BIBLIOTECA

INSTITUTO DE QUÍMICA
UNIVERSIDADE DE SÃO PAULO

BIBLIOTECA

INSTITUTO DE QUÍMICA

Universidade de São Paulo

20637

UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE QUÍMICA

ESTERÓIDES: ANÁLISE ESPECTRAL, GERAÇÃO DE ESTRUTURAS E SIMULAÇÃO DE DADOS DE RMN ^{13}C

Marcelo José Pena Ferreira

Tese de Doutorado

Prof. Dr. Vicente de Paulo Emerenciano

Orientador

SÃO PAULO

2003

(Nota da BCQ: Não foi possível capturar fielmente a imagem das figuras desta tese)

DEDALUS - Acervo - CQ

30100005730

Ficha Catalográfica

Elaborada pela Divisão de Biblioteca e
Documentação do Conjunto das Químicas da USP.

Ferreira, Marcelo José Pena
F383e Esteróides : análise espectral, geração de estruturas e
simulação de dados de RMN ^{13}C / Marcelo José Pena
Ferreira. -- São Paulo, 2003.
307p.

Tese (doutorado) – Instituto de Química da Universidade
de São Paulo. Departamento de Química Fundamental.
Orientador: Emerenciano, Vicente de Paulo

1. Produtos naturais : Química orgânica 2. Computação
aplicada à Química I. T. II. Emerenciano, Vicente de
Paulo, orientador.

547.7 CDD

“Esteróides: análise espectral, geração de estruturas e simulação de dados de RMN ^{13}C ”

MARCELO JOSÉ PENA FERREIRA

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Química – Área: Química Orgânica

Aprovado por:

Prof. Dr. VICENTE DE PAULO EMERENCIANO
IQ – USP
(Orientador e Presidente)

Profa. Dra. LILIANA MARZORATI
IQ – USP

Profa. Dra. DAISY DE BRITO REZENDE
IQ – USP

Profa. Dra. MARIA AUXILIADORA COELHO KAPLAN
UFRJ

Prof. Dr. MARIA CLAUDIA MARX YOUNG
IB – SP

SÃO PAULO
24 DE OUTUBRO 2003.

Aos meus pais, José e Maria

AGRADECIMENTOS

Ao Prof. Dr. Vicente de Paulo Emerenciano pela orientação e, principalmente, por toda a liberdade, responsabilidade, confiança e sugestões que permitiram o bom desenvolvimento desse trabalho e inúmeros outros durante estes anos no seu grupo de pesquisas.

Ao Prof. Dr. Gilberto do Vale Rodrigues pela criação e aperfeiçoamento dos programas descritos nessa tese.

Ao Antônio Jedson Caldeira Brant por sua amizade, sugestões, paciência e auxílio nos diversos trabalhos desenvolvidos.

À Profa. Dra. Patricia de Almeida Telles Macari por sua amizade e por todo o auxílio durante os primeiros meses de uso do SISTEMAT.

Às Dras. Luz Consuelo G. Alonso Panzarini e Joana D'arc Felício de Souza por todo o incentivo e apoio. É um privilégio ser amigo de vocês.

Aos colegas de grupo Sandra, Mara e Alessandra pelos bons momentos proporcionados.

A todos os funcionários da Secretaria de Pós-Graduação pelo atendimento tão solícito durante todos esses anos no Instituto.

À minha família pela confiança e apoio constantes.

A meu assessor por todos os comentários e apoio nas solicitações junto ao órgão financiador.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, pela bolsa e reserva técnica concedidas.

RESUMO

O sistema especialista SISTEMAT tem por objetivo auxiliar pesquisadores da área de produtos naturais no processo de determinação estrutural de substâncias. Para tanto, utilizando dados provenientes de várias técnicas espectrométricas e espectroscópicas, principalmente RMN ^{13}C , inúmeros programas foram desenvolvidos com a finalidade de propor o provável esqueleto de uma substância. Essa informação, juntamente com as subestruturas apresentadas a partir de um conjunto de dados, é utilizada por geradores estruturais como grandes restrições, a fim de impedir a explosão combinatória e a geração de propostas estruturais incompatíveis com produtos naturais, além de reduzir o elevado tempo computacional gasto durante uma análise.

Esse trabalho descreve o desenvolvimento e utilização dos módulos de reconhecimento de esqueletos, determinação e geração estrutural e simulação de dados de RMN ^{13}C de esteróides. Assim, foi elaborada uma base de dados com 1436 substâncias distribuídas entre 119 tipos de esqueletos provenientes das mais diversas fontes naturais. Vários testes foram realizados e bons percentuais de acerto foram obtidos para o reconhecimento de esqueletos e geração de propostas estruturais através da sobreposição dos tipos de anéis encontrados em esqueletos de esteróides. Para validar as propostas estruturais apresentadas pelo gerador, bem como para prever os dados de deslocamentos químicos de novos esteróides, o simulador de dados de RMN ^{13}C foi usado e, quando comparado a um programa comercial de mesma finalidade, apresentou maior exatidão na previsão dos dados.

ABSTRACT

The aim of the expert system SISTEMAT is to aid natural product researchers in the process of structural determination of organic substances. For that, using data from various spectrometric and spectroscopic techniques, mainly ^{13}C NMR, countless programs were developed to propose the most probable skeleton of a substance. This information together with the substructures shown from the data set are utilized by structural generators as important constraints in order to avoid the combinatorial explosion problem and the generation of incompatible structural proposals for natural products, besides reducing the computational time spent during the analysis.

This work describes the development and use of the modules of skeleton identification, structural determination and generation, and the ^{13}C NMR data prediction of steroids. Thus, was built a database containing 1436 steroids distributed in 119 different skeletons originated from the most varied natural sources. Several tests were performed, wherein good hit percentuals were obtained for the skeleton identification and structural generation through the overlapping of the types of rings found in the steroid skeletons. For validation of the structural proposals shown by the generator as well as for prediction of the chemical shift data of new substances, the simulator of ^{13}C NMR data was used and next compared with a commercial program of the same purpose, and exhibited higher accuracy in the data prediction.

SUMÁRIO

I – INTRODUÇÃO	1
II – SISTEMAS ESPECIALISTAS EM DETERMINAÇÃO ESTRUTURAL	
2.1. O Sistema DENDRAL	6
2.2. O Sistema SpecInfo	9
2.3. O Sistema SENECA	14
III – O SISTEMA ESPECIALISTA SISTEMAT	
3.1. Metodologia	18
3.1.1. A Revisão Bibliográfica	18
3.1.2. A Organização dos Bancos de Dados	18
3.2. Programas de Criação da Base de Dados	20
3.2.1. O Programa DATASIS	20
3.2.2. O Programa ARQUISIS	25
3.3. Programas Aplicativos	37
3.3.1. O Programa SISLIST	37
3.3.2. O Sistema PESQUISA	37
3.3.2.1. O Programa MACRONO	42
3.3.2.2. O Programa SISCONST	55
3.3.2.3. O Programa C13MACH	58
3.3.2.4. O Programa SISOCBOT	59
3.3.2.5. O Programa REGRAS	60
IV – MÉTODOS AUTOMÁTICOS DE PREVISÃO DE DADOS DE RMN ¹³ C	
4.1. O Simulador Tridimensional de Dados de RMN ¹³ C	69
4.1.1. Metodologia	69
4.1.2. A obtenção de modelos tridimensionais	69
4.1.3. O programa FOCOS	77
4.1.3.1. A descrição dos átomos	77
4.1.3.2. A criação do arquivo para simulação de dados de RMN ¹³ C ...	80
4.1.3.3. A previsão de dados de RMN ¹³ C	82
V – O GERADOR DE ESTRUTURAS DE ESTERÓIDES	
5.1. A busca de sinais característicos de RMN ¹³ C	86
5.1.1. O Programa TIPCARB	86

5.1.2. O Sistema PICKUP	87
5.2. O Programa ESTGERA	91
5.3. Validação de propostas estruturais	95
VI – RESULTADOS	98
VII – DISCUSSÃO DE RESULTADOS	269
7.1. Programas aplicativos do SISTEMAT	269
7.2. Programas de simulação de dados de RMN ^{13}C	275
7.3. Programa de geração de estruturas – ESTGERA	279
VIII – CONCLUSÕES	281
IX – REFERÊNCIAS BIBLIOGRÁFICAS	284

SUMÁRIO DE FIGURAS

Figura 1. Estrutura do warburganal	7
Figura 2. Fragmentos estruturais fornecidos ao DENDRAL para gerar a estrutura do warburganal	8
Figura 3. Estrutura da diasina	9
Figura 4. Fluxograma do processo de geração de estruturas do SpecSolv	12
Figura 5. Estrutura da uspalatina	13
Figura 6. Estrutura da 15-oxo-23,24-di-hidrocucurbitacina F	13
Figura 7. Estruturas utilizadas para testar o sistema SENECA	15
Figura 8. Formulários apropriados para arquivamento dos dados	19
Figura 9. Tela do menu principal do programa DATASIS	20
Figura 10. Tela de desenho do programa DATASIS	21
Figura 11. Direção das ligações do vetor passos	23
Figura 12. Tela de codificação do programa DATASIS	24
Figura 13. Tela de leitura dos dados do programa DATASIS	24
Figura 14. Esquema dos bancos distribuídos e interligados do SISTEMAT	26
Figura 15. Esqueletos de esteróides presentes na base de dados	27
Figura 16. Tela do menu principal do sistema PESQUISA	41
Figura 17. Estruturas dos macronós presentes na base de dados	43
Figura 18. Esteróide utilizado para testar o sistema PESQUISA	53
Figura 19. Grupos substituintes propostos pelo programa MACRONO	54
Figura 20. Menu de programas aplicativos do sistema	54
Figura 21. Primeira subestrutura proposta pelo programa SISCONST	56
Figura 22. Segunda subestrutura proposta pelo programa SISCONST	57
Figura 23. Terceira subestrutura proposta pelo programa SISCONST	57
Figura 24. Estrutura com maior índice de similaridade apresentada pelo programa C13MACH	59
Figura 25. Exemplo do cálculo dos tipos de átomos de carbono	60
Figura 26. Tela inicial de desfuncionalização de dados de RMN ¹³ C	65
Figura 27. Tela final de desfuncionalização de dados de RMN ¹³ C	66
Figura 28. Probabilidade global de esqueleto calculada pelo sistema PESQUISA ...	66
Figura 29. Tela do programa HyperChem	70

Figura 30. Estrutura de menor energia para o colesterol	70
Figura 31. Arquivo gerado pelo HyperChem para a estrutura do colesterol	71
Figura 32. Distâncias observadas experimentalmente <i>versus</i> distâncias previstas	72
Figura 33. Esteróides utilizados para validar o modelo tridimensional do simulador	73
Figura 34. Tela inicial do simulador FOCOS	77
Figura 35. Arquivo estruturado gerado através do programa FOCOS	79
Figura 36. Lista completa do primeiro foco gerado através do programa FOCOS ...	81
Figura 37. Lista compacta dos focos presentes na base de dados	81
Figura 38. Dados de RMN ¹³ C previstos para o esteróide da Figura 18	83
Figura 39. Valores experimentais versus valores previstos pelo simulador FOCOS	84
Figura 40. Esqueleto base de um esteróide	85
Figura 41. Esqueleto cardanolídeo	86
Figura 42. Tela inicial do programa ESTGERA	91
Figura 43. Prováveis grupos substituintes propostos pelo programa ESTGERA	92
Figura 44. Fragmentos propostos pelo programa ESTGERA	93
Figura 45. Primeira proposta estrutural gerada pelo programa ESTGERA	93
Figura 46. Segunda proposta estrutural gerada pelo programa ESTGERA	94
Figura 47. Teceira proposta estrutural gerada pelo programa ESTGERA	94
Figura 48. Fluxograma de análise de dados dos programas do SISTEMAT	97
Figura 49. Dados experimentais vs. dados previstos pelo simulador FOCOS	277
Figura 50. Dados experimentais vs. dados previstos pelo simulador da ACD Labs	277

SUMÁRIO DE TABELAS

Tabela I. Conteúdo dos bancos de dados distribuídos e interligados do SISTEMAT	26
Tabela II. Listagem do número de ocorrências de esqueletos de esteróides	38
Tabela III. Relação do número total de carbonos e dos números de carbonos quaternários, metínicos, metilênicos e metílicos dos esqueletos de esteróides presentes no banco de dados	61
Tabela IV. Faixas de deslocamentos químicos utilizadas para desfuncionalização dos dados de RMN ^{13}C	64
Tabela V. Distâncias experimentais e previstas para os esteróides	74
Tabela VI. Códigos dos grupamentos químicos utilizados na codificação do simulador	78
Tabela VII. Dados de RMN ^{13}C previstos por diferentes programas simuladores	84
Tabela VIII. Listagem obtida através do programa TIPCARB para o esqueleto cardanolídeo	87
Tabela IX. Códigos de subestruturas utilizados pelo programa SISPICK2	89
Tabela X. Codificação da subestrutura do esqueleto cardanolídeo	90
Tabela XI. Resultados obtidos para a subestrutura selecionada do esqueleto cardanolídeo	90
Tabela XII. Avaliação do sistema de validação de estruturas	96
Tabela XIII. Faixas de deslocamento químico características de esteróides – Anel A	100
Tabela XIV. Faixas de deslocamento químico características de esteróides – Anel B	117
Tabela XV. Faixas de deslocamento químico características de esteróides – Anel C	130
Tabela XVI. Faixas de deslocamento químico características de esteróides – Anel D	140
Tabela XVII. Faixas de deslocamento químico características de esteróides – Cadeia Lateral	151
Tabela XVIII. Fragmentos estruturais propostos pelo programa ESTGERA	266
Tabela XIX. Resumo das análises realizadas através do programa SISCONST	270
Tabela XX. Resumo das análises realizadas através do programa C13MACH	271

Tabela XXI. Novas faixas de deslocamentos químicos de RMN ^{13}C usadas para desfuncionalização dos dados	272
Tabela XXII. Resumo das análises realizadas através do programa MACRONO	273
Tabela XXIII. Resumo das análises realizadas com esteróides através dos simuladores de dados de RMN ^{13}C	275
Tabela XXIV. Resumo das análises realizadas com o gerador ESTGERA	279

ABREVIATURAS

EM	–	Espectro de Massas
IV	–	Infra-vermelho
RMN ^{13}C	–	Ressonância Magnética Nuclear de ^{13}C
RMN ^1H	–	Ressonância Magnética Nuclear de ^1H
IS	–	Índice de Similaridade
CSS	–	Correlações Substrutura-Subespectro
HMBC	–	
DEPT	–	Distorção por Transferência de Polarização
COSY	–	Espectroscopia de Correlação
Å	–	Ångstrom
s	–	Singleto
d	–	Dubleto
t	–	Tripleto
q	–	Quarteto
CDCl_3	–	Clorofórmio deuterado
CD_3OD	–	Metanol tetradeuterado
$\text{C}_5\text{D}_5\text{N}$	–	Piridina pentadeuterada
DMSO-d^6	–	Dimetilsulfóxido hexadeuterado
Ac	–	Acetila
MeO	–	Metoxila
Glc	–	Glicose
Rha	–	Rhamnose
Dig	–	Digitalose
Ara	–	Arabinose
Gal_1	–	3-O-Metil-6-Deoxigalactose
Gal_2	–	Galoíla
EP	–	Epóxido
OXI	–	Éter cíclico diferente de epóxido
OXO	–	Carbonila
EN	–	Ligação dupla
RLM	–	Regressão Linear Múltipla

I - INTRODUÇÃO

O problema da determinação de estruturas químicas de substâncias orgânicas é tão antigo quanto a própria Química Orgânica. Embora tenham existido tentativas anteriores, somente em 1859, August Kekulé, em uma racionalização intuitiva dos dados existentes na época, descreveu a estrutura de hidrocarbonetos simples e, posteriormente, a partir do conceito de valência dos átomos frequentemente encontrados em substâncias orgânicas e da fórmula molecular de uma substância desconhecida, descreveu a existência de diferentes substâncias de mesma fórmula molecular e elaborou todos os possíveis isômeros estruturais teóricos para uma dada estrutura. Nasce então, a mais antiga abordagem da determinação estrutural, através da qual a designação da estrutura correta de uma substância foi reduzida ao problema de distinção entre os diferentes isômeros, o que exigia um estudo químico da substância desconhecida e/ou a síntese de um ou mais isômeros. Este processo foi limitado às substâncias mais simples e a pouca informação química disponível nesta época. Assim, desde o princípio, os procedimentos utilizados na determinação estrutural de substâncias orgânicas foram de natureza empírica e baseados na experiência e observação.

