Pharmaceutical Research, Vol. 5, No. 7, 1988

Research Article

Searching for the Endogenous Benzodiazepine Using the Graph Theoretical Approach[1]

Rodney S. Markin[2] and Wallace J. Murray[3,4]

Received December 10, 1987; accepted January 29, 1988

The graph theoretical indices of several compounds with reported benzodiazepine receptor binding affinities were calculated. Our results demonstrate a structural similarity among diazepam, triazolam, and the β-carboline nucleus and a structural dissimilarity to the purines and nicotinamide. This result correlates with their respective binding affinities. Using the graph theoretical indices as structural descriptors of the benzodiazepines and the significant ligands of the β-carbolines, a search for peptide sequences as potential ligands was explored. Single amino acids through pentapeptides with all possible amino acid substitutions and chemical modifications were calculated. The peptides generated were subjected to graph theoretical analysis, and their indices were compared to those of the benzodiazepines. Comparisons resulted in seven dipeptides and six tripeptides that are topologically similar to the benzodiazepines and β-carbolines. The dipeptides are histidine- or tryptophan-containing compounds with pyroglutamine, phenylalanine, and tyrosine residues in the second position. The tripeptides have two aromatic amino acid residues and a pyroglutamine or glycyl terminal residue. These structures are promising candidates because (1) they are structurally (topologically) similar to the benzodiazepines, represented by diazepam and triazolam, and to the β-carbolines; and (2) they are sequences that may reasonably form a part of a larger peptide or that may be formed metabolically by proteolysis.

KEY WORDS: topological approach; endogenous benzodiazepine; QSAR, similarity.

INTRODUCTION

The 1,4-benzodiazepines, of which diazepam (1) and triazolam (2) (Scheme I) represent prototypical structures, are used for their anxiolytic, sedative-hypnotic, antiepileptic, and muscle relaxant properties. The mechanism of action of the benzodiazepines has been closely linked to that of γ-aminobutyric acid (GABA), and the specific binding sites that have been demonstrated for the benzodiazepines seem to be coupled in a supramolecular complex with GABA receptors (1,2).

The specific binding site for the benzodiazepines has been called the "benzodiazepine receptor," and there have been many attempts to identify an endogenous ligand. In addition to other classes of drugs, several substances have been isolated from biological material and their interactions with the benzodiazepine receptor have been characterized (3). They include all of the L-amino acids; all known neurotransmitters, including the enkephalins; various coenzymes, including nicotinamide adenine dinucleotide (NAD^+); purine and pyrimidine nucleotides and nucleosides and their respective bases; intestinal and other visceral peptides; the prostaglandins; and hormones—all of which exhibited no significant binding to the benzodiazepine receptor. Most attention has been focused on purines such as inosine, hypoxanthine, and adenosine, which antagonize the interaction of diazepam with the benzodiazepine receptor; nicotinamide, which has five times more affinity than inosine or hypoxanthine for the receptor; and the β-carbolines. The β-carbolines (3) (Scheme II) are a family of tricyclic, aromatic hydrocarbons that have a higher affinity for the benzodiazepine receptor than does diazepam. They have been isolated from large quantities of human urine, and homogenized brains extracted with ethanol have been shown to have β-carbolines present. The A and B rings of the β-carbolines are formed by condensing the side chain of tryptophan with the indole nucleus.

Another promising candidate is a peptide termed DBI (diazepam binding inhibitor) which has been isolated from non-GABAergic neurons in rat cerebral cortex (4,5). DBI consists of 104 amino acid residues, and furthermore, proteolytic digestion of DBI gives peptide fragments that exhibit activity similar to that of the parent peptide. These were sufficiently stimulating data for us to assume that the endogenous ligand is a peptide and, on the basis of this assumption, to begin a "search" for an endogenous ligand or for peptide sequences that were similar in structure to the

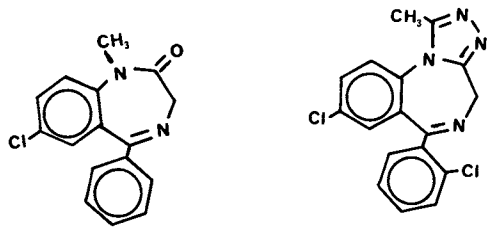
[1] Presented in part at the American Association of Pharmaceutical Scientists Meeting, November 1986, Washington, D.C.

[2] Department of Pathology and Microbiology, College of Medicine, University of Nebraska Medical Center, Omaha, Nebraska 68105-1065.

[3] Department of Pharmaceutical Sciences, College of Pharmacy, University of Nebraska Medical Center, Omaha, Nebraska 68105-1065.

