antiaggregative potency.

last variable contributes poorly to a linear correlation between activity and structure but was included due to its contribution to multiple relationships.

First, *backpropagation neural network* analysis [18, 28, 29] (BP) was applied. As model parameters, the following layers were used: (i) the input layer with linear transfer function and 3 nodes (physicochemical descriptors, Tab. I); the first hidden layer with sigmoidal transfer function and 3 nodes; and (iii) the output layer with sigmoidal transfer function and 1 node (percentage of formation of amyloid aggregation). All possible connections were analyzed. The sigmoidal backpropagation functions were solved by the nonlinear Levenberg-Marquardt algorithm. Optimization was achieved by (i) estimating the global error vector prior to adjusting weights, and (ii) updating successively the weights until convergence was reached. As goodness-of-fit criteria, the squared multiple correlation coefficient (R^2) was used and statistically tested.

The results are listed in Table II. Table III collects the theoretically calculated percentages of a formation of Aβ(1–42) peptide aggregation. The squared multiple correlation coefficient $R^2 = 0.956$ is statistically significant at the 5% level or less (the critical quantile of the maximum-likelihood criterion is $\Lambda = 0.563$). The statistical MASCA approach led to $R^2 = 0.716$, and partial-least squares regression yielded $R^2 = 0.677$. The backpropagation neural network gave better fitting statistics, therefore.

Second, the *generalized-regression genetic-neural network* model (GRNN) was used [29]. As parameters, the following layers were used: (i) the input layer with linear transfer function and 3 nodes (physicochemical descriptors,

TABLE II Network weights, current adjustment deltas, and goodness-of-fit criteria of the backpropagation (BP) network analysis

Layer	*Node*	*Connection*	*Weight*	*Weight delta*
2	1	1	4.50558	0.000039
2	1	2	−8.28399	−0.000055
2	1	3	4.86191	0.000030
2	2	1	−4.32409	−0.000002
2	2	2	−3.70402	−0.000002
2	2	3	−4.68957	−0.000003
2	3	1	−14.34891	−0.000070
2	3	2	−12.29154	−0.000053
2	3	3	11.72296	0.000063
3	1	1	−4.14275	0.000043
3	1	2	2.85004	−0.000003
3	1	3	4.01875	−0.000026
Standard deviation				4.151
Bias				0.035
Maximum error				11.669
Multiple correlation				0.978

TABLE III Recognition (training) set: results of backpropagation neural network (BP), generalized-regression genetic-neural network (GRNN), self-organizing map neural network (SOM): experimentally obtained [3] (Obtd.) and theoretically calculated (Calcd.) percentages of an inhibition of Aβ(1 – 42) aggregation after administration of peptide ligands

Compd.	*Obtd.*	*BP** *Calcd.*	*GRNN*** *Calcd.*	*SOM#* *Calcd.*
1	57.9	56.3	59.9	80.3
2	85.7	84.1	83.7	81.2
3	61.3	53.0	61.1	76.6
4	87.1	84.7	86.1	78.3
5	82.4	83.9	82.6	78.3
6	88.9	89.8	89.0	83.6
7	93.3	93.2	94.6	87.2
8	45.4	57.1	46.4	77.5
9	98.5	97.5	97.2	86.2
10	80.9	81.0	80.9	89.7
11	101.3	102.5	101.3	93.6
12	125.3	126.5	125.3	95.4
13	94.4	93.1	94.4	91.4

* "Second-best" model.
** Best model but it makes sometimes problems in prediction sets.
# No useful results.

Tab. I); the first hidden layer with sigmoidal transfer function and 3 nodes; and (iii) the output layer with direct transfer function and 1 node (percentage of formation of amyloid aggregation). All possible connections were analyzed. The following net parameters were used: generation run was 10, population size was 42, the minimum network training passes for each network were 20, the cut-off for network training passes was 20, the input and hidden neural node influence factors were 0, the limit on hidden neurons was 8; selection of relevant descriptors was performed by the top 50% surviving, refilling of the population was done by cloning the survivors, mating was performed by using the TailSwap technique (the system picks up a cut point and exchanges "genetic material" between the cut point and the end of the string of the "parents", essentially swapping tails). Mutations were performed using the random bit exchange technique at rate 25%.