Em meados de 1950, uma ferramenta nova e poderosa para a investigação química começa a emergir: a espectroscopia molecular. O desenvolvimento dessa área foi bastante acelerado e a determinação estrutural começa a ser mais dependente dos dados espectrais do que dos dados químicos para a identificação da possível estrutura, uma vez que a velocidade de aquisição dos dados espectrais era maior em relação aos dados químicos.

A identificação de uma substância através da determinação estrutural convencional não tem sido um problema para os especialistas. No entanto, com o aumento do número de substâncias sintetizadas e isoladas, e da conseqüente demanda por determinação estrutural, ficou evidente que o tempo requerido pelo espectroscopista para a interpretação completa de um conjunto de dados espectrais e redução da informação estrutural deduzida para uma estrutura molecular plausível, tornou-se o fator limitante na produtividade. Desta forma, através de consecutivas gerações de espectrômetros fornecendo rapidamente dados espectrais, começou a surgir o problema da determinação incorreta de estruturas.

Dada a necessidade de processar eficientemente grande quantidade de dados espectrais, foi natural que os pesquisadores se interessassem pelo computador e,

realmente, por mais de trinta anos, o computador tem sido o centro de várias tentativas para aumentar a produtividade de químicos e espectroscopistas empenhados na determinação estrutural. A tarefa não tem sido fácil pois o objetivo é bastante ambicioso: desenvolver uma máquina capaz de simular o raciocínio humano. Buscando essa meta, as técnicas de Inteligência Artificial que, no sentido mais amplo, procuram simular esse raciocínio para resolver problemas complexos¹, têm se mostrado um campo de trabalho muito explorado, pois a determinação estrutural, embora de domínio vasto e complexo, é relativamente bem focada e definida, e também um campo no qual existe um conhecimento grande e coerente. Por estas razões, a determinação estrutural de substâncias orgânicas foi um dos primeiros e mais extensivos estudos de aplicação da Inteligência Artificial².

Os programas que utilizam técnicas de Inteligência Artificial são denominados de Sistemas Especialistas³ pois operam na solução de problemas de domínio restrito, usando um método particular, que é baseado em um modelo computacional do raciocínio de um especialista. Independentemente da área de atuação, esses sistemas apresentam em comum características como: uma grande base de dados, flexibilidade para aquisição de novos conhecimentos, facilidade de interação com o usuário, um grau de confiabilidade da resposta apresentada e, principalmente, a disseminação do conhecimento humano especializado, raro e de alto custo.

Muitos sistemas especialistas criados realizam buscas exaustivas e sistemáticas que acabam resultando em um elevado tempo computacional ou em uma explosão combinatória. Para diminuir esses problemas e reduzir o espaço de busca são utilizadas regras, que foram desenvolvidas a partir da observação de resultados típicos³, com o objetivo que estas forneçam uma orientação sobre como resolver um problema particular. Essas regras são denominadas de regras heurísticas. Por exemplo, a afirmação de que “substâncias voláteis são constituintes de baixo peso molecular” é uma heurística, pois não define que estrutura a substância apresenta, nem como sintetizá-la, mas auxilia na redução do espaço de busca durante a resolução de um problema⁴.

Vários tipos de problemas são abordados por sistemas especialistas, principalmente nas áreas científicas e tecnológicas. A seguir, são apresentados alguns exemplos e aplicações de sistemas especialistas em diversas áreas do conhecimento.

A - Área Industrial

1. Planejamento de produtos: o sistema seleciona os processos e materiais que devem ser utilizados durante a produção de produtos e estima os custos do procedimento⁵.
2. Controle de qualidade: o sistema TIREDX⁶ identifica defeitos durante a produção de arcos, sendo capaz de diagnosticar as prováveis causas do defeito trilhando a qualidade adquirida em vários estágios do processo de fabricação.

B - Área de Engenharia Civil

1. Segurança: O sistema é utilizado para detectar fraturas mecânicas no concreto⁷.

C - Área Geográfica

1. Informações: o sistema GEDAS⁸ fornece dados geográficos sobre uma determinada área e localiza possíveis regiões onde se pode desenvolver um tipo específico de atividade.
2. Planejamento urbano: o sistema MATISSE⁹ é um sistema de suporte cuja finalidade é auxiliar no planejamento urbano, em particular, na seleção de áreas industriais, avaliando locais potenciais para construção, a partir de um conjunto de requisitos locacionais.

D - Área Geológica

1. Prospecção: o sistema PROSPECTOR¹⁰ fornece assessoria na prospecção de petróleo e minerais em variados tipos de solo.
2. Prospecção: o sistema MODELER¹¹ auxilia o geólogo na identificação de jazidas de ferro.

E - Área Ecológica

1. Mapeamento de áreas: o sistema faz o mapeamento da distribuição da vegetação de uma determinada área e pode ser usado para investigar o funcionamento do ecossistema presente. O tamanho e o movimento de populações animais pode ser previsto interligando os recursos fornecidos por diferentes espécies vegetais. O sistema foi recentemente testado na área do parque nacional de Yellowstone¹².

F - Áreas de Apicultura, Piscicultura e Agricultura

1. Identificação de pragas: o sistema identifica pragas de abelhas e sugere tratamentos apropriados, providenciando um diagnóstico baseado na descrição da aparência externa ou da conduta da colônia afetada¹³.

2. Diagnóstico de doenças: o sistema Fish-Expert¹⁴ é composto de inúmeras regras, imagens e gráficos dos diferentes tipos de sintomas e males podendo diagnosticar vários tipos de doenças em espécies de peixes. Atualmente, o sistema tem sido testado em uso piloto por piscicultores da região norte da China.
3. Identificação de pragas: o sistema DIARES-IPM¹⁵ auxilia na identificação de insetos nocivos e benéficos, doenças e deficiências nutricionais em plantações de espécies de Solanaceae. O objetivo é prevenir pragas que podem afetar economicamente a safra.

G - Área Médica

1. Diagnóstico de doenças infecciosas bacterianas: o sistema MYCIN¹⁶ apresenta grupos de regras heurísticas que incluem sintomas, diagnósticos confirmados e excluídos, sugestões e propostas de exames e terapias adicionais além dos tipos de tratamento e medicamentos que devem ser utilizados para o tratamento de doenças infecciosas bacterianas.
2. Diagnóstico de doenças nas válvulas cardíacas: o sistema realiza a identificação de pacientes que apresentam possíveis doenças nas válvulas coronarianas, através da análise dos sinais Doppler^{17,18}.
3. Diagnóstico de microcalcificações: o sistema, através das imagens de mamografias, faz a classificação automática de agrupamentos de microcalcificações para a detecção prematura de câncer de mama¹⁹.
4. Diagnóstico de infarto agudo do miocárdio: o sistema analisa os sintomas, os fatores de risco, o histórico do paciente e interpreta os resultados do eletrocardiograma, fornecendo como resultado faixas de probabilidade do paciente sofrer um infarto agudo do miocárdio²⁰.
5. Tratamento oncológico: o sistema é parte de um conjunto integrado de informações médicas que auxilia no diagnóstico e definição das terapias, ajustando-as às características individuais do paciente que tornarão o tratamento oncológico mais adequado²¹.

H - Área Química

1. Estudo da relação estrutura química-atividade biológica: o sistema MultiCASE^{22,23} foi construído para prever a atividade de novas substâncias como ligantes do receptor estrógeno.

2. Detecção de amostras termicamente degradáveis: o sistema desenvolvido auxilia na identificação de produtos gerados através da degradação térmica de amostras de fármacos²⁴.
3. Separação de misturas: o sistema LabExpert²⁵ controla a instrumentação, a seleção da coluna, escolhe a fase móvel e outros parâmetros experimentais a fim de otimizar a separação de misturas complexas através de Cromatografia Líquida de Alta Eficiência (CLAE). Foi desenvolvido, também, outro sistema²⁶ que permite a escolha automática de condições analíticas que resultam em um melhor desempenho na separação de formulações farmacêuticas através de cromatografia líquida.
4. Polímeros: o sistema EX-PIMM²⁷ foi desenvolvido para determinar os parâmetros mais apropriados para a moldagem de materiais termoplásticos.
5. Caracterização da origem e autenticidade de alimentos: a partir de dados de Ressonância Magnética Nuclear de ^1H (RMN ^1H) o sistema desenvolvido auxilia na caracterização geográfica e conseqüente autenticidade de óleos de oliva²⁸.
6. Processos químicos: o sistema é utilizado para diagnosticar, em tempo real, defeitos durante o processo de craqueamento em refinarias, fornecendo sugestões ao operador quando ocorrem situações anormais²⁹.
7. Monitoramento de reatores nucleares: o sistema monitora a pressão da água em reatores, detectando anomalias mais rapidamente do que sistemas convencionais de alarme³⁰.
8. Síntese orgânica: podemos destacar os programas CAMEO³¹, que faz a previsão lógica dos produtos de reações orgânicas, o LHASA³² que, por retrossíntese, prevê quais reagentes de partida devem ser utilizados para a síntese de uma substância e o MECHGEN³³ que, através de cálculos cinéticos e considerações termodinâmicas, identifica o possível mecanismo da reação.
9. Determinação estrutural: os sistemas desenvolvidos para a determinação estrutural de substâncias orgânicas serão abordados posteriormente.

Os objetivos desse trabalho são: a partir do sistema especialista SISTEMAT, apresentar os módulos criados de reconhecimento de esqueletos e determinação estrutural de esteróides; apresentar o simulador tridimensional de dados de RMN ^{13}C bem como o gerador de estruturas de esteróides.

II - SISTEMAS ESPECIALISTAS EM DETERMINAÇÃO ESTRUTURAL

Nas últimas décadas, têm sido desenvolvidos inúmeros sistemas especialistas em determinação estrutural de substâncias orgânicas. Embora esses sistemas atuem utilizando dados provenientes de várias técnicas espectrométricas e espectroscópicas, todos eles são fundamentalmente dependentes dos dados de Ressonância Magnética Nuclear (RMN) e, a partir desses dados, os sistemas podem ser divididos em dois grandes grupos: os sistemas que utilizam exclusivamente dados unidimensionais de RMN e aqueles que utilizam, também, dados bidimensionais de RMN. Entre os principais sistemas especialistas desenvolvidos na área de determinação estrutural, podemos citar o DENDRAL², o ACCESS³⁴, o DARC/EPIOS^{35,36}, o SpecInfo³⁷, o LSD³⁸, o X-PERT^{39,40}, o ASSEMBLE⁴¹ e o SENECA⁴².

Nesse trabalho, será realizada uma breve descrição do sistema especialista DENDRAL, pois esse foi o pioneiro na área de determinação estrutural e inspirou o desenvolvimento de vários outros sistemas. Posteriormente, será descrito o sistema SpecInfo, por apresentar a maior base de dados e, finalmente, será descrito o mais recente sistema, o SENECA.

2.1. O Sistema DENDRAL

O sistema DENDRAL² teve sua origem em 1960, na Universidade de Stanford, quando o Prof. Lederberg, interessado pela isomeria química, criou um algoritmo combinatorial⁴³ capaz de gerar todas as estruturas possíveis a partir de uma dada composição química.

O sistema utiliza computadores de grande porte e dados provenientes de várias técnicas como, por exemplo, a espectrometria de massas (EM), e as espectroscopias no infra-vermelho (IV) e de ressonância magnética nuclear de ¹³C e ¹H (RMN ¹³C e RMN ¹H), e foi desenvolvido para auxiliar espectroscopistas nas investigações dos problemas de elucidação estrutural. Para atingir esta meta, foram desenvolvidos programas que realizam as partes mais difíceis da análise estrutural, como a geração de estruturas e a previsão de propriedades espectrais para uma substância-problema.

A metodologia utilizada por esse sistema consiste no método básico de “planejar-gerar-testar”, que são as fases distintas desenvolvidas pelo espectroscopista no processo

de determinação estrutural: a interpretação dos dados, a geração e a avaliação das estruturas.

Na etapa de planejamento, é essencial para a solução dos problemas que o usuário forneça os fragmentos subestruturais presentes e ausentes, que são inferidos a partir da análise dos dados químicos e espectrais da substância desconhecida. Na fase de geração de estruturas, o sistema utiliza o módulo gerador denominado GENOA⁴⁴, que necessita da definição da fórmula molecular do composto e das restrições subestruturais da fase anterior, para gerar todos os isômeros possíveis a partir dos fragmentos. Esta fase é puramente combinatorial e exaustiva, de modo a garantir que todas as combinações sejam realizadas, portanto, o número de soluções e o tempo computacional são muito elevados.

Na etapa de testes, as estruturas geradas são avaliadas para exclusão de propostas improváveis. Esse procedimento é realizado confrontando-se as estruturas com bancos de dados de esqueletos de produtos naturais, no caso da substância ser de origem natural, ou através da previsão de espectros e comparação dos dados teóricos com os dados do espectro questionado.

Para exemplificar a aplicação do sistema em determinação estrutural⁴⁵, foi utilizado o warburganal (Figura 1), um sesquiterpeno com esqueleto drimano, isolado de *Warburgia ugandensis*⁴⁶, que apresenta fórmula molecular $C_{15}H_{22}O_3$. O usuário, nesse teste, forneceu os fragmentos estruturais presentes e ausentes no composto, Figura 2, inferidos a partir dos dados dos espectros de IV, RMN 1H e RMN ^{13}C , além da fórmula molecular, obtida a partir do espectro de massas.

Figura 1. Estrutura do warburganal

A partir das restrições impostas, o programa gerou 42 estruturas, algumas totalmente incompatíveis com produtos naturais mas, dentre elas, a estrutura correta do warburganal.

FRAGMENTOS

Figura 2. Fragmentos estruturais fornecidos ao DENDRAL para gerar a estrutura do warburganal

Após a geração completa das estruturas, estas são transferidas para o programa STRCHK⁴⁵ que faz a avaliação. Esse programa, através do módulo SURVEY, examinou se as estruturas incorporavam esqueletos ou feições (fragmentos de esqueletos) de produtos naturais. De todas as substâncias geradas, apenas uma incorporava um esqueleto completo de produto natural, e esta correspondia à estrutura do warburganal. As substâncias geradas também foram avaliadas através da comparação dos espectros de massas e de RMN ¹³C reais com os previstos e, nos dois casos, a primeira substância na escala de similaridade correspondia à estrutura do warburganal.

Um outro produto natural usado para avaliar o gerador de estruturas do sistema DENDRAL foi a diasina (Figura 3), um diterpeno com esqueleto labdano rearranjado, isolado de *Croton diasii*⁴⁷, que apresenta fórmula molecular C₂₁H₂₄O₇. O gerador forneceu 145 estruturas⁴⁸ que, posteriormente analisadas pelo programa STRCHK e seu módulo SURVEY, foram reduzidas a duas estruturas prováveis, das quais uma foi

eliminada através de considerações biogenéticas, restando apenas a estrutura correta da diasina.

Figura 3. Estrutura da diasina

2.2. O Sistema SpecInfo

O sistema de informação espectroscópica multidimensional SpecInfo tem sido desenvolvido desde 1970, pela BASF, como uma ferramenta para auxiliar o espectroscopista analítico, visto que a automatização dos laboratórios espectroscópicos e analíticos gerou, durante as últimas décadas, um número maior de espectros que necessitam ser avaliados rapidamente e a um custo mínimo. Esse sistema encontra-se em sua terceira versão, utilizando uma interface gráfica do tipo X-Windows e sendo executado em estações de trabalho^{37,49}.

A base de dados do SpecInfo apresentava, em 1990, 16.000 espectros de massas, 110.000 espectros de RMN¹³C, 15.000 espectros de RMN ¹¹B, ¹⁵N, ¹⁷O, ¹⁹F e ³¹P e 20.000 espectros de IV³⁷. Para assegurar a confiabilidade do SpecInfo, cada espectro a ser incluído na base de dados passa por uma rotina de controle de qualidade, que realiza testes automatizados e, também, por uma inspeção final de um espectroscopista experiente, de modo que espectros de baixa qualidade sejam descartados antes de contaminar a base de dados.

O SpecInfo tem como princípio a idéia de que ambientes químicos de um determinado átomo em uma molécula influenciam as suas propriedades espectroscópicas. Baseado nesse princípio, foi construído um sistema de codificação de subestruturas utilizando o código HOSE (Hierarchically Ordered Spherical Description

of Environment)⁵⁰. Esse código caracteriza o ambiente de cada átomo como uma esfera ao redor de um centro específico de interesse, isto é, a partir de um nó central são descritos todos os átomos que o circundam até a terceira esfera. Desta forma, os efeitos sobre o acoplamento e o deslocamento químico do átomo central fazem parte da informação calculada e armazenada para qualquer átomo na base de dados.