[4] To whom correspondence should be addressed.

0724-8741/88/0700-0408$06.00/0



Scheme I

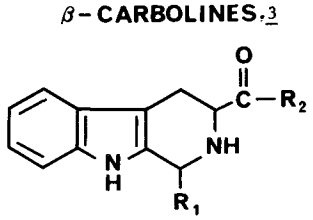
benzodiazepines and the β-carbolines. "Synthesis" of the peptides consisted of constructing molecular structures of hypothesized peptides and determining by the graph theoretical approach their similarity to the benzodiazepines and β-carbolines.

METHODS

The graph theoretical approach is a molecular topological approach. Molecular topology has been described as a mathematical description of molecular structure (6), and the various applications of topological methods can be found in Ref. 7. The details of applying the graph theoretical approach to arrive at a numerical representation of molecular structure have been described previously (8–13).

The underlying premise is that there is more than one type of molecule with the desired biological activity, e.g., drug molecules are not endogenous compounds for which endogenous receptors were formed to fit. Since a computer database is used, the method is useful in screening large numbers of molecular candidates. Implicit in the approach is that structurally similar compounds will have similar properties. The purpose is to eliminate random choice of structures to test, although it is arbitrary which sets of structures are used for comparison.

The graph theoretical approach uses the molecular or skeletal graph of a compound and enumerates all self-avoiding paths. A self-avoiding path is a sequence of edges in which no atom or bond is counted more than once. The algorithm does not include hydrogen atoms in the graph structure (hydrogen-suppressed graph), and no distinction is made between skeletal atoms, e.g., C, N, and O. The general protocol is as follows:



	R_1	R_2
A	—H	—OH
B	—H	—OCH_3
C	—H	—OC_2H_5
D	—CH_3	—OH
E	—CH_3	—OC_2H_5

Scheme II

(1) number all vertices,
(2) construct a table of path lengths vs atoms, and
(3) total the number of paths of each length for the entire molecule.

The method is illustrated for a series of hexane isomers (Table I).

For this study, a database of the 22 naturally occurring amino acids was constructed that enabled us to "synthesize," or hook together, permutations of all the possible peptides through pentapeptides. The path sequences of all synthesized peptides was calculated and structures were compared by the Euclidean distance method, which quantitates similarity (8).

In this method the path sequences are considered to be coordinates in an n-dimensional Euclidean space, described earlier by Randic and Wilkins (8), where n is the maximum path length in the set of compounds. Each molecule occupies a point in that space, and the similarity between molecules can be calculated by a mathematical difference. Equation (1) summarizes the relationship between two hypothetical sequences, a and b:

$$D_{ab} = \Sigma[(a_i - b_i)^2]^{1/2} \quad (1)$$

The more similar a compound is to the reference compound, the "closer" it is in terms of its Euclidean distance, D_{ab}. This measure may be used to rank order a given property of each compound against a reference compound.

The path sequences for the active β-carbolines, several nucleosides, and nicotinamide were also calculated for similarity comparison.

The approach has certain advantages.

(1) The technique is reduced to a computer program which rapidly and inexpensively computes the molecular path sequences and the requisite comparisons and/or statistical evaluations.
(2) No chemical synthesis or physical measurement is required.
(3) Although three-dimensional structure is not specifically addressed by the calculation, it is to some extent a function of connectivity.
(4) The ease of quantitation of similarity makes possible rational choices of potential structures with desired properties in the absence of any information other than the structures themselves.

This method has been previously used in a study of enkephalins in which the graph theoretical indices of the 10 most active enkephalins were ordered and compared to their

Table I. Path Enumerations for Hexane Isomers

	Path length					
	0	1	2	3	4	5
N-Hexane	6	5	4	3	2	1
2-Methylhexane	6	5	5	4	1	
2,3-Dimethylhexane	6	5	6	4		
2,2-Dimethylhexane	6	4	7	3		

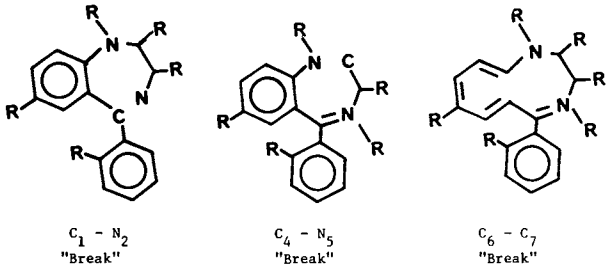


Fig. 1. Truncated graphs of benzodiazepines showing various bond breaks used in the calculation of the graph theoretical indices.

binding affinities for the opiate receptor (9). A direct correlation between the ordering of the graph theoretical indices and binding affinities was found.