The squared multiple correlation coefficient $R^2 = 0.9976$ is highly significant. Surprisingly, the second physicochemical descriptor was omitted (although the linear correlation coefficient between biological activity and electrostatic energy, $r = -0.598$, is significant; the critical quantile of Roy's largest root criterion is $\theta = 0.563$ at the 5% significance level). The theoretically calculated biological activity is listed in Table IV.

Third, the *self-organizing map network* [18] or Kohonen neural net (SOM) was applied (minimum passes = 20, maximum passes = 50). The squared

TABLE IV Prediction set: aggregation-inhibiting potency (Prdd) of geometry-optimized homologous peptide ligands (L-configuration) using the backpropagation neural network

Compd.	*Ligand*	E_{tot}	E_{el}	E_{hy}	*Prdd*
15B	KLVAF	21.8098	−30.4004	−9.799	78.2
15A	KLVFF	2.2444	−20.7144	−10.977	90.8
15F	KLVF	−4.1451	−20.0475	−12.193	102.1
15M	KHQKLVFF	−11.3533	−67.8978	−32.890	117.1
15O	KLVFFAE	−42.9722	−47.5658	−13.362	102.7
16A	LKFD	28.7781	14.5707	−9.967	87.5
16B	KFFD	20.0844	14.1319	−12.256	97.6
16C	LFFK	28.3777	19.6533	−8.331	88.9
16D	LFFDK	32.5738	19.8657	−16.863	60.5
16E	LFFKD	32.5864	20.8395	−16.168	62.6
17A	LPFFDY	45.0750	21.7598	−14.461	45.5*
17B	LPFFDR	25.5018	−1.8115	−23.804	62.2
18	Nicotine	24.5343	−3.5963	−1.143	70.9

* Highest predicted inhibition (1 − %) of the peptide ligands.

multiple correlation coefficient is only $R^2 = 0.5625$, and the theoretically calculated biological activity indicates also that SOM is not useful to analyze the data set (Tab. III). It is notable that the second physicochemical descriptor was also excluded from final network functions.

Table IV lists a number of peptide ligands with untested biological activity (prediction set). Unfortunately, SOM and GRNN cannot be applied because of the low goodness of recognition (SOM) and the known fact that GRNN produces sometimes "out-of-range errors" (related to the physicochemical descriptors). Therefore, BP is used to predict the percentages of formation of Aβ(1−42) aggregation (Tab. IV). Although the structure of nicotine is outside the structures of the training set, the effect of nicotine was also predicted for fun. Theoretically speaking, nicotine (compound 18) prevents β-sheet conformation. This agrees with experimental studies [30].

Clearly, the energies included into the quantitative structure-activity relationships (QSARs) depend on the geometry-optimized conformations of the peptide ligands. This is an advantage from the viewpoint of a 3D-QSAR analyst but a drawback from the viewpoint of an analyst who prefers traditional, simple, and clearly interpretable QSAR approaches basing on extrathermodynamic descriptors [31]. Nevertheless, the ligands listed in Tables I and IV are lead structure for subsequent research.

4. CONCLUSIONS AND FUTURE PERSPECTIVES

The central, apolar, and antigenic α-helical folding domain of Alzheimer's Aβ peptides is the molecular target of the binding of short anti-β-sheet

ligands. Peptide ligands are able to recognize and interact with this domain, preventing thus the polymerization of the monomeric, nontoxic Aβ peptides to neurotoxic, polymeric aggregates. It is found by artificial neural network approaches that the total (and electrostatic) energy of geometry-optimized conformations and the hydration energy of the ligands may simulate the antiaggregatory potency. The studied homologeous peptide ligands are probably lead structures for subsequent research.