O resultado desse processo é um “estoque” de ambientes químicos de átomos, cuja aplicação não está limitada a classes específicas de substâncias. Uma vez que cada átomo da base de dados está correlacionado a um padrão espectral, é possível prever o espectro de uma substância, pois cada átomo será convertido automaticamente em uma lista de códigos HOSE e, através de uma busca na biblioteca de RMN ^{13}C , será fornecido um deslocamento químico para cada código HOSE da lista, portanto, um deslocamento químico para cada átomo de carbono da substância.

Recentemente, foi criado um novo módulo para o sistema SpecInfo, denominado de SpecSolv⁵¹. Esse módulo trabalha exclusivamente com RMN ^{13}C e não necessita de informações adicionais provenientes de outras fontes espectroscópicas como o IV, EM e RMN ^1H . A elucidação estrutural com o SpecSolv é dividida em três etapas:

1. Aquisição do espectro de RMN ^{13}C -DEPT experimental e extração dos deslocamentos químicos, multiplicidades e intensidades.
2. A procura de subespectros em uma biblioteca de correlações subestrutura-subespectro (CSS) e geração de uma lista com os melhores resultados.
3. Reunião das subestruturas, através da sobreposição destas, seguida de um passo de validação das subestruturas intermediárias e do resultado final.

A base de dados do SpecSolv

A base de dados do SpecSolv contém mais de 400.000 subestruturas geradas a partir dos códigos HOSE de terceira esfera e mais de 100.000 a partir de códigos HOSE de segunda esfera. Esses CSS foram originados a partir da base de dados do SpecInfo que, atualmente, apresenta mais de 200.000 espectros de RMN ^{13}C . Cada CSS apresenta uma informação estrutural completa (subespectros, matriz de conectividade e código HOSE para todos os átomos) e contém todos parâmetros de RMN ^{13}C (deslocamento químico, multiplicidade e intensidade) dos átomos de carbono.

A procura de subespectros

A procura de subespectros consiste na comparação de todos subespectros da base de dados de CSS com o espectro-problema. Durante a busca, somente são selecionados CSS com deslocamento químico, multiplicidades e intensidades dentro da faixa definida pelo usuário. As subestruturas selecionadas são, então, classificadas de acordo com um fator de comparação que indica o desvio médio, em ppm, dos deslocamentos químicos entre as CSSs e o espectro-problema.

O SpecSolv não requer o conhecimento dos tipos de átomos do espectro questionado, mas restrições de heteroátomos multivalentes como, por exemplo, N, P e S, reduzem o número de CSSs propostos e, consequentemente, o tempo computacional.

O Gerador de Estruturas

As subestruturas provenientes do processo de busca de CSS são interligadas pela sobreposição dos átomos comuns às duas subestruturas. A maior subestrutura gerada é então validada por uma nova previsão espectral, baseada nos CSSs. Se o espectro previsto for compatível com o espectro questionado, a subestrutura gerada é aceita e utilizada como novo ponto de partida, até que todas as subestruturas sejam sobrepostas com a subestrutura de partida e/ou todos os deslocamentos químicos do espectro problema estejam completamente atribuídos. Caso o espectro previsto não seja compatível com o questionado, a subestrutura gerada é rejeitada e o processo inicia-se com a escolha da próxima subestrutura melhor classificada. O processo de geração de estruturas do SpecSolv é descrito no fluxograma apresentado na Figura 4.

Vale observar que, durante o processo de sobreposição das subestruturas individuais, os deslocamentos químicos da subestrutura resultante ficam disponíveis na memória central do computador, assim um novo subespectro pode ser gerado pela simples combinação dos subespectros originais de uma forma rápida, pois não é necessário o acesso a uma base de dados.

A validação instantânea das estruturas intermediárias é uma das principais diferenças entre o SpecSolv e outros programas geradores de estruturas, que realizam a validação de estruturas somente após a geração da molécula inteira. Esse processo no SpecSolv torna o espaço de solução significativamente menor, o que acarreta uma

redução crucial do tempo computacional quando comparado a outros programas convencionais.

Figura 4. Fluxograma do processo de geração de estruturas do SpecSolv

Exemplos

Para verificar o potencial do SpecSolv como uma ferramenta no campo da determinação estrutural, foram submetidas a testes as substâncias uspalatina⁵² e 15-oxo-23,24-di-hidrocucurbitacina F⁵³, apresentadas, respectivamente, nas Figuras 5 e 6.

A elucidação estrutural da uspalatina foi realizada automaticamente pelo SpecSolv, em um computador VAX 6610, em menos de três minutos. Os dados de entrada fornecidos ao programa foram os deslocamentos químicos, multiplicidades e intensidades dos sinais do espectro de RMN ¹³C, sendo atribuído um valor máximo do tipo C₁₀₀H₁₀₀O₁₀₀N₁₀₀ para a fórmula molecular. Para minimizar o tempo computacional, também foi estimado o número de átomos de nitrogênio próximo do valor real. A busca de subespectros resultou em 150 subestruturas, através das quais o

gerador construiu apenas uma molécula que apresentou espectro previsto compatível ao experimental, com um fator de comparação de 0,448, sendo essa a maior diferença observada. A estrutura gerada correspondia à estrutura correta da uspalatina.

Figura 5. Estrutura da uspalatina

Para a estrutura da 15-oxo-23,24-di-hidrocucurbitacina F, o SpecSolv encontrou 1118 CSSs para o espectro experimental, contudo o gerador de estruturas apresentou, em menos de três minutos, duas propostas estruturais, ambas com bons fatores de comparação e compatíveis com o espectro problema. No entanto, uma das estruturas foi descartada por apresentar um sistema cíclico com uma ponte incomum. A estrutura resultante correspondia à substância correta apresentada na Figura 6. Nos dois exemplos, o sistema não forneceu as estereoquímicas das substâncias.

Figura 6. Estrutura da 15-oxo-23,24-di-hidrocucurbitacina F

Limitações do SpecSolv

O módulo SpecSolv é capaz de gerar estruturas e prever espectros para substâncias com massa até aproximadamente 1000u e que apresentem somente átomos de C, H, N, O, S, P e halogênios. Moléculas maiores podem requerer um tempo computacional muito grande devido o número de possibilidades combinatoriais para sobrepor as subestruturas. Além disso, o sistema somente é capaz de elucidar estruturas que possam ser descritas pela sobreposição de subestruturas, portanto, não é possível a caracterização de componentes individuais de uma mistura com o sistema.

A previsão de deslocamentos químicos pelo sistema é afetada negativamente pelo efeito do solvente ou influências estereoquímicas, uma vez que ele não apresenta informações estereoquímicas e não descreve efeitos configuracionais. Como qualquer outro sistema que utilize esse tipo de dado, o SpecSolv depende crucialmente da qualidade e diversidade da base de dados de correlações subestrutura-subespectro.

2.3. O Sistema SENECA

Nos últimos anos, vários sistemas especialistas em determinação estrutural começaram a utilizar os dados de conectividade fornecidos por espectros bidimensionais de RMN^{54,55}. Um dos mais recentes sistemas desenvolvidos com esta abordagem, é o sistema SENECA⁴².

O SENECA apresenta um módulo denominado LUCY⁵⁶, que trabalha exclusivamente com as correlações obtidas através de experimentos de RMN, com ênfase especial nos dados de HMBC, e não necessita de informações adicionais provenientes de outras fontes espectroscópicas e espectrométricas como o IV e a EM. Esse módulo utiliza as seguintes informações no processo de elucidação estrutural: dados de RMN ¹³C desacoplados, dados dos DEPT-90 e DEPT-135 e dados do espectro C,H COSY (HMQC e HSQC), sendo opcionais os dados do espectro H,H COSY.

O processo utiliza os dados de HMBC para gerar fragmentos longos do esqueleto carbônico no primeiro passo da elucidação da estrutura. Os fragmentos CH_x (x= 0-3) e os heteroátomos, obtidos a partir dos experimentos DEPT, são combinados em um processo repetitivo para gerar todas as possíveis interpretações do acoplamento C,H a longa distância. O algoritmo⁵⁶ é baseado na hipótese que os sinais cruzados do espectro

HMBC são representações dos acoplamentos $^2J_{C,H}$ e $^3J_{C,H}$. Assim, dois átomos C_A e B (onde B pode ser um átomo de carbono ou um heteroátomo ligado a hidrogênio) entre os quais a correlação $C_A \rightarrow B-H$ existe, podem ser conectados entre si, tanto diretamente ($^2J_{C,H}$) quanto através de um átomo opcional, que posteriormente será determinado, na matrix ($^3J_{C,H}$).

A geração desses fragmentos, a partir dos dados de HMBC, conduz a uma limitação extremamente efetiva do número de soluções e, conseqüentemente, o tempo gasto para a elucidação estrutural é consideravelmente menor em comparação com os sistemas que utilizam os dados de HMBC para a seleção de estruturas, após o processo de geração estrutural^{57,58}.

Para exemplificar a aplicação e eficiência do sistema SENECA em determinação estrutural^{42,56}, foram utilizadas três substâncias: a monochaetina (Figura 7), um metabólito isolado do fungo *Monochaetia compta*⁵⁹, que apresenta fórmula molecular $C_{18}H_{20}O_5$, que já havia sido utilizado anteriormente para testar o sistema ASSEMBLE⁵⁴; o eurabidiol (Figura 7), um sesquiterpeno com esqueleto guaiano, isolado de *Euryops arabicus*⁶⁰, que apresenta fórmula molecular $C_{15}H_{28}O_2$ e, o policarpol (Figura 7), um triterpeno de esqueleto lanostano e fórmula molecular $C_{30}H_{48}O_2$, isolado de *Polyalthia oliveri*⁶¹.

Figura 7. Estruturas utilizadas para testar o sistema SENECA

Nos três casos analisados, o sistema propôs a estrutura correta da substância⁴². No primeiro teste, somente a estrutura da monochaetina foi proposta pelo sistema, enquanto o sistema previamente testado havia proposto seis prováveis estruturas⁵⁴. No caso dos terpenóides, o sistema forneceu duas e seis estruturas prováveis para o sesquiterpeno e para o triterpeno, respectivamente. Nos dois últimos casos, foi necessário utilizar um método rápido de identificação de esqueletos de produtos naturais⁶² para que o sistema selecionasse corretamente, entre as propostas apresentadas, a estrutura da substância.

III - O SISTEMA ESPECIALISTA SISTEMAT

O sistema especialista SISTEMAT^{63,64} desenvolvido no Instituto de Química da Universidade de São Paulo, a partir de 1988, tem por objetivo auxiliar pesquisadores da área de produtos naturais no processo de determinação estrutural de substâncias e em estudos quimiotaxonômicos.

Inicialmente, foram desenvolvidas técnicas de codificação das substâncias, isto é, de representação das moléculas e de arquivamento dos dados relacionados a estas. Posteriormente, foram criados vários programas aplicativos para determinação estrutural utilizando dados de diversas fontes espectroscópicas aliados a informações sobre a origem da substância.

O processo de codificação de estruturas era, inicialmente, manual e a criação da base de dados requeria vários programas, o que tornava o processo muito lento e passível de erros. Em 1992, com a criação do programa CODISIS⁶⁵, a codificação das substâncias passou a ser semi-automática e solicitava do usuário somente a definição da sequência da numeração da cadeia carbônica da substância a ser codificada. Posteriormente, esse programa foi aperfeiçoado e recebeu o nome de DATASIS, que realiza a codificação de forma totalmente automática. O reflexo imediato desta automatização foi o aumento na velocidade de arquivamento dos dados e a diminuição do número de erros.

O SISTEMAT tem sua aplicação voltada para a química de produtos naturais, sendo assim, informações como a classe química da substância e o tipo de esqueleto carbônico, que são pontos-chave no processo de determinação estrutural, foram considerados e introduzidos na base de dados, portanto cada substância arquivada no sistema apresenta, além dos dados espectrais obtidos, informações como a classe química, o tipo de esqueleto e a origem (família, gênero e espécie) da substância. Também foram desenvolvidos programas que auxiliam na busca de faixas de deslocamentos químicos de RMN ¹³C característicos dos diversos tipos de esqueletos e/ou subesqueletos, pois a utilização desse conceito durante a etapa de geração de estruturas restringe o número de propostas, evitando-se assim uma possível explosão combinatória, quando as moléculas analisadas são muito complexas como, por exemplo, diterpenos, triterpenos, esteróides e alcalóides, além de resultar em um tempo computacional menor.

Uma das grandes e principais vantagens do SISTEMAT, em relação a outros sistemas especialistas em determinação estrutural, é que ele possui um algoritmo de codificação estrutural e compactação de dados extremamente eficiente, possibilitando assim a utilização de microcomputadores em ambiente DOS, enquanto os demais sistemas são executados em computadores de grande e médio portes. Isto faz com que o SISTEMAT possa ser utilizado por qualquer usuário que disponha de um microcomputador. Esse método de codificação também permite que o próprio sistema descubra informações químicas das substâncias tais como: grupos funcionais, massa e fórmula moleculares e número de oxidação, a partir da estrutura codificada.

No grupo de pesquisa que trabalha com o SISTEMAT, o crescimento da base de dados se dá por classes químicas, de modo que cada pessoa trabalhe com substâncias afins e de interesse para o grupo. Atualmente, o SISTEMAT apresenta um total de 13.000 espectros de RMN ^{13}C , 1.900 de RMN ^1H , cerca de 1.700 espectros de massas e 27.000 ocorrências de produtos naturais, tendo sido estudadas as seguintes classes químicas de substâncias: monoterpenos⁶⁶, iridóides, sesquiterpenos lactonizados⁶⁷ e não-lactonizados⁶⁸, diterpenos⁶⁹, triterpenos⁷⁰, flavonóides⁷¹ e esteróides (objeto do presente estudo). Paralelamente a esse trabalho, outras classes de produtos naturais como os lignóides e alguns tipos de alcalóides estão sendo inseridos na base de dados. Futuramente, outras classes químicas tais como cumarinas, alcalóides indólicos, benzofuranos, benzopiranos, acetofenonas, xantonas e antibióticos serão acrescentados à base de dados.

Em 1998, foi desenvolvido um pré-gerador de estruturas do SISTEMAT, denominado SISGER⁷², que utiliza informações como a fórmula molecular, o tipo de esqueleto, inferido a partir dos outros programas do sistema, e um banco de dados de subestruturas com a descrição dos tipos de carbonos em diversos ambientes químicos. Estas subestruturas são denominadas de ELCOs⁷³ (ambientes que são limitados, concêntricos e ordenados) e são geradas a partir dos bancos fonte do SISTEMAT. O processo de geração de estruturas inicia-se com a informação do tipo de esqueleto, determinado anteriormente, e ocorre através da sobreposição progressiva dos ELCOs.

Os programas do SISTEMAT estão divididos em dois grupos distintos: os programas de criação de bancos de dados e os programas aplicativos. Esses programas foram escritos, inicialmente, em linguagem FORTRAN e, os mais recentes, em PASCAL. Esses já foram testados com sucesso em outros trabalhos e teses⁶⁶⁻⁷¹ e serão

agora aplicados à base de dados de esteróides. Novos programas aplicativos desenvolvidos também serão apresentados e testados no decorrer desse trabalho.

3.1. Metodologia

3.1.1. A Revisão Bibliográfica

A base de dados foi construída a partir de uma pesquisa bibliográfica que foi dividida em três etapas:

1. Coleta dos dados de RMN ^{13}C de esteróides a partir de um artigo de revisão⁷⁴, o qual compila os dados de literatura publicados até o ano de 1976.
2. Obtenção de dados nos periódicos “*Phytochemistry*” e “*Journal of Natural Products*”, que são periódicos especializados na publicação de produtos naturais.
3. Obtenção dos dados nas referências, disponíveis no IQ-USP, citadas no “*Chemical Abstracts*”, a partir do ano de 1976 através da palavra-chave “steroid*”.

3.1.2. A Organização dos Bancos de Dados

Os dados obtidos através da revisão bibliográfica foram transferidos para fichas apropriadas (Figura 8) e, posteriormente, arquivados em bancos de dados. Os dados físico-químicos introduzidos no sistema foram os de RMN ^{13}C . Os solventes utilizados na obtenção dos espectros também foram considerados no armazenamento dos dados, sendo que a maioria dos esteróides foram analisados em CDCl_3 . Para a realização das análises, todos os dados, independentemente do tipo de solvente, foram agrupados de modo que se obtivesse o maior conjunto de dados de RMN ^{13}C . Para a obtenção de deslocamentos químicos característicos de esqueletos ou subesqueletos esse agrupamento também foi utilizado, juntamente com a numeração biogenética estabelecida para cada esqueleto de esteróide.