RESULTS

In this study the paths and Euclidean distances of two prototypical benzodiazepines, diazepam (1) and triazolam (3), were calculated for the total molecule and their truncated fragments, as shown in Fig. 1. In another study, we found that truncating the molecule, i.e., "breaking" selected bonds in the benzodiazepine nucleus, yielded some interesting comparisons (15). The breaking of the bonds in the diazepine ring effectively decreases the number of "cycles" calculated by the algorithm. In essence, it makes the molecule more "linear." Looking at the molecule in a linear fashion leads one to postulate a linear structure, such as a peptide, which might assume a shape similar to the benzodiazepine with appropriate functional groups.

Graph theoretical calculations of several compounds with reported benzodiazepine receptor binding affinities were carried out (16). The Euclidean distances of the β-carbolines, several purines, inosine, hypoxanthine, adenine, and nicotinamide were compared, and the results of these calculations are summarized in Table II. The results showed a high degree of similarity among the benzodiazepines, as represented by diazepam and triazolam, the truncated diazepam (C_1–N_2 break; Fig. 1), and the β-carbolines. The calculations for the other compounds showed them to be structurally dissimilar. Furthermore, these results correlated well with their respective binding affinities.

The fact that the truncated or broken benzodiazepines correlated so well structurally with the β-carbolines, with the receptor binding affinities and with pharmacokinetic properties, and the biological data observed with the peptides, DBI (4,5) and nepenthin (14), another peptide that binds to the benzodiazepine receptor, caused us to hypothesize that the endogenous benzodiazepine may be a peptide. Figure 2 shows how these comparisons might be made using various amino acid residues put together in hypothesized peptide links. Using the graph theoretical indices as structural descriptors of the benzodiazepines and β-carbolines, a search for peptides using computer-generated peptide sequences as potential ligands was attempted. The computer was programmed to produce peptide sequences from single amino acids to dipeptides through pentapeptides with all possible amino acid residues and chemical modifications. Figure 3 shows how individual amino acid structures stored in computer memory were added together in matrix form for calculation by the program.

The peptides generated were subjected to graph theoretical analyses and then compared to the graph theoretical indices of diazepam and triazolam, as benzodiazepine prototypes, and to the β-carbolines. The similarity indices for those peptides that most closely resemble the structures of these prototypes are given in Table III. Those compounds with a calculated Euclidean distance, D_{ab}, of less than 20 were arbitrarily designated as having significant similarity.

Table II. Euclidean Distance Comparisons of the β-Carbolines with Diazepam, Triazolam, and Truncated Diazepam[a]

β-Carboline	Diazepam	Triazolam	Diazepam with C_1–N_2 break
A	5.13	18.56	4.58
B	5.02	18.49	4.41
C	5.01	18.48	4.38
D	4.13	18.36	3.98
E	4.76	18.38	4.28

[a] Diazepam with a C_1–N_2 break (see text). In this comparison, the indicated β-carboline is used as the reference compound, i.e., the origin in the Euclidean coordinate system.

DISCUSSION

In another study by us, we compared some pharmacokinetic parameters to the graph theoretical paths of a set of 17 benzodiazepines (15). We found that truncating the molecule, i.e., "breaking" selected bonds in the benzodiazepine nucleus, yielded structures in which one parameter, the receptor binding affinity of the benzodiazepines, showed a significant increase in correlation: from a correlation coefficient of 0.703 before truncation to one of 0.916 after truncation. The breaking of the bonds in the diazepine ring effectively decreases the number of cycles marked by the algorithm, making the molecule more "linear." Looking at the molecule in a linear fashion leads one to postulate a linear structure, such as a peptide. In addition, the truncated or broken benzodiazepines were structurally similar to the β-carbolines, using Euclidean distance as a quantitative measure of similarity (Table II).

Following these initial studies showing that the graph theoretical indices are useful structural descriptors of the benzodiazepines and β-carbolines, a search for peptides using computer-generated peptide sequences as potential ligands was attempted. The computer was programmed to produce peptide sequences from single amino acids to dipeptides through pentapeptides with all possible amino acid residues and chemical modifications. The Euclidean distance comparison of peptide sequences of greater than six amino acid residues exhibited D_{ab} values of greater than 20, the D_{ab} value selected to discriminate between similar and dissimilar structures.

