

INFORMATIZAÇÃO DA BUSCA DE INFORMAÇÕES PARA A DETERMINAÇÃO ESTRUTURAL DE PRODUTOS NATURAIS

Alphonse Kelecom e Valéria L. Teixeira

Departamento de Biologia Geral UFF; C.P. 100.183; 24000 – Niterói (RJ)

(Recebido em 24/2/88; Cópia revisada em 15/9/88)

ABSTRACT

A data base on marine diterpenes has been built up using the "TOTALWORKS" program. This data base includes structural information which allows rapid identification of known substances or characterization of a given metabolite as a new compound. This methodology can be extended easily to other classes of natural products.

RESUMO

Um banco de dados sobre diterpenos marinhos foi elaborado utilizando os recursos do programa "TOTALWORKS". Este banco de dados contém informações estruturais que permitem a identificação imediata de substâncias conhecidas ou a caracterização de um metabólito como sendo original. A metodologia aqui apresentada pode ser facilmente aplicada a outras classes de produtos naturais.

INTRODUÇÃO

A caracterização imediata de um metabólito recém isolado, como substância original ou não, é fator de eficiência em pesquisa de produtos naturais. Esta caracterização, baseada numa análise espectroscópica preliminar (EM, IV, UV e RMN ^1H e ^{13}C), deve se fazer com alto grau de certeza e com rapidez. Uma metodologia informatizada de pesquisa numa biblioteca de padrões parece ser a opção mais certa para executar esta tarefa.

Existem, descritos na literatura, vários programas de computador que foram desenvolvidos para fins de determinação estrutural. Estes programas não atendem diretamente à nossa finalidade, além de necessitar, em geral, de computadores de grande porte. Por outro lado, o químico de produtos naturais é frequentemente especialista de uma determinada classe de metabólitos. Ele necessita em geral de um banco de dados que reúna aquelas informações ligadas diretamente a sua área de atuação e não de sistemas informatizados mais sofisticados.

O presente trabalho visa a descrever uma nova rotina seguida na determinação estrutural de diterpenos isolados de algas pardas da família Dictyotaceae.

MATERIAL E MÉTODOS

O Equipamento

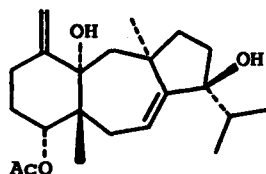
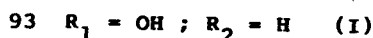
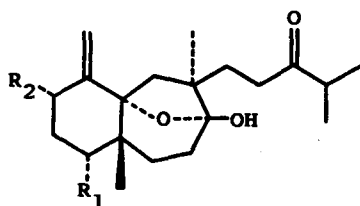
Alguns critérios básicos nortearam a informatização desta rotina de determinação estrutural. Primeiro, foi escolhido um "hardware" de fabricação nacional com grande capacidade de memória RAM, acentuação da língua portuguesa e com um teclado numérico independente, uma vez que se irá manipular muitos números. Segundo, já que estudantes de vários níveis (doutorandos, mestrados e até estagiários) devem utilizar o sistema, é bom escolher um software integrado, em português, e caso necessário adaptá-lo para nossos fins. Optou-se pelo computador TK 3000//e da Microdigital, equipado com uma placa de expansão de memória de 512K, e pelo programa "TOTALWORKS" da Royal Digital (versões brasileiras do Apple IIe e do Appleworks respectivamente).

Metodologia

Foi elaborado um banco de dados que contém, de maneira codificada, informações sobre as estruturas dos diterpenos fichados. A busca das informações se faz utilizando os recursos existentes no programa "TOTALWORKS" descritos detalhadamente no manual do programa. Estes recursos se mostraram amplamente suficientes para nossas necessidades.