Ficha N^o .:

Família:

Gênero/Espécie:

Referência:

Estrutura da Substância	Parte da planta:
	Dados de coleta:

RMN ¹³ C		RMN ¹ H		EM	IV
1.	26.	1.	26.		
2.	27.	2.	27.		
3.	28.	3.	28.		
4.	29.	4.	29.		
5.	30.	5.	30.		
6.	31.	6.	31.		
7.	32.	7.	32.		
8.	33.	8.	33.		
9.	34.	9.	34.		
10.	35.	10.	35.		
11.	36.	11.	36.		
12.	37.	12.	37.		
13.	38.	13.	38.		
14.	39.	14.	39.		
15.	40.	15.	40.		
16.	41.	16.	41.		
17.	42.	17.	42.		
18.	43.	18.	43.		
19.	44.	19.	44.		
20.	45.	20.	45.		
21.	46.	21.	46.		
22.	47.	22.	47.		
23.	48.	23.	48.		
24.	49.	24.	49.		
25.	50.	25.	50.		

Figura 8. Formulários apropriados para arquivamento dos dados

3.2. Programas de Criação da base de dados

3.2.1. O Programa DATASIS

O programa DATASIS⁶⁷ realiza a codificação automática das substâncias a partir do desenho da estrutura da molécula na tela do computador.

Ao iniciarmos o programa DATASIS, é apresentada uma tela “solicitando” o nome do arquivo de banco-fonte, que é um banco de dados onde as estruturas das substâncias e seus dados obtidos da literatura serão inicialmente armazenados. A seguir, serão “solicitados” os nomes dos arquivos de esqueletos e substâncias. No arquivo de esqueletos, encontram-se gravados os desenhos dos esqueletos de esteróides encontrados na revisão bibliográfica, o mesmo ocorrendo para o arquivo de substâncias. Desta forma o processo de entrada de uma nova substância é agilizado e facilitado, pois o esqueleto-base da estrutura já está desenhado. Logo em seguida, o programa apresenta a tela do menu principal (Figura 9).

Figura 9. Tela do menu principal do programa DATASIS

Através das opções 0 e 8 pode-se voltar ao Sistema Operacional, sendo que pela opção 0 a volta é definitiva, de forma que para se iniciar o programa é necessário definir, novamente, os arquivos de banco-fonte, de substâncias e de esqueletos. A opção 8 permite que a saída para o Sistema Operacional seja temporária, de forma que para se retornar ao programa não existe a necessidade da definição dos arquivos, pois o programa volta, automaticamente, para a tela do menu principal. A opção 1 serve para redefinir arquivos de banco-fonte, de substâncias ou de esqueletos, caso, na definição inicial, esta tenha sido feita de forma errônea.

A opção 3 permite a entrada de uma nova substância. Para que esta ocorra, deve-se, inicialmente, fornecer o nome trivial, que é usado para verificar se a substância já existe no banco de dados. Caso não exista nome trivial, o programa pode reconhecer se uma substância já está arquivada usando um identificador matemático, que será discutido posteriormente. Em caso negativo, o programa solicita o número da ficha (Figura 8) na qual os dados foram arquivados e apresenta a tela de desenho, Figura 10, onde se encontram as opções para manipulação dos arquivos e os símbolos utilizados para representação dos átomos e ligações.

Figura 10. Tela de desenho do programa DATASIS

Após o desenho da substância, o programa realiza automaticamente a codificação da estrutura. A representação de uma substância precisa ser traduzida em uma linguagem matemática para ser compreendida pelo computador. Este processo é denominado de codificação molecular⁶³.

No método adotado pelo SISTEMAT, utilizou-se a técnica do vetor reduzido. Esta técnica consiste na criação de um vetor que define a sequência de valores e posições dos elementos, que foi denominado de vetor nó. Para a construção do vetor nó foi criado um método de codificação arbitrário, através do qual a substância é numerada respeitando-se as seguintes regras⁶³:

A - Inicialmente, são numerados os nós monoatômicos, estes são átomos ligados por uma única ligação a um único átomo que pertence a uma cadeia (Fig. 12, átomos 01, 02, 03, 04, 05 e 06).

B - A numeração deve ser adequada para que o programa possa encontrar as ligações alfa e beta. Assim, foi estabelecido que, se um átomo estiver em posição alfa, o programa irá considerar que todas as suas ligações com átomos de maior numeração serão alfa, o mesmo sendo aplicado às ligações beta.

Após a numeração da estrutura, a próxima etapa é a elaboração do vetor nó, que deverá apresentar, sequencialmente, os seguintes dados :

1. Inicialmente serão indicados os números iniciais e finais de cada cadeia (Fig. 12, átomos 07 e 28).

2. Precedido por -1 é indicado, em ordem numérica dos nós monoatômicos, o número dos átomos aos quais estes estão ligados (Fig. 12, -1110713071720), seguidos dos números dos átomos das ligações extras da cadeia principal, isto é, dos átomos com numeração não sequencial que estão interligados (Fig. 12, 1228132616251722).

3. Precedido por -2 são indicados os heteroátomos através de seu número atômico e posições que ocupam (Fig. 12, -20806).

4. Precedido por -3 são indicadas as posições dos átomos das ligações duplas (Fig. 12, -32223).

5. Precedido por -4, indica-se a posição de um átomo pertencente a cada ciclo aromático, uma vez que o programa é capaz de reconhecer os demais átomos do ciclo.

6. Precedido por -5 são indicadas as posições dos átomos das ligações triplas.

7. Precedidos por -6-1 são indicados os átomos que apresentam estereoquímica beta (Fig. 12, -6-103050611), por -6-2 os átomos em alfa (Fig. 12, -6-201), por -6-3 os átomos com carga positiva e por -6-4 os átomos com carga negativa.

8. Precedidos por -7, indica-se os átomos da junção *cis* de anéis.
9. Precedidos por -8 são indicados as posições e códigos dos grupos substituintes.
10. Através de -9, o vetor nó é encerrado.

A partir do vetor nó, o programa reconhece os tipos de ligação, de átomos e as cadeias existentes na molécula. No entanto, ainda é necessário que esta seqüência numérica seja transformada, pelo computador, no desenho da estrutura da substância. Para tanto, as direções das ligações precisam ser indicadas. Esta informação é encontrada no vetor passos, apresentado após o vetor nó, onde, inicialmente, indicam-se as coordenadas do átomo inicial da cadeia, de acordo com a sua posição na tela (Fig. 10 e 12, 1441), seguida das orientações das ligações que compõem a estrutura, codificadas de acordo com a Figura 11.

Figura 11. Direção das ligações do vetor passos

Após o desenho e codificação da substância, o programa DATASIS “solicita” a confirmação da mesma. Em caso negativo, o programa retorna para a tela de desenho (Figura 10), em caso afirmativo, é apresentado o vetor completo da substância (vetor nó seguido do vetor passos), construído a partir das regras descritas anteriormente, o tempo e o número de bytes usados na compactação, o número de oxidação da substância e o identificador matemático (Figura 12). O identificador matemático permite a identificação de cada estrutura, além de possibilitar a rápida e fácil localização no banco de dados.

Depois da apresentação dos dados codificados, o programa apresenta a tela onde é realizada a leitura dos dados (Figura 13). Nesta, são solicitados os dados obrigatórios (classe e esqueleto da substância, código da revista, ano e página) e se existirem, os dados optativos (família, gênero, espécie e índice de evolução morfológica da planta).

Figura 12. Tela de codificação do programa DATASIS

The screenshot shows the DATASIS program interface. The background is red. At the top, a blue box contains the text "LEITURA DOS DADOS". Below this, a blue box contains the following data:

CLASSE	:	ESTEROIDE	1	1
ESQUELETO	:	COLESTANO		
CODIGO DA REVISTA	:	[91]		
ANO	:	[91]		
PAGINA	:	[3615]		
FAMILIA	:	SOLANACEAE	1	
GENERO	:	IPETUNIA	1	
ESPECIE	:	INTEGRIFOLIA	1	
ITEM	:	[]		

Below the blue box, a red box contains the text "OBSERVACOES:" followed by three lines of text:

1- HAVENDO DADOS BOTANICOS, ITEM = 1
 2- NAO HAVENDO DADOS BOTANICOS, ITEM = -1
 3- SEM FAMILIA E COM GENERO, FAMILIA = IGNORADA

Figura 13. Tela de leitura dos dados do programa DATASIS

As revistas foram definidas em um banco denominado de BREV. Neste se encontram os nomes das revistas utilizadas na pesquisa bibliográfica, sendo que cada uma apresenta um código já conhecido pelo SISTEMAT. Ao término da leitura dos dados, o programa retorna à tela de entrada do nome trivial e uma nova substância poderá ser codificada.

Em caso de entrada de algum dado errado, esse pode ser corrigido através da opção 6 do menu principal do programa DATASIS (Figura 10) e, se for necessário, pode-se eliminar, definitivamente, uma substância através da opção 7. Isto é necessário, normalmente, quando houver algum erro no desenho da estrutura, que não foi percebido durante a entrada dos dados.

3.2.2. O programa ARQUISIS

Após a codificação e arquivamento das estruturas é necessário realizar-se a distribuição dos dados dos bancos-fonte em bancos interligados, descritos na Figura 14, que apresentam apenas um tipo de dado, de acordo com as informações contidas na Tabela I. Essa etapa é feita através do programa ARQUISIS e, com este procedimento, o processo de busca de informações no sistema é agilizado, pois quando um determinado tipo de dado é pesquisado, a procura não é feita em toda a base de dados, mas sim nos bancos interligados que apresentam o dado requerido. Os bancos descritos na Figura 14 são criados a partir da opção 2 do menu principal do programa DATASIS e a distribuição dos dados é feita através da opção 5 (Figura 9, pg. 20). Somente após este procedimento, os dados físico-químicos podem ser introduzidos, utilizando-se a opção 4.

A base de dados construída contém 1436 esteróides, distribuídos em 119 tipos de esqueletos, dos quais 1405 substâncias apresentam dados de RMN ^{13}C . Os esqueletos de esteróides são apresentados na Figura 15.

Figura 14. Esquema dos bancos distribuídos e interligados do SISTEMAT

Tabela I. Conteúdo dos bancos de dados distribuídos e interligados do SISTEMAT

Banco de dados	Dados apresentados
BNUM	Número matemático criado para cada substância
BTIP	Classe das substâncias e tipos de dados
BESQ	Nome dos esqueletos
BREF	Referências bibliográficas
BREV	Nome das revistas
BESP	Nome das espécies
BGEN	Nome dos gêneros
BFAM	Nome das famílias
BCOMP	Código das substâncias
BDTA	Dados físico-químicos

Figura 15. Esqueletos de esteróides presentes na base de dados

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

Figura 15. Continuação

3.3. Programas Aplicativos

Os programas aplicativos⁶⁴ constituem a “parte inteligente” do sistema especialista SISTEMAT, pois permitem a análise dos dados do sistema e a obtenção de informações úteis que podem auxiliar no processo de determinação estrutural. A partir dos dados físico-químicos arquivados, esses programas podem auxiliar na elucidação do provável esqueleto de uma substância desconhecida.

3.3.1. O Programa SISLIST

O programa SISLIST⁷⁰ fornece dois tipos de listagem dos conteúdos dos bancos de dados. As listagens podem ser realizadas por número de substâncias e tipos de esqueletos ou por dados da origem da substância. No primeiro caso, o programa fornece uma lista do total do número de substâncias na classe química estudada e a distribuição destas por tipo de esqueleto e, no segundo caso, o programa apresenta os dados da origem da substância, isto é, para cada família presente no banco de dados, é listado o número de ocorrências, por gênero e espécie estudados.

Na Tabela II é apresentada a listagem, obtida através do programa SISLIST, do número de ocorrências por classe e tipo de esqueleto presentes no banco de dados. A listagem total das famílias mostra que existem 1.436 ocorrências de esteróides distribuídos em 94 famílias de plantas, corais, algas, fungos e bactérias, dentre outros.

3.3.2. O Sistema PESQUISA

Para facilitar o acesso da pesquisa de esqueletos no banco de dados de esteróides, foi desenvolvido, durante esse trabalho, o sistema PESQUISA que consiste na reunião de sete programas aplicativos do SISTEMAT desenvolvidos para a determinação final de esqueletos e subestruturas.

Este sistema apresenta uma tela de menus, onde o usuário digita apenas uma vez o espectro em teste e o nome da família e/ou gênero a partir da qual a substância foi isolada, sendo o conjunto de dados analisado através dos diferentes programas aplicativos. Ao final da análise, a probabilidade de que a substância teste pertença a um determinado esqueleto é calculada e apresentada ao usuário, considerando todos os programas aplicativos do SISTEMAT.

Tabela II. Listagem do número de ocorrências de esqueletos de esteróides

Classe : ESTERÓIDE	Total : 1436
Esqueletos	Número de ocorrências
1. Androstano	167
2. 3-Metilandrostando	5
3. 2-Metilandrostando	7
4. 1-Metilandrostando	3
5. Estrano	25
6. 7-Metilestrano	1
7. Cardanolídeo	100
8. 17-Metilandrostando	36
9. 16-Metilandrostando	4
10. 16,16-Dimetilandrostando	1
11. Colestano	371
12. 5-Metil-colestano	2
13. 6-Metil-colestano	14
14. 2-Metil-colestano	5
15. 3,5-Ciclocolestano	2
16. Colano	38
17. Ergostano	100
18. Pregnano	121
19. 16-Metil-pregnano	2
20. 16-Metilestrano	4
21. 16,16-Dimetilestrano	1
22. 1-Metilestrano	3
23. 4-Metilestrano	5
24. Vitanolídeo	61
25. 5,19-Cicloandrostando	16
26. Estigmastano	34
27. 3-Metil-colestano	8
28. 19-Norergostano	3
29. Fusidano	2
30. Campestando	16
31. Poriferastano	14
32. 20-Metil-pregnano	5
33. 17,18-Ciclopregnano	3
34. 3,5-Cicloandrostando	12
35. Gonano	3
36. 27-Nor-colestano	9
37. 14-Metil-poriferastano	1
38. 23,28-Cicloestigmastano	5
39. 26,27-Dinor-colestano	1
40. 2,2-Dimetil-androstando	1
41. 7-Feniletil-androstando	2
42. 3,20-Dimetil-pregnano	2
43. 4-Metil-5,19-Cicloandrostando	1
44. 3,5-Abeo-3,4-Secoandrostando	2
45. 19(10→5)-Androstando	1
46. 1,19-Cicloandrostando	3

Tabela II. Continuação

Esqueletos	Número de ocorrências
47. 12,23-Cicloergostano	1
48. 4,14-Dimetil-campestano	1
49. 4,14-Dimetil-ergostano	3
50. 5,7-Abeo-5,6-Seco-colestano	1
51. 4-Metilergostano	7
52. 1,10-Seco-1,6-Cicloergostano	1
53. 4,24,24-Trimetil-colestano	1
54. 1-Nor-ergostano	7
55. 1-Nor-9,19-Cicloergostano	1
56. 13,17-Secoergostano	1
57. 24-Propil-colestano	1
58. 9,11-Seco-colestano	16
59. 9,11-Seco-19-Nor-colestano	1
60. 5,6-Seco-colestano	1
61. 3,4-Seco-colestano	1
62. 4,5-Seco-colestano	1
63. 22-Metil-colestano	2
64. 23-Metil-colestano	2
65. Gorgostano	4
66. 30-Nor-gorgostano	2
67. 9,11-Secogorgostano	3
68. 9,11-Secoergostano	6
69. 22-Metil-campestano	2
70. 23-Metil-ergostano	3
71. 3,5-Cicloestigmastano	5
72. 14,15-Secopregnano	2
73. 3,5-Cicloergostano	3
74. 22-Metil-23,28-Ciclocolestano	2
75. 23,28-Cicloergostano	1
76. 23,28-Ciclocampestano	1
77. 24-Metil-28,30-Cicloporiferastano	1
78. 24-Metil-28,30-Cicloestigmastano	7
79. 26-Metil-26,27-Cicloergostano	1
80. 6,7-Secoestigmastano	4
81. 5,6-Secoergostano	1
82. 13-Etil-gonano	3
83. 11,18-Dimetil-19-Nor-pregnano	9
84. 26,27-Dinor-ergostano	5
85. 25-Metil-ergostano	4
86. 1,10-Secoergostano	2
87. 9,10-Secoergostano	1
88. 9,10-Secocolestano	3
89. 13,17-Secocolestano	4
90. 8,9-Secoergostano	2
91. 9,11-Secocampestano	2
92. 5,6-Secocampestano	2
93. 4-Metil-poriferastano	1

Tabela II. Continuação

Esqueletos	Número de ocorrências
94. 19-Nor-pregnano	3
95. 24-Metil-estigmastano	5
96. 24,24-Dimetil-colestano	3
97. 13,14-Seco-colestano	6
98. 21-Nor-ergostano	5
99. 24-Isopropil-25-Metil-colestano	6
100. 26,27-Dimetil-ergostano	2
101. 4-Metil-8,14-Secocampestano	1
102. 4-Metil-8,14-Secoporiferastano	1
103. 9,10-Seco-26,27-Dinor-ergostano	1
104. 9,11-Seco-26,27-Dinor-ergostano	5
105. 3,5-Ciclo-4,5-Seco-colestano	3
106. 26-Metil-26,27-Ciclocampestano	3
107. 16,23-Ciclocolestano	1
108. 1,2,3,4,5,10,19-Heptanor-colestano	5
109. 3,5-Ciclo-27-Norcampestano	1
110. 22-Metilergostano	2
111. 22-Metil-26,27-Dinorergostano	1
112. 28-Metilestigmastano	1
113. 19-Nor-vitanolídeo	1
114. 17(13→18)-Abeo-ergostano	7
115. 17(13→18)-Abeo-27-Nor-ergostano	2
116. Vitanolídeo-II	6
117. Vitanolídeo-III	2
118. Vitanolídeo-IV	3
119. Vitanolídeo-V	3

Ao iniciar-se o sistema PESQUISA é apresentada a tela do menu principal (Figura 16). O sistema e, conseqüentemente, as interfaces das telas foram escritos em Inglês para facilitar a posterior avaliação, distribuição e/ou comercialização dos programas nos mais diversos grupos de fitoquímica nacionais e internacionais.