The peptides generated were subjected to graph theoretical analyses and then compared to the graph theoretical indices of diazepam and triazolam, as benzodiazepine prototypes, and to the β-carbolines. The similarity indices for those peptides that most closely resemble the structures of these prototypes are given in Table III. The comparison

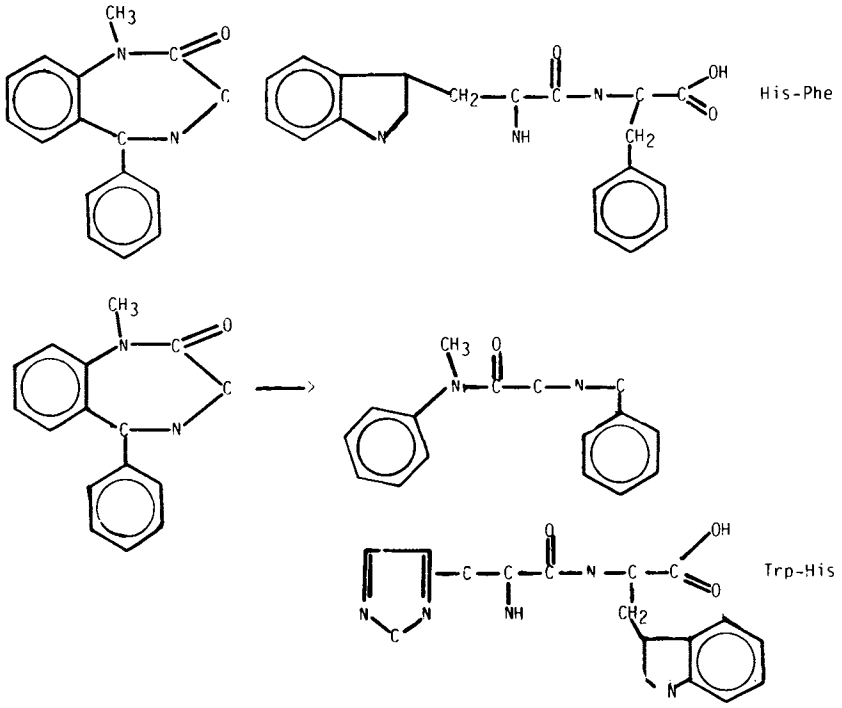


Fig. 2. Conversion of a benzodiazepine represented by diazepam into a peptide sequence that is topologically similar in structure.

shows that several di- and tripeptides are topologically similar to the benzodiazepines chosen and to the β-carbolines. There were, of course, thousands of compounds which were dissimilar to the compounds listed. The dipeptides with predicted benzodiazepine receptor activity are histidine- or tryptophan-containing compounds with pyroglutamine, phenylalanine, tyrosine, or histidine residues in the second position. In addition, several tripeptides where shown to have structural or topological similarity to the benzodiazepines and β-carbolines. These peptides have at least two ring structures donated by separate amino acid residues and pyroglutamine or glycyl terminal residue.

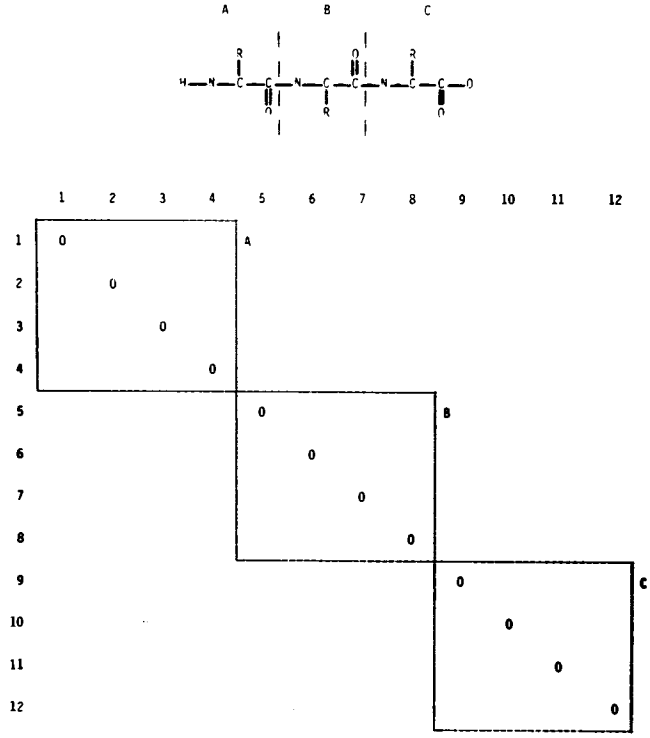


Fig. 3. Depiction of the technique used to "synthesize" peptide sequences by computer for subsequent calculation of their graph theoretical indices.