However, further molecular manipulations are required to improve the pharmacokinetic properties. Beside the use of D-amino acids and conformationally restricted cyclic derivatives, the introduction of space linkers may lead to better pharmacokinetic properties. Other possibilities are probably the replacement of proline of the ligands by pseudoprolines (ΨPro, 2,2-dimethylated thiazolidines) and N-methylpyrrolidines.

References

[1] Simmonis, L. W., Becker, G. W., Brems, D. N., Brigham, E. F., Fuson, K. S., Li, W. Y., Lieberburg, I., May, P. C., Reydel, R. E., Tomaselli, K. J. and Wright, S. (1994). Secondary structure of amyloid β peptide correlates with neurotoxic activity *in vitro*, *Mol. Pharmacol.*, **45**, 373–379.

[2] Mager, P. P. (1998). Molecular simulation of the primary and secondary structures of the Aβ(1–42)-peptide of Alzheimer's disease, *Med. Res. Rev.*, **18**, 403–430.

[3] Soto, C., Kindy, M. S., Baumann, M. and Frangione, B. (1996). Inhibition of Alzheimer's amyloidosis by peptides that prevent β-sheet conformation, *Biochem. Biophys. Res. Commun.*, **226**, 672–680.

[4] Soto, C. and Frangione, B., Inhibition of Alzheimer's amyloidogenesis by anti-β-sheet peptides, In: *Alzheimer's Disease*: *Biology*, *Diagnosis and Therapeutics*, Iqbal, K., Winblad, B., Nishimura, T., Takeda, M. and Wisniewski, H. M. Eds., Wiley, New York, 1997, pp. 711–716.

[5] Soto, C., Sigurdsson, E. M., Morelli, L., Kumar, R. A., Castano, E. M. and Frangione, B. (1998). β-Sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: Implications for Alzheimer's therapy, *Nature Med.*, **4**, 822–826.

[6] Ghanta, J., Shen, C. L., Kiessling, L. L. and Murphy, R. M. (1996). Strategy for designing inhibitors of β-amyloid toxicity, *J. Biol. Chem.*, **22**, 29525–29528.

[7] Tjernberg, L. O., Näslund, J., Lindqvist, F., Johannson, J., Karlström, A. R., Thyberg, J., Terenius, L. and Nordstedt, C. (1996). Arrest of β-amyloid fibril formation by a pentapeptide ligand, *J. Biol. Chem.*, **271**, 8545–8548.

[8] Tjernberg, L. O., Lilliehöök, C., Callaways, D. J. E., Näslund, J., Hahne, S., Thyberg, J., Terenius, L. and Nordstedt, C. (1997). Controlling amyloid β-peptide fibril formation with protease-stable ligands, *J. Biol. Chem.*, **272**, 12601–12605.

[9] Allinger, N. L. and Yan, L. (1993). Molecular mechanics (MM3). Calculation of furan, vinyl, ethers, and related compounds, *J. Am. Chem. Soc.*, **115**, 11918–11925.

[10] Tai, J. C., Yang, L. and Allinger, N. L. (1993). Molecular mechanics (MM3). Calculation on nitrogen-containing aromatic heterocyclics, *J. Am. Chem. Soc.*, **115**, 11906–11917.

[11] Gasteiger, J. and Marsili, M. (1980). Iterative partial equalization of orbital electronegativity – a rapid access to atomic charges, *Tetrahedron*, **36**, 3219–3288.

[12] Frey, R. F., Coffin, J., Newton, S. Q., Ramek, M., Cheng, V. K. W., Momany, F. A. and Schäfer, L. (1992). Importance of correlation-gradient geometry optimization for molecular conformational analyses, *J. Am. Chem. Soc.*, **114**, 5369–5377.