Figura 1. Exemplo de ficha do banco de dados

nome	: isolinearol
produto número	: 93
fórmula molecular	: $\text{C}_{20}\text{H}_{32}\text{O}_4$
tipo de esqueleto	: secodolastano (grupo 2)
estrutura	: $2\text{C}(6, 7) - 1\text{L} = (1\text{ exo}) - 4\text{Me}$ $(2\text{s}, 2\text{d}) - 2\text{OH}(1\text{ sec}, 1\text{ terc}) -$ $- 1\text{CO} / - 1\text{ROR}$
posições oxigenadas	: 2, 6, 10, 14
índices químicos-sistemáticos	: I. O. (1): -1.200 E.(1): 0.20 I. O. (2): -1.200 E.(2): 0.50
ocorrência	: <i>Dictyoya cervicomis</i> (1)
distrib. geográfica	: Atlântico (1)
referências	: (1) <i>J. Nat. Prod.</i> (1986) 49, 570.



41 (II)

O Banco de Dados

Foram reunidos num banco de dados mais de 330 diterpenos de algas pardas (Phaeophyta) e de falsos corais (Octocorallia). Cada ficha reúne informações químicas, biológicas e quimiosistemáticas como aparece na Figura 1 abaixo.

Tabela 1. Codificação das estruturas.

Grupo	Código	Observações
ciclos	$xC(M, N, P, \dots)$	x : número total de ciclos hidrocarbonados M, N, P : anéis de M, N, P... membros
lig. duplas	$xL = (y_{exo}, z_{di}, v_{tri}, w_{tetra})$	x : número total de ligações duplas (L =) $x = y + z + v + w$ exo, di, tri, tetra : natureza da L=
metilas	$xMe(y_s, z_d, v_t, w_ =)$	x : número total de grupos metila $x = y + z + v + w$ s, d, t: sinais simples, duplo, triplo em $RMN^{-1}H$ = : metilas em ligação dupla
funções oxigenadas: hidroxila	$xOH(y_{prim}, z_{sec}, v_{terc})$	x : número total de hidroxilas $x = y + z + v$ prim, sec, ...: primária, secundária...
acetoxi	$xOAc(y_{prim}, z_{sec}, v_{terc})$	como para hidroxilas
cetonas	$xCO/$	x : número total da função
aldeídos	$xCHO$	x : ídem
ácidos	$xRCOOH$	x : ídem; R : resíduo do esqueleto
ésteres	$xRCOOQ$	x, R : ídem; Q : resíduo álcool
lactonas	$xRCOOR$	x, R : ídem
éter linear	$xROQ$	x, R, Q : ídem
éter cíclico	$xROR$	x, R : ídem
éter metílico	$xOMe$	x : ídem

Da codificação das estruturas

O campo "estrutura" deste banco de dados merece maiores comentários. Resumidamente, trata-se de uma codificação das estruturas feita de maneira a poder ser manipulada pelo programa "TOTALWORKS". A codificação das estruturas dos diterpenos não visa a descrever por completo a estrutura mas sim a arquivar os elementos estruturais facilmente reconhecíveis a partir de uma análise espectroscópica preliminar. Isto compreende os ciclos, as ligações duplas, os grupos metila e as funções oxigenadas. Outras funções poderão no futuro ser adicionadas caso sejam necessárias.

CICLOS:

- número e tamanho dos ciclos (exceto os heterociclos)
exemplo: isolinearol (93) tem 2 ciclos um de 6 outro de 7 membros, o que é codificado $2C(6, 7)$

LIGAÇÕES DUPLAS:

- número e tipo de ligações duplas
exemplo: isolinearol (93) possui apenas uma ligação dupla exometilênica codificada $1L = (1_{exo})$

METILAS:

- número de grupos metila
- multiplicidade em RMN^1H ou metila em ligação dupla

exemplo: no isolinearol (93) existem 4 metilas, 2 em carbono quaternário e 2 em carbono terciário, o que é codificado 4Me(2s, 2d)

FUNÇÕES OXIGENADAS:

- número, tipo e natureza das funções oxigenadas
exemplo: isolinearol (93) contém uma hidroxila secundária, uma terciária, uma carbonila e um éter cíclico, o que é codificado
2OH(1sec, 1terc) – 1CO/–1ROR

Assim, o campo “estrutura” da ficha representada na Figura 1 é apenas a justaposição das codificações acima mencionadas. A Tabela 1 resume os vários códigos.

RESULTADOS

Esta metodologia é ilustrada, a seguir, para dois dos diterpenos isolados da alga parda *Dictyota cervicornis* durante a tese de mestrado de um dos autores (VLT).