SUMMARY

Several peptides are calculated by the graph theoretical approach to be structurally and topologically similar to the benzodiazepines and β-carbolines, and therefore, they are

Table III. Similarity Indices (Euclidean Distance Comparison) of Peptides with Diazepam, Triazolam, and β-Carboline (D)[a]

Peptide	Diazepam	Triazolam	β-Carboline (D)
His-Tyr	15.73	18.84	11.48
His-Phe	15.71	18.81	11.42
His-Trp	6.16	8.01	6.48
Trp-His	5.86	7.51	5.23
Trp-PyroGlu	5.47	7.35	5.43
Trp-Phe	5.72	7.70	6.12
Trp-Tyr	5.74	7.72	6.15
Phe-Tyr-PyroGlu	17.76	8.76	9.12
Tyr-Phe-PyroGlu	17.79	8.81	9.21
His-Pro-Glu	15.45	18.71	10.93
His-Phe-PyroGlu	16.61	9.42	9.93
His-Tyr-PyroGlu	16.63	9.44	9.94
Tyr-Phe-Gly	6.73	8.62	6.91

[a] The indicated di- or tripeptide in Column 1 is used as the reference compound. A Euclidean distance, D_{ab} = 20, was arbitrarily selected as the cutoff between "similar" and "dissimilar."

predicted to have a high affinity for the benzodiazepine receptor. These structures appear to be promising possibilities for three reasons: (i) they are structurally (topologically) similar; (ii) they are peptide sequences that may reasonably be found as part of a larger protein molecule which functionally interacts with the receptor; and (iii) they are residues that may be formed from proteolysis of higher peptide chains and, as such, functionally interact with the receptor. Finally, the synthesis of the small peptide fragments predicted to have similarity to the benzodiazepines and β-carbolines and the determination of their binding affinities to the benzodiazepine receptor can be easily accomplished. Such studies will give us an indication of the usefulness of the graph theoretical approach in predicting and quantitating similarity in structure.

REFERENCES

1. S. Mazzari, M. Massotti, A. Guidotti, and E. Costa. In E. Costa, G. DiChiara, and G. L. Gessa (eds.), *GABA and Benzodiazepine Receptors,* Raven Press, New York, 1981, pp. 1–8.
2. J. F. Tallman, P. Mallorga, J. W. Thomas, and D. W. Gallager. In E. Costa, G. DiChiara, and G. L. Gessa (eds.), *GABA and Benzodiazepine Receptors,* Raven Press, New York, 1981, pp. 9–18.
3. C. Braestrup, M. Nielsen, H. Skovbjerg, and O. Gredal. In E. Costa, G. DiChiara, and G. L. Gessa (eds.), *GABA and Benzodiazepine Receptors,* Raven Press, New York, 1981, pp. 147–155.
4. E. Costa and A. Guidotti. *Biochem. Pharmacol.* 34:3399–3403 (1985).
5. H. Alho, E. Costa, P. Ferrero, M. Fujimoto, D. Cosenza-Murphy, and A. Guidotti. *Science* 229:179–182 (1985).
6. H. E. Simmons and R. E. Merrifield. *Proc. Natl. Acad. Sci.* 74:2616–2623 (1977).
7. L. B. Kier and L. H. Hall. *Molecular Connectivity in Structure-Activity Analysis,* Research Studies Press, John Wiley and Sons, New York, 1986.
8. M. Randic and C. L. Wilkins. *J. Chem. Inf. Comp. Sci.* 19:31–37 (1979).
9. C. L. Wilkins, M. Randic, S. M. Shuster, R. S. Markin, S. Steiner, and L. Dorgan. *Ana. Chim. Acta.* 133:637–645 (1981).
10. M. Randic and C. L. Wilkins. *J. Phys. Chem.* 83:1525–1531 (1979).
11. M. Randic and C. L. Wilkins. *Int. Symp. Quant. Chem. Quant. Biol.* 6:55–71 (1979).
12. C. L. Wilkins and M. Randic. *Theor. Chem. Acta.* 58:45–51 (1980).
13. M. Randic, G. M. Brissey, R. B. Spencer, and C. L. Wilkins. *Comput. Chem.* 4:27–39 (1979).
14. J. H. Woolf and J. C. Nixon. *Biochemistry* 20:4263–4269.
15. R. S. Markin, W. J. Murray, and H. Boxenbaum. *Pharm. Res.* 5:201–208 (1988).
16. R. S. Markin and W. J. Murray. Manuscript in preparation.