[13] Kyte, J. and Doolite, R. F. (1982). Simple method for displaying the hydrophobic character of a protein, *J. Mol. Biol.*, **157**, 105–132.
[14] Dixon, S. L. and Jurs, P. C. (1993). Estimation of pK_a for organic oxyacids using calculated atomic charges, *J. Comput. Chem.*, **14**, 1460–1467.
[15] Meylan, W. M. and Howard, P. H. (1995). Atom/fragment contribution method for estimating octanol-water partition coefficients, *J. Pharm. Sci.*, **84**, 83–92.
[16] Eisenhaber, F., Argos, P., Lijnzaad, P., Sander, C. and Scharf, M. (1995). The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies, *J. Comput. Chem.*, **16**, 273–284.
[17] Stout, J. M. and Dykstra, C. E. (1995). Static dipole polarizabilities and a predictive model, *J. Am. Chem. Soc.*, **117**, 5127–5132.
[18] Manallack, D. T. and Livingstone, D. J. (1999). Neural networks in drug discovery: have they lived up to their promise?, *Eur. J. Med. Chem.*, **34**, 195–208.
[19] Mager, P. P., *Design Statistics in Pharmacochemistry*, Wiley, New York, 1991.
[20] Jiang, H., Burdick, D., Glabe, C. G., Cotman, C. W., Tenner, A. J. (1994). β-amyloid activates complement by binding to a specific region of the collagen-like domain of the C1q A chain, *J. Immunol.*, **152**, 5050–5059.
[21] Terzi, E., Holzemann, G. and Seelig, J. (1994). Alzheimer β-amyloid peptide 25–35: electrostatic interactions with phospholipid membranes, *Biochemistry*, **33**, 7431–7441.
[22] Mason, R. P., Estermeier, J. D., Kelley, J. S. and Mason, P. E. (1996). Alzheimer's disease amyloid β-peptide 25–35 is localized in the membrane hydrocarbon core: X-ray diffraction analysis, *Biochem. Biophys. Res. Commun.*, **222**, 78–82.
[23] Mager, P. P. and Fischer, K., Simulation of the lipophilic and antigenic cores of the Aβ(1–42) peptide of Alzheimer's disease, *Mol. Simul.*, in press.
[24] Hilbich, C., Kisters-Wolke, B., Masters, C. L. and Reed, J. (1991). Aggregation and secondary structure of synthetic amyloid βA4 peptide of Alzheimer's disease, *J. Mol. Biol.*, **218**, 149–163.
[25] Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C. L. and Beyreuther, K. (1992). Substitutions of hydrophobic amino acids reduce the amyloidgenicity of Alzheimer's disease βA4 peptides, *J. Mol. Biol.*, **228**, 460–473.
[26] Barrow, C. J., Kenny, P. T. M., Ykikazu, A. and Zagorski, M. G. (1992). Solution conformations and aggregational properties of synthetic amyloid β-peptides of Alzheimer's disease, *J. Mol. Biol.*, **225**, 1075–1093.
[27] Chaney, M. O., Webster, S. D., Kuo, Y.-M. and Roher, A. E. (1998). Molecular modeling of the Aβ1–42 peptide from Alzheimer's disease, *Protein Engineering*, **11**, 761–767.
[28] Benzinger, T. L. S., Gregory, D. M., Burkoth, T. S., Miller-Auer, H., Lynn, D. G., Botto, R. E. and Meredith, S. C. (1998). Propagating structure of Alzheimer's β-amyloid (10–35) is parallel β-sheet with residues in exact register, *Proc. Natl. Acad. Sci. USA*, **95**, 13407–13412.
[29] Mager, P. P. (1998). Neural network approaches applied to selective A_{2a} adenosine receptor agonists, *Med. Chem. Res.*, **8**, 277–290.
[30] Mager, P. P. and Reinhardt, R. (1999). A comparison of backpropagation and generalized-regression genetic-neural network models, *Drug Design and Discovery*, **16**, 49–53.
[31] Salomon, A. R., Marcinowski, K. J., Friedland, R. P. and Zagorski, M. G. (1996). Nicotine inhibits amyloid formation by the β-peptide, *J. Am. Chem. Soc.*, **35**, 13568–13578.
[32] Fujita, T., The Extrathermodynamic Approach to Drug Design, In: *Comprehensive Medicinal Chemistry*, Hansch, C. and Taylor, J. B. Eds., Pergamon, Oxford, 1990, pp. 497–560.