Todos os diterpenos isolados desta alga apresentaram, em RMN ¹H, quatro grupos metila, sendo dois aparecendo como sinais simples e dois como sinais duplos, o que foi codificado: 4Me(2s, 2d). Utilizando os recursos de busca do programa “TOTALWORKS”, foi procurado, no campo “estrutura” do banco de dados geral, o elemento estrutural acima mencionado. Dos c.a. de 330 diterpenos arquivados foram selecionados apenas 21, sendo 18 dolastanos, 1 seco-dolastano, 1 asbestinano e 1 eunicellano.

DITERPENOS I:

O espectro no IV do diterpeno I indica a presença de grupos OH (3520 cm⁻¹) e de carbonila (1715 cm⁻¹). Buscando, entre as 21 entradas acima referidas, o elemento estrutural CO (função carbonilada diferente de acetato ou de aldeído), selecionamos apenas o produto 92 (Tabela 2).

Tabela 2. Diterpenos com 4Me(2s, 2d) e CO”

prod	fórmula mol.	tipo de esqueleto	ocorrência	ref.
92	C ₂₀ H ₃₂ O ₄	seco-dolastano	<i>D. linearis</i>	1

Esta dupla seleção retém o único diterpeno de esqueleto seco-dolastano conhecido até então. Embora tendo características espectrais muito próximas de linearol (92), o diterpeno I difere de 92 pela posição do hidrogênio carbinólico (δ 3,57) mais desprotegido em 92 (δ 4,24). Portanto o diterpeno I é um metabólito novo o qual foi identificado a isolinearol (93)².

DITERPENOS II:

O espectro de IV de II mostra absorção de grupos OH e OAc em 3500 cm⁻¹ e 1745 cm⁻¹ respectivamente. O espectro de RMN ¹H indica que o OAc é secundário, de-

finindo o elemento estrutural “1 OAc (1sec)”. Este espectro, além das metilas já mencionadas acima, ainda revela a presença de um grupo exometilênico e de um hidrogênio olefínico definindo assim o elemento estrutural 2L = (1exo, 1tri). Porém, como o espectro de RMN ¹³C de II não foi obtido, a existência de uma ligação dupla tetrasubstituída não pode ser descartada com absoluta segurança. A condição de seleção das ligações duplas deve ser então apenas: “1exo e 1tri”, i.e., apenas as informações absolutamente certas. Introduzindo as condições de busca “4Me(2s, 2d), 1 OAc(1sec) e 1tri” foi selecionando o produto da Tabela 3 idêntico espectroscopicamente ao diterpeno II identificado então de maneira rápida e não equivocada a 41².

Tabela 3. Seleção “4Me(2s, 2d), 1 OAc (1sec) e 1tri”

prod	fórmula mol.	esqueleto	ocorrência	ref.
41	C ₂₂ H ₃₄ O ₄	dolastano	<i>Dictyota sp.2</i> <i>D. divaricata</i> <i>D. linearis</i>	3 4 4

CONCLUSÕES

A busca na biblioteca de padrões dos dois diterpenos relatados neste trabalho não demorou mais de alguns minutos, o que indica a rapidez desta abordagem. O diterpeno I foi apontado como sendo novo ao passo que II já era conhecido. É preciso destacar que estas informações foram obtidas usando apenas os espectros de RMN ¹H e de IV, ou seja respostas seguras podem ser obtidas com informações espectrais limitadas. O grau de confiança nas respostas de busca depende exclusivamente de utilizar como regras de seleção apenas aquelas informações absolutamente certas e depende também de reunir informações exatas no banco de dados.

É claro que esta metodologia pode ser utilizada para todas as classes de produtos naturais, bastando apenas elaborar códigos adequados para as funções não citadas aqui.

Se este sistema ainda precisa de aprimoramentos, ele não deixa de se revelar extremamente útil na busca rápida e segura de informações.

AGRADECIMENTOS

Agradecemos o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) que financiou a compra do microcomputador e as pesquisas fitoquímicas (Processos 40.2543-85/QU, 40.2221-86/QU e 40.6341-87.7).

REFERÊNCIAS

- Ochi, M.; Miura, I., Tokoroyama, T., *J. Chem. Soc., Chem. Comm.* (1981) 100.