Figura 16. Tela do menu principal do sistema PESQUISA

Através das opções 0 e 6 pode-se retornar ao sistema operacional de forma definitiva ou momentânea, respectivamente, assim como ocorre no programa DATASIS.

Através da opção 1, o sistema apresenta um sub-menu onde, através das respectivas opções, são introduzidos os dados de RMN ^{13}C , RMN ^1H , EM e dados sobre a origem natural de uma amostra. Após a entrada dos dados, retorna-se ao menu principal (Figura 16).

A partir da opção 2 do menu principal é possível pesquisar a presença de grupos substituintes em uma amostra, através do programa MACRONO. No caso da existência de um grupo substituinte, os sinais atribuídos a esse são automaticamente removidos da amostra para que somente sejam pesquisados pelos demais programas os sinais referentes ao esqueleto carbônico das substâncias.

3.3.2.1. O Programa MACRONÓ

Os produtos naturais, em geral, apresentam grupos substituintes, tais como, angeloíla, butiroíla e acetila, dentre outros, ligados ao esqueleto carbônico da substância por meio de um heteroátomo. Esses substituintes são, normalmente, provenientes da esterificação das hidroxilas presentes no esqueleto da substância, ou da esterificação destas com ácidos carboxílicos. Estes grupos, no SISTEMAT, são denominados de macronós.

Durante a análise espectral realizada pelos programas do SISTEMAT, a presença dos sinais pertencentes aos macronós pode confundir os algoritmos de identificação de esqueletos e estruturas, conduzindo a propostas incorretas e também à explosão combinatória. Para solucionar esse problema, foi desenvolvido o programa MACRONÓ⁶⁷, que fornece os possíveis substituintes ligados ao esqueleto carbônico de uma substância, a partir da análise dos dados do espectro de RMN ¹³C. Assim, antes que o espectro-problema seja submetido aos programas que realizam a identificação de esqueletos e fornecem propostas estruturais, é realizada uma análise prévia para identificação e remoção dos sinais dos substituintes.

Para o desenvolvimento deste programa, foi criado um banco de dados que apresenta as faixas de deslocamento químico dos carbonos dos 58 tipos de substituintes mais comuns encontrados em lactonas sesquiterpênicas⁷⁵. No entanto, quando o programa MACRONÓ foi utilizado para a identificação dos substituintes presentes em várias classes de produtos naturais apresentou baixo nível de desempenho em comparação ao observado para os sesquiterpenos lactonizados. A causa desse problema estava na base de dados do programa, que apresentava duas falhas principais:

a- durante o arquivamento dos substituintes, haviam sido utilizadas as faixas de deslocamentos químicos dos ácidos livres, que não correspondem às faixas dos respectivos ésteres;

b- a variedade de substituintes arquivada era muito pequena quando comparada à encontrada em terpenóides.

A solução encontrada para este problema foi a construção de uma nova base de dados que apresenta, atualmente, 173 tipos de substituintes freqüentemente encontrados em terpenóides. As estruturas dos macronós presentes na base de dados são apresentadas na Figura 17. A eficiência da base de dados foi intensamente avaliada para identificar grupos substituintes presentes em variadas classes de terpenóides⁷⁶.

Figura 17. Estruturas dos macronós presentes na base de dados

Figura 17. Continuação

Figura 17. Continuação

Figura 17. Continuação

Figura 17. Continuação

Oc1cc(O)c(O)c2c1O[C@H](O[C@@H]3[C@H](O)[C@H](O)[C@@H](O)[C@H]3O[C@H]2CO[C@@H]4[C@H](O)[C@H](O)[C@@H](O)[C@H]4O)[C@H](O)[C@H](O)[C@H]5O[C@H](CO[C@@H]6[C@H](O)[C@H](O)[C@@H](O)[C@H]6O)[C@H](O)[C@H](O)[C@H]5O

2'-[2"-hidroxi-6"-(*m*-hidroksifenil)-*p*-hidroksibenzoil]-glicosila

The image displays two chemical structures. The left structure is *p*-glicosil-*trans*-cafeoil, which consists of a glucose molecule in its pyranose form linked via its C1 position to the C4 position of a phenyl ring. The phenyl ring is substituted with a hydroxyl group at the C3 position and a *trans*-cafeoyl group at the C1 position. The *trans*-cafeoyl group is a propenoic acid derivative. The right structure is xilosila-(1->3)-galactosila (sinuatosila), which consists of a galactose molecule in its pyranose form linked via its C1 position to the C3 position of a xylose molecule in its furanose form.

The image displays four chemical structures of terpenoids, each with its name below it. The structures are drawn in a skeletal format. Each molecule features a carboxylate group (represented by a carbonyl group double-bonded to an oxygen and a single bond to an oxygen with a wavy line) and a hydroxyl group (OH). The side chains vary in their branching and functional groups.

- neroloila**: A branched chain with a double bond and a hydroxyl group.
- foliamentoila**: A branched chain with a double bond and a hydroxyl group.
- di-hidrofoliamentoila**: A branched chain with a double bond and a hydroxyl group.
- mentiafoloila**: A branched chain with a double bond, a hydroxyl group, and a methyl group.

Figura 17. Continuação

Figura 17. Continuação

Figura 17. Continuação

Figura 17. Continuação

Figura 17. Continuação

A partir dos deslocamentos químicos e multiplicidades dos sinais obtidos de um espectro de RMN ^{13}C problema, o programa MACRONÓ pesquisa quais os possíveis grupos substituintes presentes na estrutura da substância. Como resultado, o programa apresenta os nomes dos prováveis macronós, os sinais da amostra a ele atribuídos, o número de átomos de carbono do substituinte e a faixa de erro médio.

O número de átomos de carbono do substituinte é um dado importante pois se, por exemplo, o espectro-problema apresentar 34 sinais e for de um esteróide, isto indica que de seis a sete sinais não fazem parte do esqueleto carbônico, assim se o programa apresentar entre os possíveis macronós, um substituinte que apresente nove átomos de carbono, este pode ser automaticamente descartado. O substituinte mais provável será aquele que apresentar a menor faixa de erro médio aliada a um número de átomos de carbono compatível com os do espectro-problema, pois a faixa de erro médio é a média dos desvios entre os deslocamentos químicos do espectro-problema e os da base de dados.

Para demonstrar a utilização do sistema foi escolhido o esteróide de esqueleto ergostano, isolado de *Clavularia viridis* – Coral⁷⁷, apresentado na Figura 18.

Dados de RMN ^{13}C (CDCl_3): 207,7s; 129,1d; 141,0d; 36,4t; 76,9s; 75,3d; 29,2t; 29,2d; 47,1d; 53,7s; 68,4d; 51,1t; 43,1s; 55,8d; 23,8t; 28,3t; 55,9d; 13,6q; 14,6q; 36,3d; 18,6q; 33,5t; 30,6t; 39,1d; 31,5d; 20,5q; 17,6q; 15,5q; Ac: 169,9s; 21,3q

Figura 18. Esteróide utilizado para testar o sistema PESQUISA

Os resultados fornecidos pelo programa são apresentados na Figura 19, através da qual é possível verificar que o PESQUISA selecionou, com menor faixa de erro médio, o grupo substituinte presente na estrutura da substância. Os sinais referentes ao grupo acetoila são então retirados da análise, restando apenas os sinais que correspondem ao esqueleto carbônico do esteróide.

Figura 19. Grupos substituintes propostos pelo programa MACRONO

Através da opção 3 da Figura 16 (p. 41), é apresentado um sub-menu, Figura 20, onde são selecionadas as consultas aos demais programas aplicativos do SISTEMAT.

Figura 20. Menu de programas aplicativos do sistema

3.3.2.2. O Programa SISCONST

O programa SISCONST⁷⁸ foi desenvolvido para auxiliar o químico no processo de determinação estrutural de produtos naturais, a partir dos deslocamentos químicos e multiplicidades dos sinais do espectro de RMN ^{13}C . Esse programa prevê o tipo de esqueleto mais provável para uma substância-problema e fornece grandes subestruturas, com a atribuição dos sinais de RMN ^{13}C compatíveis com os dados espectrais.

O programa confronta os sinais do espectro questionado com todos os espectros pertencentes à base de dados. Se um sinal do espectro questionado está presente, por exemplo, no espectro da primeira substância do banco e tem a mesma multiplicidade, os deslocamentos químicos dos carbonos vizinhos interligados são confrontados com os do espectro-problema. Esse processo se repete até que sejam obtidos os maiores fragmentos das substâncias da base de dados que apresentem deslocamentos químicos compatíveis com os dados do espectro-problema.

O usuário deve fornecer os seguintes dados para o programa SISCONST:

- deslocamentos químicos e multiplicidades do espectro experimental;
- o número mínimo de átomos para proposição de subestruturas;
- a faixa de erro admitida pelo programa, geralmente, δ 1,0;
- o gradiente, que permite o aumento automático da faixa de erro;
- o espaço de busca da pesquisa, ou seja, se esta deve ser realizada sobre todos os compostos da classe ou apenas sobre os compostos de um esqueleto específico.

Na busca do provável esqueleto, o programa somente escolhe substâncias que apresentem subestruturas com pelo menos metade do número total de carbonos. A partir das substâncias selecionadas, é calculada a probabilidade de esqueleto do espectro-problema. Para a realização das pesquisas foi fixado como intervalo de erro, no deslocamento químico de RMN ^{13}C , 1,0 δ para subestruturas e 5,0 δ para esqueletos. Se a busca não for bem sucedida, os intervalos são automaticamente aumentados de acordo com um gradiente de 0,5 e 1,0 δ para subestruturas e esqueletos, respectivamente. Caso esse intervalo alcance valor maior que 3,0 δ para subestruturas e 10,0 δ para esqueletos, o programa cessa a busca e não apresenta previsão de subestruturas e/ou esqueletos.

Para exemplificar a utilização do programa SISCONST foi utilizado o espectro de RMN ^{13}C do esteróide de esqueleto ergostano apresentado na Figura 18. Após a escolha dos parâmetros, o programa SISCONST inicia a pesquisa de subestruturas e esqueletos.

O programa encontrou quatro subestruturas, as quais são apresentadas uma a uma para o usuário, que seleciona quais são boas propostas de estrutura. Após a seleção, o programa separa os deslocamentos químicos atribuídos a essas propostas e não os utiliza em uma análise posterior. Nas Figuras 21 e 22 são apresentadas as duas primeiras subestruturas localizadas pelo programa SISCONST; as demais subestruturas não são apresentadas uma vez que se tratam de subestruturas menores, que já se encontram incorporadas às subestruturas das Figuras 21 e 22.

Ao término da análise, o usuário tem a opção de fazer uma nova consulta somente com os deslocamentos químicos não atribuídos às estruturas anteriormente selecionadas. Para o teste em questão, essa opção foi requerida e, através da nova análise realizada, o programa SISCONST forneceu a subestrutura exibida na Figura 23.

Nas Figuras 21-23 pode-se visualizar, no lado direito da tela, a atribuição de sinais de RMN ^{13}C realizada pelo programa. Os deslocamentos químicos experimentais encontram-se marcados na cor azul, enquanto os δ em branco referem-se à estrutura presente na base de dados, obtida da respectiva referência citada acima da atribuição dos sinais.

Figura 21. Primeira subestrutura proposta pelo programa SISCONST

Figura 22. Segunda subestrutura proposta pelo programa SISCONST

Figura 23. Subestrutura apresentada após a nova análise pelo programa SISCONST

Através da sobreposição das subestruturas apresentadas, Figuras 21-23, é obtida uma subestrutura maior, compatível com a estrutura proposta na literatura⁷⁷, e na qual faltam apenas a carbonila da posição 1 e os carbonos metilênicos das posições 7 e 12, para completar a estrutural final da substância.

Após a identificação das subestruturas, o programa SISCONST fornece ainda a probabilidade de esqueleto, que é armazenada pelo programa para cálculo da probabilidade final. A partir da opção 2 (Run C13MACH) da Figura 20, inicia-se a utilização do próximo programa aplicativo do sistema.

3.3.2.3. O Programa C13MACH

O programa C13MACH^{67,70} prevê, a partir dos deslocamentos químicos e multiplicidades dos sinais do espectro de RMN ¹³C experimental, o tipo de esqueleto da substância e seleciona as *x* substâncias (*x* pode variar de 1 a 50) do banco de dados que apresentam maior índice de similaridade⁷⁹ com a substância testada, dentro da faixa de erro estabelecida. No entanto, para selecionar uma estrutura o programa não requer a condição de que os átomos de carbono da amostra estejam interligados, o que pode conduzir a falsos resultados.

Para testar o programa foram utilizados os dados do esteróide da Figura 18. Foi estabelecido 1,08 como faixa de erro a ser admitida pelo programa para o deslocamento químico de RMN ¹³C e em dez o número de substâncias a serem selecionadas, com melhores índices de semelhança. Após a definição desses parâmetros, o programa processa as informações e apresenta as estruturas de maior índice de semelhança com o espectro em questão. Na Figura 24 é mostrada a estrutura que apresenta o maior índice de similaridade e, também, no lado direito da tela, o número de carbonos da estrutura, o índice de semelhança, a referência bibliográfica e a provável classe química à qual pertence o espectro-problema.

Ao final da análise, o programa apresenta a probabilidade de esqueleto para os dados questionados, que é armazenada pelo sistema para cálculo da probabilidade final.

As opções 3 e 4 da Figura 20 permitem acessar os programas H1MACH e MASMACH, respectivamente. O programa H1MACH realiza a análise dos dados de RMN ¹H enquanto que o programa MASMACH analisa dados provenientes de EM. Os

dois programas não foram utilizados nesse trabalho, pois os respectivos dados que analisam não foram arquivados no sistema.

Figura 24. Estrutura com maior índice de similaridade apresentada pelo programa C13MACH

3.3.2.4. O Programa SISOCBOT

O programa SISOCBOT^{66,80}, opção 5 da Figura 20, permite a análise dos dados de origem natural utilizando, como dado de entrada obrigatório, o nome da família na qual se deseja realizar a pesquisa, sendo o gênero e a espécie dados optativos. Como resultado da análise, o programa fornece o percentual de ocorrência dos esqueletos na família e/ou gênero escolhidos. Essa informação foi testada para diversas classes de produtos naturais⁸⁰ e pode ser utilizada como restrição no processo de determinação estrutural, isto é, esses percentuais podem indicar quais tipos de esqueletos são encontrados frequentemente em um determinado taxon sendo, assim, úteis como auxiliares em sistemas especialistas^{66,81}. O programa foi também testado com o esteróide da Figura 18 e a probabilidade de esqueleto armazenada pelo sistema.

3.3.2.5. O Programa REGRAS

O programa REGRAS é um novo programa desenvolvido, durante esse trabalho, para o sistema especialista SISTEMAT. Esse programa realiza a identificação de esqueletos através da desfuncionalização dos dados de RMN ^{13}C . Para tanto, a partir dos esqueletos encontrados durante a revisão bibliográfica, foi criado um arquivo que apresenta os seguintes dados:

- a- o nome de cada esqueleto presente no banco de dados;
- b- o número e tipos de átomos de carbono de cada esqueleto, isto é, o número total de carbonos de cada esqueleto e o número de carbonos quaternários, metínicos, metilênicos e metílicos.

Na Figura 25, é apresentado um exemplo do cálculo dos tipos de átomos de carbono presentes no esqueleto ergostano e, na Tabela III, são apresentados os tipos de átomos de carbono para os esqueletos de esteróides. O arquivo criado para o programa REGRAS pode ainda conter as faixas de deslocamentos químicos características para cada esqueleto e/ou subesqueleto estudado, as quais serão posteriormente descritas e apresentadas.

Substância : Yonarasterol B

Esqueleto : Ergostano

Para o esqueleto, número de: C_{total} : 28; CH_3 : 6; CH_2 : 12; CH : 8; C : 2

Figura 25. Exemplo do cálculo dos tipos de átomos de carbono

Tabela III. Relação do número total de carbonos e dos números de carbonos quaternários, metínicos, metilênicos e metílicos dos esqueletos de esteróides presentes no banco de dados

Esqueletos	No. C	C	CH	CH ₂	CH ₃
Gonano	17	0	6	11	0
Estrano	18	1	5	11	1
13-Etil-gonano	19	1	5	12	1
1-Metilestrano	19	1	6	10	2
4-Metilestrano	19	1	6	10	2
7-Metilestrano	19	1	6	10	2
16-Metilestrano	19	1	6	10	2
Androstano	19	2	4	11	2
19(10→5)-Androstano	19	2	4	11	2
1,19-Cicloandrostano	19	2	5	11	1
3,5-Abeo-3,4-Secoandrostano	19	3	3	10	3
5,19-Cicloandrostano	19	3	3	12	1
3,5-Cicloandrostano	19	3	4	10	2
Heptanor-Colestano	20	1	5	9	5
19-Nor-pregnano	20	1	6	11	2
1-Metilandrostando	20	2	5	10	3
2-Metilandrostando	20	2	5	10	3
3-Metilandrostando	20	2	5	10	3
16-Metilandrostando	20	2	5	10	3
17-Metilandrostando	20	2	5	10	3
16,16-Dimetilestrano	20	2	5	10	3
4-Metil-5,19-Cicloandrostano	20	3	4	11	2
14,15-Secopregnano	21	2	4	11	4
Pregnano	21	2	5	11	3
16,16-Dimetil-androstano	21	3	4	10	4
2,2-Dimetil-androstano	21	3	4	10	4
17,18-Ciclopregano	21	3	4	12	2
11,18-Dimetil-19-Nor-pregnano	22	1	7	11	3
16-Metil-pregnano	22	2	6	10	4
20-Metil-pregnano	22	2	6	10	4
Cardanolídeo	23	2	6	11	4
3,20-Dimetil-pregnano	23	2	7	9	5
Colano	24	2	6	12	4
26,27-Dinorcolestando	25	2	6	13	4
9,10-Seco-26,27-Dinorergostano	26	1	7	13	5
19-Nor-9,11-Secocolestando	26	1	7	13	5
9,11-Seco-26,27-Dinorergostano	26	2	6	12	6
27-Norcolestando	26	2	6	14	4
26,27-Dinorergostano	26	2	7	12	5
22-Metil-26,27-Dinorergostano	26	2	7	12	5
9,10-Secocolestando	27	1	7	14	5
13,14-Secocolestando	27	1	7	14	5
13,17-Secocolestando	27	1	7	14	5

Tabela III. Continuação

Esqueletos	No. C	C	CH	CH ₂	CH ₃
17(13→18)-Abeo-27-Norergostano	27	1	8	14	4
1-Norergostano	27	1	9	10	7
19-Norergostano	27	1	9	12	5
19-Norvitanolídeo	27	1	9	12	5
4,5-Secocolestano	27	2	6	13	6
5,6-Secocolestano	27	2	6	13	6
9,11-Secocolestano	27	2	6	13	6
7-Feniletil-androstano	27	2	6	17	2
3,4-Secocolestano	27	2	7	11	7
Colestano	27	2	7	13	5
21-Norergostano	27	2	7	13	5
3,5-Ciclo-4,5-secocolestano	27	2	8	11	6
1-Nor-9,19-Cicloergostano	27	2	8	11	6
5,7-Abeo-5,6-Seco-colestano	27	2	8	11	6
16,23-Ciclocolestano	27	2	9	11	5
3,5-Ciclocolestano	27	3	7	12	5
3,5-Ciclo-27-Nor-campestano	27	3	7	12	5
13,17-Secoergostano	28	1	8	13	6
9,10-Secoergostano	28	1	8	13	6
1,10-Secoergostano	28	1	9	11	7
17(13→18)-Abeo-ergostano	28	1	9	13	5
1,10-Seco-1,6-Cicloergostano	28	1	10	11	6
8,9-Secoergostano	28	2	6	14	6
5,6-Secoergostano	28	2	7	12	7
9,11-Secoergostano	28	2	7	12	7
9,11-Secocampestano	28	2	7	12	7
5,6-Secocampestano	28	2	7	12	7
2-Metilcolestano	28	2	8	12	6
3-Metilcolestano	28	2	8	12	6
6-Metilcolestano	28	2	8	12	6
Ergostano	28	2	8	12	6
Campestano	28	2	8	12	6
Vitanolídeo-I	28	2	8	12	6
22-Metilcolestano	28	2	8	12	6
23-Metilcolestano	28	2	8	12	6
Vitanolídeo-V	28	2	8	12	6
23,28-Ciclocampestano	28	2	9	12	5
23,28-Cicloergostano	28	2	9	12	5
22-Metil-23,28-Ciclocolestano	28	2	9	12	5
12,23-Cicloergostano	28	2	10	10	6
Vitanolídeo-II	28	2	10	10	6
5-Metilcolestano	28	3	6	13	6
Vitanolídeo-III	28	3	7	13	5
Vitanolídeo-IV	28	3	8	11	6
3,5-Cicloergostano	28	3	8	11	6
6,7-Secoestigmastano	29	2	8	11	8
Estigmastano	29	2	8	13	6

Tabela III. Continuação

Esqueletos	No. C	C	CH	CH ₂	CH ₃
Poriferastano	29	2	8	13	6
4-Metil-8,14-Secocampestano	29	2	7	13	7
4-Metilergostano	29	2	9	11	7
22-Metil-campestano	29	2	9	11	7
23-Metilergostano	29	2	9	11	7
22-Metil-ergostano	29	2	9	11	7
26-Metil-26,27-Cicloergostano	29	2	9	13	5
26-Metil-26,27-Ciclocampestano	29	2	9	13	5
23,28-Cicloestigmastano	29	2	10	11	6
30-Norgorgostano	29	2	10	11	6
25-Metilergostano	29	3	7	12	7
Fusidano	29	3	7	12	7
24,24-Dimetilcolestano	29	3	7	12	7
3,5-Cicloestigmastano	29	3	8	12	6
4-Metil-8,14-Secoporiferastano	30	2	7	14	7
24-Metilestigmastano	30	2	8	14	6
24-Propil-colestano	30	2	8	14	6
26,27-Dimetilergostano	30	2	8	14	6
4-Metil-poriferastano	30	2	9	12	7
28-Metilestigmastano	30	2	9	12	7
14-Metil-poriferastano	30	3	7	13	7
14-Metil-poriferastano	30	3	7	13	7
4,14-Dimetil-campestano	30	3	8	11	8
4,14-Dimetil-ergostano	30	3	8	11	8
4,24,24-Trimetil-colestano	30	3	8	11	8
9,11-Secogorgostano	30	3	8	11	8
24-Metil-28,30-Cicloprioriferastano	30	3	8	13	6
24-Metil-28,30-Cicloestigmastano	30	3	8	13	6
Gorgostano	30	3	9	11	7
24-Isopropil-25-Metil-colestano	31	3	8	12	8

A partir dos dados de RMN ¹³C de centenas de substâncias arquivadas no sistema especialista SISTEMAT, foram obtidas faixas de deslocamentos químicos normalmente observadas para um dado grupo funcional. Estas faixas de deslocamentos químicos são apresentadas na Tabela IV, juntamente com a multiplicidade inicial dos sinais que estas apresentam no espectro de RMN ¹³C e a multiplicidade final dos sinais que estas devem apresentar após a desfuncionalização dos dados. Esta informação, aliada aos tipos de átomos de carbono existentes em cada esqueleto, foram arquivadas em bases de dados distintas no programa REGRAS.

Para exemplificar o modo de operação e utilização do programa REGRAS foi utilizado o esteróide da Figura 18. Após a entrada dos dados espectrais é iniciada a

pesquisa de esqueletos no programa REGRAS, o qual realiza a desfuncionalização dos dados de RMN ^{13}C através das faixas apresentadas na Tabela IV, para posterior comparação com os tipos de carbono dos esqueletos obtidos na literatura (Tabela III).

Tabela IV. Faixas de deslocamentos químicos utilizadas para desfuncionalização dos dados de RMN ^{13}C

Função Química	Multiplicidade Inicial	Multiplicidade Final	Faixas de deslocamento químico
C=O	1	3	190,0 - 250,0
CHO	2	4	190,0 - 250,0
R-COO-R'	1	4	165,5 - 190,0
C=C	1	2	113,0 - 165,5
	2	3	104,0 - 167,0
	3	4	100,0 - 167,0
C-OH	1	2	61,0 - 100,0
	2	3	58,0 - 90,0
	3	4	54,0 - 90,0
C(OR) ₂	1	3	100,0 - 113,0
CH(OR) ₂	2	4	90,0 - 104,0

O programa REGRAS, inicialmente, apresenta uma tela (Figura 26) com o deslocamento químico que será desfuncionalizado e a faixa na qual esse se encaixa com a respectiva multiplicidade, além de mostrar a função existente antes da desfuncionalização e o tipo de carbono que será gerado após a desfuncionalização. Na Figura 26, esse passo corresponde à “desfuncionalizando C=O para CH₂”. No final da tela, o programa ainda apresenta o número de ocorrências de cada multiplicidade no espectro não desfuncionalizado e o número total de carbonos. Esse processo repete-se até que todos os deslocamentos químicos que se encontram nas faixas apresentadas na Tabela IV sejam completamente desfuncionalizados.

Após a desfuncionalização completa do espectro de RMN ^{13}C , o REGRAS apresenta o número de ocorrências de cada multiplicidade no espectro desfuncionalizado e solicita a faixa de tolerância para a pesquisa nas faixas de deslocamentos químicos características para cada esqueleto e/ou subesqueleto estudado

(Figura 27). Prosseguindo a análise, o programa pesquisa o esqueleto por tipos de carbonos encontrados após o término da desfuncionalização e armazena a probabilidade de esqueleto encontrada para a substância. O programa REGRAS foi testado com substâncias das mais variadas classes de terpenóides e exibiu resultados excelentes durante o processo de previsão de esqueletos⁸².

Figura 26. Tela inicial de desfuncionalização de dados de RMN ¹³C

Após a análise realizada pelos programas, as previsões de esqueleto são combinadas através da opção 4 (Global Results) do menu principal do programa PESQUISA (Figura 16), resultando em uma probabilidade global, que é a média ponderada das probabilidades fornecidas pelos programas utilizados, sendo atribuídos, com base em resultados prévios, os seguintes pesos aos programas: SISCONST – 0,9; C13MACH – 0,7; SISOCBOT – 0,5 e REGRAS – 1,0. Na Figura 28, é apresentada a tela final do programa PESQUISA com a probabilidade global de esqueleto para o esteróide da Figura 18.

Outros testes foram realizados com os programas aplicativos do SISTEMAT e serão apresentados em Resultados.

Figura 27. Tela final de desfuncionalização de dados de RMN ^{13}C

Prompt do MS-DOS - PESQUISA

Auto

SKELETONS PROBABILITY (%)

SKELETON	1	2	3	4	5	6	7
1) ERGOSTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
2) COLESTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
3) CAMPESTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
4) 22ME-23,22CICLO-ERGOSTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
5) ESTIGMASTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
6) 3,5CICLO-ESTIGMASTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
7) WITHANOLIDE	72.0	6.4	87.0	0.0	0.0	100.0	81.5
8) PORIFERASTANO	72.0	6.4	87.0	0.0	0.0	100.0	81.5
9) OTHER	1.8	0.0	0.0	0.0	0.0	0.0	0.0

1-SISCONST; 2-C13MACH; 3-SISOCBOT; 4-NASHACH; 5-HINACH; 6-DESFUNC.; 7-GLOBAL

Figura 28. Probabilidade global de esqueleto calculada pelo sistema PESQUISA

IV - MÉTODOS AUTOMÁTICOS DE PREVISÃO DE DADOS DE RMN ^{13}C

A Ressonância Magnética Nuclear de ^{13}C é uma das ferramentas mais poderosas para a elucidação ou verificação da estrutura de substâncias orgânicas. Regularmente, espectroscopistas experientes necessitam de meios para estimar a posição de sinais no espectro, uma vez que os deslocamentos químicos de RMN ^{13}C são sensíveis à presença de grupos substituintes próximos ao átomo de carbono analisado. Assim, um procedimento geral e confiável para estimar deslocamentos químicos é extremamente útil tanto para o espectroscopista quanto para validar propostas de um gerador de estruturas. Além disso, a simulação espectral pode auxiliar o químico na solução de problemas complexos em determinação estrutural bem como na verificação da atribuição de deslocamentos químicos, particularmente quando espectros de referência não são disponíveis.

Os três métodos mais comuns de simular dados espectrais de RMN ^{13}C envolvem relações lineares de aditividade, consultas à base de dados espectroscópicas e modelos empíricos.

A simulação espectral por relações lineares de aditividade tem sido bastante investigada⁸³⁻⁸⁶. Essa técnica utiliza uma base de dados de átomos de carbono e seus respectivos deslocamentos químicos para extrair parâmetros empíricos para uma variedade de estruturas e funcionalizações. O deslocamento químico de um novo átomo de carbono, que apresenta um ambiente químico particular, pode ser calculado através de equações lineares simples, usando os parâmetros empíricos obtidos a partir de ambientes químicos similares. Devido à aplicação trabalhosa da técnica, têm sido desenvolvidos vários programas com essa proposição⁸⁷⁻⁹¹. A vantagem desse método é que este pode ser aplicado a uma ampla faixa de estruturas. No entanto, faixas de erro iguais ou superiores a 5δ são comuns, o que pode atrapalhar a interpretação desejável dos dados simulados. Além disso, parâmetros que descrevem a configuração e a conformação das moléculas não são considerados, sendo essas as principais desvantagens do método.

Similarmente, técnicas de consulta à base de dados espectroscópicas estimam o deslocamento químico de átomos de carbono de uma substância desconhecida, extraíndo o deslocamento do átomo mais similar contido na base de dados, para cada carbono da substância⁹²⁻⁹⁶. Embora essa metodologia possa ser aplicada a um grande

número de substâncias orgânicas, a exatidão do resultado é altamente dependente do tamanho e da variedade da base de dados, bem como do grau de similaridade utilizado.

Em contraste, as técnicas de modelos empíricos são tipicamente usadas para desenvolver previsões de deslocamentos químicos exatas, para uma classe de átomos pequena e homogênea. As previsões são baseadas no desenvolvimento de modelos lineares matemáticos a partir de um conjunto de deslocamentos químicos conhecidos. Esses modelos relacionam o deslocamento químico de um átomo com vários parâmetros estruturais do átomo (descriptores) e têm a seguinte forma: $S = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n$ onde S é o deslocamento químico de um átomo de carbono, os X_i são os valores dos descriptores e os b_i são os coeficientes determinados por regressão linear múltipla. Uma vez conhecidos, esses modelos podem ser utilizados para prever os deslocamentos químicos de carbonos não contidos no conjunto de dados original⁹⁷⁻⁹⁹. A limitação dessa técnica surge porque as equações-modelo podem ser usadas somente para prever deslocamentos químicos de átomos que são estruturalmente similares aos átomos usados para desenvolver os modelos. Entretanto, ao contrário das técnicas de consulta à base de dados, esses modelos têm a capacidade para interpolar deslocamentos químicos. Além disso, essa técnica permite simular dados de espectro com erros tipicamente na faixa de 1,0–1,5δ.

Recentemente, as análises por regressão linear múltipla estão sendo combinadas ou substituídas por outras técnicas numéricas e estatísticas tais como a análise de componentes principais^{100,101} e as redes neurais¹⁰²⁻¹⁰⁴, as quais também estão sendo utilizadas para previsão de dados de deslocamentos químicos de RMN ¹H^{105,106}. As principais classes químicas já estudadas através de modelos empíricos são as seguintes: alcanos lineares e ramificados¹⁰⁷, cicloalcanos¹⁰⁸, cicloalcanóis¹⁰⁹, hidroxí-esteróides¹¹⁰, ciclopentanos e ciclopentanóis¹¹¹, nor-bornanóis¹¹², ciclohexanonas e ciclodecalonas¹¹³, piperidinas¹¹⁴, bifenilas policloradas¹¹⁵, benzenos alquil substituídos e poliaromáticos¹¹⁶, ceto-esteróides¹¹⁷, quinolinas e isoquinolinas¹¹⁸, ciclopentanonas e cicloheptanonas¹¹⁹, alcalóides indólicos¹²⁰, polissacarídeos¹²¹, tetrahidropiranos¹²², trissacarídeos¹²³, monossacarídeos¹²⁴, ribonucleosídeos¹²⁵ e dibenzofuranos¹²⁶.

No entanto, nos três métodos abordados são raros os casos onde são considerados parâmetros que descrevem a configuração e a conformação das moléculas. Isso também se verifica em programas comerciais desenvolvidos para o mesmo fim. Sendo assim, decidimos desenvolver um simulador de dados de RMN ¹³C que suprimisse as falhas observadas nos análogos comerciais existentes.

4.1. O SIMULADOR TRIDIMENSIONAL DE DADOS DE RMN ^{13}C

A proposta de construção de um simulador tridimensional de dados de RMN ^{13}C originou-se devido à influência exercida pela proximidade espacial de um grupo funcional sobre o deslocamento químico de um átomo de carbono. Aliada à influência espacial temos, ainda, a influência da conformação da molécula sobre os deslocamentos. Como esses dois fatores não são contemplados em um simulador bidimensional, o qual abrangeria apenas os átomos α , β , γ e δ ligados a um determinado carbono, nossa opção foi o desenvolvimento de um simulador tridimensional, que associa a estereoquímica, a conformação e a influência espacial dos grupos funcionais a deslocamento químico e seus respectivos átomos diretamente ligados.

4.1.1. Metodologia

Os esteróides obtidos através da revisão bibliográfica são inicialmente desenhados para obtenção de sua respectiva estrutura tridimensional. Após esta etapa, as estruturas são codificadas no simulador de dados de RMN ^{13}C e, a cada átomo de carbono presente na estrutura, é associado seu respectivo deslocamento químico que, posteriormente, servirá de base para a previsão dos dados. Os solventes utilizados na obtenção dos espectros, assim como o esqueleto e a referência bibliográfica também foram considerados no armazenamento dos dados.

4.1.2. A obtenção de modelos tridimensionais

As substâncias inseridas no simulador de dados de RMN ^{13}C são inicialmente desenhadas no programa HyperChem¹²⁷. Para cada substância este programa busca as conformações que levam a molécula a seu estado de menor energia. Para esse procedimento são usados cálculos de mecânica molecular¹²⁸.

Para exemplificar a utilização do programa HyperChem e do simulador utilizaremos o esteróide colesterol, o qual é apresentado na tela do programa HyperChem (Figura 29). As ligações em azul representam ligações carbono-carbono e, em vermelho, representam ligações carbono-oxigênio. Após o desenho da substância, adicionam-se os átomos de hidrogênio e inicia-se o processo de obtenção da estrutura de menor energia, a qual é apresentada na Figura 30, para o colesterol.

Figura 29. Tela do programa HyperChem

Figura 30. Estrutura de menor energia para o colesterol

Ao término desse processo, o programa HyperChem gera um arquivo que contém, para cada átomo da substância, as suas respectivas coordenadas x,y,z. Na Figura 31, é possível visualizar parte do arquivo gerado para a estrutura do colesterol onde, a partir da sexta linha do arquivo, são descritas as coordenadas para cada átomo da estrutura. Utilizando como exemplo o átomo 1, verifica-se que este é um átomo de carbono e que apresenta coordenadas x,y,z respectivamente em -2,562937, 0,315651 e 1,19365. A seguir é indicado o número de ligantes presentes no átomo em questão (para o átomo 1, 4 ligantes) e os números dos átomos que se encontram ligados ao respectivo átomo, bem como o tipo de ligação entre os mesmos (Fig. 31, átomos 2, 10, 34 e 35 todos apresentando ligações simples (s)).

```

forcefield mm+
sys 0 0 1
view 40 0.1342 64.87963 24.87963 0.7066959 -0.2615143 0.6574125 0.672858 0.5356729 -0.5102123
seed -1084
mol 1
atom 1 - C ** - 0 -2.562937 0.315651 -1.19365 4 2 s 10 s 34 s 35 s
atom 2 - C ** - 0 -3.668547 -0.703873 -0.8120853 4 1 s 3 s 36 s 37 s
atom 3 - C ** - 0 -3.061547 -2.106332 -0.5569252 4 2 s 4 s 28 s 72 s
atom 4 - C ** - 0 -1.995755 -2.028907 0.5663308 4 3 s 5 s 38 s 39 s
atom 5 - C ** - 0 -0.9493289 -0.9751011 0.2427825 3 4 s 6 d 10 s
atom 6 - C ** - 0 0.3497142 -1.322478 0.3484643 3 5 d 7 s 74 s
atom 7 - C ** - 0 1.500584 -0.3948371 0.06096879 4 6 s 8 s 40 s 41 s
atom 8 - C ** - 0 1.089127 1.101103 0.04444306 4 7 s 9 s 14 s 29 s
atom 9 - C ** - 0 -0.2264757 1.257556 -0.7863932 4 8 s 10 s 11 s 30 s
atom 10 - C ** - 0 -1.410608 0.4384123 -0.1342344 4 9 s 1 s 5 s 19 s
atom 11 - C ** - 0 -0.5744719 2.760211 -1.05966 4 9 s 12 s 42 s 43 s
atom 12 - C ** - 0 0.6228232 3.589068 -1.611195 4 11 s 13 s 44 s 45 s
atom 13 - C ** - 0 1.869037 3.485944 -0.6772864 4 12 s 14 s 17 s 18 s
atom 14 - C ** - 0 2.208207 1.961745 -0.6076915 4 13 s 8 s 15 s 31 s
atom 15 - C ** - 0 3.613188 1.914273 0.03136156 4 14 s 16 s 46 s 47 s
atom 16 - C ** - 0 4.309337 3.177867 -0.5404145 4 15 s 17 s 48 s 49 s
atom 17 - C ** - 0 3.21858 4.029781 -1.280328 4 16 s 13 s 20 s 32 s
atom 18 - C ** - 0 1.579219 4.126316 0.7233894 4 13 s 50 s 51 s 52 s
atom 19 - C ** - 0 -1.950591 1.129264 1.162385 4 10 s 53 s 54 s 55 s
atom 20 - C ** - 0 3.515908 5.571136 -1.197778 4 17 s 21 s 22 s 33 s
atom 21 - C ** - 0 2.413252 6.417917 -1.897778 4 20 s 56 s 57 s 58 s
atom 22 - C ** - 0 4.907057 5.893585 -1.841635 4 20 s 23 s 59 s 60 s
atom 23 - C ** - 0 5.442302 7.308022 -1.480054 4 22 s 24 s 61 s 62 s

```

Figura 31. Arquivo gerado pelo HyperChem para a estrutura do colesterol

A partir dessas coordenadas, o programa simulador calcula as distâncias entre um átomo e os demais átomos da substância através da equação apresentada abaixo e, através das distâncias obtidas, realiza a codificação das estruturas.

$$\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

Como todo o procedimento de codificação das substâncias utiliza as distâncias calculadas através dos dados provenientes da estrutura minimizada pelo programa HyperChem, é imprescindível, nesta fase do trabalho, validar os resultados apresentados pelo programa, isto é, assegurar que a estrutura minimizada e, conseqüentemente, as distâncias calculadas entre os átomos, correspondem às distâncias existentes na molécula. Desta forma, pode-se garantir que a base de dados não foi construída em um modelo tridimensional incorreto.

Para tanto, foram selecionados na literatura 12 esteróides¹²⁹⁻¹³⁹ (Figura 33) que apresentam as distâncias entre os átomos da estrutura obtidas por análise de difração de Raios-X. Essas substâncias foram desenhadas e tiveram suas energias minimizadas através do programa HyperChem, sendo as distâncias calculadas para os átomos da estrutura comparadas com as distâncias experimentais obtidas por Raios X, como apresentado na Tabela V. A partir desses dados, foi realizada uma análise por regressão linear das distâncias previstas *versus* as distâncias experimentalmente observadas, onde se obteve um coeficiente de correlação (R) de 0,9771, como apresentado na Figura 32, com um F= 7258,92, o que demonstra uma boa correlação entre os dados.

Figura 32. Distâncias observadas experimentalmente *versus* distâncias previstas

Figura 33. Esteróides utilizados para validar o modelo tridimensional do simulador

Tabela V. Distâncias experimentais e previstas para os esteróides

I	Raio-X	Teórico	II	Raio-X	Teórico	III	Raio-X	Teórico
C ₁ -C ₂	1,348	1,342	C ₃ -O	1,430	1,439	C ₁ -C ₂	1,531	1,547
C ₁ -C ₁₀	1,515	1,520	C ₄ -O	1,424	1,437	C ₁ -C ₁₀	1,552	1,575
C ₂ -C ₃	1,465	1,462	C ₁₇ -O	1,204	1,231	C ₂ -C ₃	1,499	1,503
C ₂ -Br	1,885	1,873	C ₁ -C ₂	1,504	1,548	C ₃ -O	1,233	1,231
C ₃ -C ₄	1,485	1,468	C ₁ -C ₁₀	1,529	1,571	C ₃ -C ₄	1,448	1,466
C ₃ -O	1,211	1,231	C ₂ -C ₃	1,507	1,549	C ₄ -C ₅	1,345	1,349
C ₄ -C ₅	1,362	1,351	C ₃ -C ₄	1,504	1,555	C ₅ -C ₆	1,510	1,517
C ₄ -Br	1,859	1,878	C ₄ -C ₅	1,531	1,570	C ₅ -C ₁₀	1,519	1,537
C ₅ -C ₆	1,495	1,518	C ₅ -C ₆	1,530	1,560	C ₆ -C ₇	1,534	1,546
C ₅ -C ₁₀	1,517	1,532	C ₅ -C ₁₀	1,553	1,580	C ₇ -C ₈	1,526	1,554
C ₆ -C ₇	1,539	1,549	C ₆ -C ₇	1,522	1,551	C ₈ -C ₉	1,548	1,583
C ₇ -C ₈	1,527	1,555	C ₇ -C ₈	1,520	1,553	C ₈ -C ₁₄	1,520	1,555
C ₈ -C ₉	1,551	1,574	C ₈ -C ₁₄	1,510	1,553	C ₉ -O	1,437	1,452
C ₈ -C ₁₄	1,531	1,553	C ₈ -C ₉	1,536	1,575	C ₉ -C ₁₀	1,578	1,597
C ₉ -C ₁₀	1,558	1,572	C ₉ -C ₁₁	1,544	1,571	C ₉ -C ₁₁	1,544	1,577
C ₉ -C ₁₁	1,538	1,567	C ₉ -C ₁₀	1,559	1,581	C ₁₀ -C ₁₉	1,563	1,572
C ₁₀ -O	1,429	1,443	C ₁₀ -C ₁₉	1,545	1,571	C ₁₉ -O	1,405	1,436
C ₁₁ -C ₁₂	1,532	1,558	C ₁₁ -C ₁₂	1,523	1,559	C ₁₁ -C ₁₂	1,529	1,560
C ₁₂ -C ₁₃	1,525	1,558	C ₁₂ -C ₁₃	1,506	1,552	C ₁₂ -C ₁₃	1,523	1,553
C ₁₃ -C ₁₄	1,545	1,559	C ₁₃ -C ₁₇	1,512	1,513	C ₁₃ -C ₁₄	1,532	1,558
C ₁₃ -C ₁₇	1,530	1,562	C ₁₃ -C ₁₈	1,531	1,564	C ₁₃ -C ₁₇	1,509	1,513
C ₁₃ -C ₁₈	1,533	1,567	C ₁₃ -C ₁₄	1,539	1,558	C ₁₃ -C ₁₈	1,548	1,564
C ₁₄ -C ₁₅	1,527	1,548	C ₁₄ -C ₁₅	1,527	1,550	C ₁₄ -C ₁₅	1,533	1,551
C ₁₅ -C ₁₆	1,557	1,556	C ₁₅ -C ₁₆	1,528	1,557	C ₁₅ -C ₁₆	1,534	1,557
C ₁₆ -C ₁₇	1,529	1,560	C ₁₆ -C ₁₇	1,513	1,514	C ₁₆ -C ₁₇	1,519	1,515
C ₁₇ -O	1,441	1,435						

IV	Raio-X	Teórico	V	Raio-X	Teórico	VI	Raio-X	Teórico
C ₁ -C ₂	1,521	1,547	C ₁ -C ₂	1,391	1,404	C ₁ -C ₂	1,524	1,552
C ₁ -C ₁₀	1,531	1,571	C ₁ -C ₁₀	1,386	1,411	C ₁ -C ₁₀	1,554	1,570
C ₂ -C ₃	1,507	1,502	C ₂ -C ₃	1,375	1,403	C ₂ -C ₃	1,513	1,549
C ₃ -O	1,210	1,231	C ₃ -C ₄	1,372	1,405	C ₃ -C ₄	1,509	1,550
C ₃ -C ₄	1,465	1,465	C ₃ -O	1,375	1,355	C ₃ -O	1,430	1,437
C ₄ -C ₅	1,323	1,349	C ₄ -C ₅	1,399	1,409	C ₄ -C ₅	1,519	1,520
C ₅ -C ₆	1,496	1,517	C ₅ -C ₆	1,518	1,515	C ₅ -C ₆	1,327	1,349
C ₅ -C ₁₀	1,527	1,537	C ₆ -C ₇	1,535	1,547	C ₅ -C ₁₀	1,529	1,534
C ₆ -C ₇	1,522	1,546	C ₇ -C ₈	1,516	1,550	C ₆ -C ₇	1,500	1,506
C ₇ -C ₈	1,534	1,553	C ₈ -C ₉	1,535	1,567	C ₇ -C ₈	1,530	1,551
C ₈ -C ₉	1,550	1,572	C ₈ -C ₁₄	1,517	1,553	C ₈ -C ₉	1,542	1,565
C ₈ -C ₁₄	1,519	1,557	C ₉ -C ₁₀	1,538	1,528	C ₈ -C ₁₄	1,524	1,554
C ₉ -C ₁₀	1,552	1,582	C ₉ -C ₁₁	1,529	1,565	C ₉ -C ₁₀	1,556	1,581
C ₉ -C ₁₁	1,530	1,528	C ₁₁ -C ₁₂	1,546	1,560	C ₉ -C ₁₁	1,544	1,568
C ₁₀ -C ₁₉	1,548	1,564	C ₁₂ -C ₁₃	1,535	1,559	C ₁₀ -C ₁₉	1,541	1,565
C ₁₁ -C ₁₂	1,512	1,516	C ₁₃ -C ₁₄	1,544	1,561	C ₁₁ -C ₁₂	1,534	1,557
C ₁₁ -O	1,218	1,231	C ₁₃ -C ₁₇	1,551	1,574	C ₁₂ -C ₁₃	1,526	1,558
C ₁₂ -C ₁₃	1,535	1,556	C ₁₃ -C ₁₈	1,528	1,571	C ₁₃ -C ₁₄	1,539	1,558
C ₁₃ -C ₁₄	1,536	1,564	C ₁₄ -C ₁₅	1,543	1,547	C ₁₃ -C ₁₇	1,560	1,571
C ₁₃ -C ₁₇	1,554	1,566	C ₁₅ -C ₁₆	1,561	1,556	C ₁₃ -C ₁₈	1,536	1,569
C ₁₃ -C ₁₈	1,544	1,569	C ₁₆ -C ₁₇	1,556	1,569	C ₁₄ -C ₁₅	1,532	1,546
C ₁₄ -C ₁₅	1,524	1,549	C ₁₇ -C ₂₀	1,514	1,549	C ₁₅ -C ₁₆	1,559	1,562

Tabela V. Continuação

IV	Raio-X	Teórico	V	Raio-X	Teórico	VI	Raio-X	Teórico
C ₁₅ -C ₁₆	1,532	1,556	C ₂₀ -C ₂₁	1,474	1,516	C ₁₆ -C ₁₇	1,538	1,537
C ₁₆ -C ₁₇	1,531	1,562	C ₂₀ -O	1,446	1,520	C ₁₆ -Me	1,520	1,553
C ₁₇ -C ₂₀	1,507	1,516	C ₂₁ -O	1,458	1,522	C ₁₇ -C ₂₀	1,515	1,520
C ₂₀ -C ₂₁	1,498	1,509				C ₂₀ -C ₂₁	1,498	1,509
C ₂₀ -O	1,208	1,232				C ₂₀ -O	1,207	1,233

VII	Raio-X	Teórico	VIII	Raio-X	Teórico	IX	Raio-X	Teórico
C ₁ -C ₂	1,528	1,547	O ₁ -C ₃	1,438	1,438	O ₁ -C ₃	1,432	1,438
C ₁ -C ₁₀	1,559	1,570	O ₂ -C ₁₄	1,440	1,446	O ₂ -C ₁₄	1,445	1,447
C ₂ -C ₃	1,523	1,502	O ₃ -C ₂₁	1,371	1,361	O ₃ -C ₂₁	1,370	1,361
C ₃ -C ₄	1,432	1,465	O ₃ -C ₂₄	1,362	1,356	O ₃ -C ₂₄	1,375	1,357
C ₃ -O	1,285	1,231	O ₄ -C ₂₄	1,223	1,230	O ₄ -C ₂₄	1,211	1,230
C ₄ -C ₅	1,332	1,349	O ₅ -C ₁₁	1,439	1,437	O ₅ -C ₁₁	1,424	1,437
C ₅ -C ₆	1,527	1,516	C ₁ -C ₂	1,510	1,550	O ₆ -C ₁₂	1,204	1,234
C ₅ -C ₁₀	1,542	1,535	C ₁ -C ₁₀	1,553	1,573	C ₁ -C ₂	1,512	1,549
C ₆ -C ₇	1,523	1,545	C ₂ -C ₃	1,497	1,549	C ₁ -C ₁₀	1,541	1,573
C ₇ -C ₈	1,535	1,554	C ₃ -C ₄	1,505	1,552	C ₂ -C ₃	1,511	1,549
C ₈ -C ₉	1,541	1,572	C ₄ -C ₅	1,526	1,564	C ₃ -C ₄	1,523	1,522
C ₈ -C ₁₄	1,523	1,555	C ₅ -C ₆	1,513	1,553	C ₄ -C ₅	1,529	1,563
C ₉ -C ₁₀	1,544	1,583	C ₅ -C ₁₀	1,561	1,577	C ₅ -C ₆	1,518	1,553
C ₉ -C ₁₁	1,554	1,568	C ₆ -C ₇	1,530	1,546	C ₅ -C ₁₀	1,551	1,576
C ₁₀ -C ₁₉	1,553	1,563	C ₇ -C ₈	1,509	1,561	C ₆ -C ₇	1,522	1,546
C ₁₁ -C ₁₂	1,538	1,556	C ₈ -C ₉	1,563	1,576	C ₇ -C ₈	1,528	1,561
C ₁₂ -C ₁₃	1,521	1,560	C ₈ -C ₁₄	1,544	1,575	C ₈ -C ₉	1,565	1,576
C ₁₃ -C ₁₄	1,534	1,560	C ₉ -C ₁₀	1,564	1,593	C ₈ -C ₁₄	1,534	1,575
C ₁₃ -C ₁₇	1,561	1,578	C ₉ -C ₁₁	1,532	1,565	C ₉ -C ₁₀	1,568	1,593
C ₁₃ -C ₁₈	1,571	1,570	C ₁₀ -C ₁₉	1,523	1,569	C ₉ -C ₁₁	1,541	1,564
C ₁₄ -C ₁₅	1,568	1,547	C ₁₁ -C ₁₂	1,539	1,552	C ₁₀ -C ₁₉	1,529	1,569
C ₁₅ -C ₁₆	1,530	1,556	C ₁₂ -C ₁₃	1,538	1,564	C ₁₁ -C ₁₂	1,528	1,515
C ₁₆ -C ₁₇	1,582	1,571	C ₁₃ -C ₁₄	1,546	1,569	C ₁₂ -C ₁₃	1,541	1,525
C ₁₇ -C ₂₀	1,538	1,523	C ₁₃ -C ₁₇	1,562	1,577	C ₁₃ -C ₁₄	1,533	1,571
C ₁₇ -O	1,444	1,410	C ₁₃ -C ₁₈	1,510	1,563	C ₁₃ -C ₁₇	1,563	1,582
C ₂₀ -C ₂₁	1,492	1,510	C ₁₄ -C ₁₅	1,541	1,559	C ₁₃ -C ₁₈	1,537	1,561
C ₂₀ -O	1,261	1,233	C ₁₅ -C ₁₆	1,539	1,548	C ₁₄ -C ₁₅	1,535	1,558
O ₁₇ -C ₂₂	1,379	1,362	C ₁₆ -C ₁₇	1,544	1,558	C ₁₅ -C ₁₆	1,527	1,545
C ₂₂ -O ₂₂	1,212	1,232	C ₁₇ -C ₂₀	1,529	1,525	C ₁₆ -C ₁₇	1,546	1,557
C ₂₂ -C ₂₃	1,515	1,511	C ₂₀ -C ₂₁	1,332	1,351	C ₁₇ -C ₂₀	1,519	1,524
C ₂₃ -C ₂₄	1,527	1,551	C ₂₀ -C ₂₂	1,435	1,479	C ₂₀ -C ₂₁	1,339	1,351
C ₂₄ -C ₂₅	1,461	1,551	C ₂₂ -C ₂₃	1,343	1,343	C ₂₀ -C ₂₂	1,425	1,479
C ₂₅ -C ₂₆	1,475	1,551	C ₂₃ -C ₂₄	1,409	1,464	C ₂₂ -C ₂₃	1,337	1,343
C ₂₆ -C ₂₇	1,475	1,547				C ₂₃ -C ₂₄	1,408	1,464

Tabela V. Continuação

X	Raio-X	Teórico	XI	Raio-X	Teórico	XII	Raio-X	Teórico
C ₁ -C ₂	1,540	1,555	C ₁ -C ₂	1,480	1,507	O ₁ -C ₁	1,215	1,232
C ₁ -C ₁₀	1,553	1,573	C ₁ -C ₁₀	1,529	1,568	O ₂ -C ₅	1,449	1,529
C ₂ -C ₃	1,526	1,550	C ₂ -C ₃	1,330	1,343	O ₂ -C ₆	1,440	1,519
C ₃ -C ₄	1,530	1,560	C ₃ -C ₄	1,470	1,508	O ₃ -C ₁₂	1,426	1,442
C ₃ -O ₃	1,416	1,438	C ₄ -C ₅	1,518	1,565	O ₄ -C ₁₇	1,426	1,445
C ₄ -C ₅	1,539	1,573	C ₅ -O	1,418	1,446	O ₅ -C ₁₂	1,420	1,447
C ₄ -O ₄	1,426	1,437	C ₅ -C ₆	1,536	1,520	O ₅ -C ₂₂	1,430	1,447
C ₅ -C ₆	1,518	1,566	C ₅ -C ₁₀	1,564	1,581	O ₆ -C ₂₃	1,445	1,408
C ₅ -C ₁₀	1,565	1,586	C ₆ -O	1,214	1,232	O ₆ -C ₂₆	1,360	1,359
C ₅ -O ₅	1,461	1,453	C ₆ -C ₇	1,496	1,510	O ₇ -C ₂₆	1,207	1,229
C ₆ -C ₇	1,528	1,564	C ₇ -C ₈	1,548	1,556	C ₁ -C ₂	1,477	1,470
C ₆ -O ₆	1,436	1,440	C ₈ -C ₁₄	1,512	1,554	C ₁ -C ₁₀	1,530	1,532
C ₇ -C ₈	1,539	1,575	C ₈ -C ₉	1,549	1,577	C ₂ -C ₃	1,307	1,342
C ₇ -O ₇	1,444	1,441	C ₉ -C ₁₁	1,526	1,569	C ₃ -C ₄	1,498	1,503
C ₈ -C ₉	1,550	1,588	C ₉ -C ₁₀	1,562	1,588	C ₄ -C ₅	1,509	1,543
C ₈ -C ₁₄	1,558	1,585	C ₁₀ -C ₁₉	1,510	1,570	C ₅ -C ₆	1,457	1,518
C ₈ -O ₈	1,419	1,442	C ₁₁ -C ₁₂	1,544	1,558	C ₅ -C ₁₀	1,534	1,563
C ₉ -C ₁₀	1,557	1,589	C ₁₂ -C ₁₃	1,516	1,556	C ₆ -C ₇	1,492	1,531
C ₉ -C ₁₁	1,528	1,568	C ₁₃ -C ₁₈	1,524	1,569	C ₇ -C ₈	1,533	1,553
C ₁₀ -C ₁₉	1,543	1,571	C ₁₃ -C ₁₄	1,543	1,562	C ₈ -C ₉	1,538	1,572
C ₁₁ -C ₁₂	1,537	1,559	C ₁₃ -C ₁₇	1,554	1,567	C ₈ -C ₁₄	1,516	1,554
C ₁₂ -C ₁₃	1,537	1,562	C ₁₄ -C ₁₅	1,542	1,549	C ₉ -C ₁₀	1,579	1,586
C ₁₃ -C ₁₄	1,578	1,587	C ₁₅ -C ₁₆	1,519	1,553	C ₉ -C ₁₁	1,539	1,574
C ₁₃ -C ₁₇	1,553	1,580	C ₁₆ -O	1,424	1,438	C ₁₀ -C ₁₉	1,541	1,564
C ₁₃ -C ₁₈	1,535	1,571	C ₁₆ -C ₁₇	1,558	1,557	C ₁₁ -C ₁₂	1,518	1,568
C ₁₄ -C ₁₅	1,551	1,558	C ₁₇ -C ₂₀	1,548	1,553	C ₁₂ -C ₁₃	1,536	1,562
C ₁₄ -O ₁₄	1,453	1,457	C ₂₀ -C ₂₁	1,515	1,551	C ₁₃ -C ₁₄	1,550	1,563
C ₁₅ -C ₁₆	1,528	1,557	C ₂₀ -C ₂₂	1,532	1,563	C ₁₃ -C ₁₇	1,547	1,579
C ₁₅ -O ₁₅	1,450	1,436	C ₂₂ -O ₃	1,417	1,447	C ₁₃ -C ₁₈	1,541	1,574
C ₁₆ -C ₁₇	1,561	1,567	C ₂₂ -O ₄	1,431	1,448	C ₁₄ -C ₁₅	1,537	1,542
C ₁₇ -C ₂₀	1,551	1,573	C ₂₂ -C ₂₃	1,506	1,556	C ₁₅ -C ₁₆	1,538	1,553
C ₂₀ -C ₂₁	1,545	1,557	C ₂₃ -C ₂₄	1,534	1,550	C ₁₆ -C ₁₇	1,545	1,576
C ₂₀ -C ₂₂	1,541	1,567	C ₂₄ -C ₂₅	1,504	1,554	C ₁₇ -C ₂₀	1,550	1,582
C ₂₂ -C ₂₃	1,541	1,556	C ₂₅ -C ₂₆	1,508	1,553	C ₂₀ -C ₂₁	1,531	1,558
C ₂₃ -C ₂₄	1,507	1,556	C ₂₅ -C ₂₇	1,520	1,553	C ₂₀ -C ₂₂	1,535	1,567
C ₂₄ -C ₂₅	1,502	1,557	C ₂₆ -O ₄	1,433	1,441	C ₂₂ -C ₂₃	1,534	1,558
C ₂₄ -O ₂₄	1,455	1,439				C ₂₃ -C ₂₄	1,492	1,518
C ₂₅ -C ₂₆	1,514	1,548				C ₂₄ -C ₂₅	1,316	1,345
						C ₂₄ -C ₂₈	1,487	1,506
						C ₂₅ -C ₂₆	1,480	1,461
						C ₂₅ -C ₂₇	1,498	1,503

4.1.3. O programa FOCOS

O simulador tridimensional de dados de RMN ^{13}C desenvolvido para o sistema foi denominado de FOCOS. O programa trabalha com um sistema de três arquivos nomeados de: arquivos do usuário, arquivos estruturados e arquivos focos. Essa divisão em três tipos de arquivos permite que novos dados sejam adicionados sem a necessidade de recompilar todos os dados já existentes na base de dados, o que agiliza o processo. Na Figura 34, é apresentada a tela inicial do programa FOCOS, que permite ao usuário selecionar com que tipo de arquivo estará trabalhando (opções 1 a 3), ou ainda a opção 4, simular espectros. As interfaces das telas, assim como no sistema PESQUISA, encontram-se escritas em Inglês para facilitar a posterior avaliação, distribuição e/ou comercialização do programa nos mais diversos grupos de fitoquímica.

Figura 34. Tela inicial do simulador FOCOS

4.1.3.1. A descrição dos átomos

A opção 2 refere-se aos arquivos do usuário, que são os arquivos gerados pelo programa HyperChem que contêm as coordenadas x, y, z de cada átomo da substância,

os quais são importados pelo simulador. Nessa opção, o usuário insere as informações sobre a estereoquímica de cada centro quiral, os deslocamentos químicos de RMN ^{13}C , o solvente no qual os dados foram obtidos, o tipo de classe, esqueleto e a referência bibliográfica da substância em questão. Para representar a estereoquímica dos grupos, foi padronizado que o número 0 representaria um grupo qualquer sem designação de estereoquímica, enquanto os números 1 e 2 representariam, respectivamente, grupos com estereoquímica alfa (α) e beta (β).

A partir da importação do arquivo das coordenadas x,y,z, o programa simulador inicia o processo de codificação de cada grupamento atômico de acordo com códigos pré-estabelecidos que são apresentados na Tabela VI.

Tabela VI. Códigos dos grupamentos químicos utilizados na codificação do simulador

Átomos	Código	Átomos	Código	Átomos	Código
-CH ₃	01	=C=	12	-F	23
-CH ₂ -	02	=O	13	-Cl	24
$\begin{array}{c} \\ -\text{CH}- \end{array}$	03	-OH	14	-Br	25
$\begin{array}{c} \\ -\text{C}- \\ \end{array}$	04	-O-	15	-I	26
=CH ₂	05	-NH ₂	16	-SH	27
=CH-	06	-NH-	17	-S-	28
$\begin{array}{c} -\text{C}= \\ \end{array}$	07	$\begin{array}{c} \\ -\text{N}- \end{array}$	18	=S	29
TCH-	08	=NH	19	$\begin{array}{c} \\ -\text{S}= \\ \end{array}$	30
TC-	09	=N-	20	$\begin{array}{c} \\ =\text{S}= \\ \end{array}$	31
HC*	10	TN	21	P	32
C*	11	N*	22		

T = representação de ligação tripla; // * = representação de carbono aromático

Através da opção 2 da Figura 34, são utilizados os arquivos estruturados, que foram criados a partir dos arquivos do usuário. Na Figura 35, é apresentado o arquivo estruturado gerado para o colesterol. Os arquivos estruturados apresentam, para cada

substância, os tipos e subtipos de átomos (coluna Type), que são codificados de acordo com a Tabela VI, e suas respectivas estereoquímicas (coluna Stere). Para o átomo 1 do colesterol (Figura 35), o tipo 1(2) representa, respectivamente, um carbono do subtipo -CH₂- sem designação de estereoquímica (coluna Stere=0); já para o átomo 28, o tipo 3(14) representa, respectivamente, um oxigênio do subtipo -OH e de configuração β (coluna Stere=2). Também são apresentados o número de ligantes – No Bnd – (Figura 35, No Bnd= 4), suas posições e tipos de ligações na estrutura – Bnd 1, Bnd 2, Bnd 3, Bnd 4 – (Figura 35, respectivamente, 2, 10, 34 e 35, todos com ligações simples (1)), as coordenadas X, Y, Z do átomo em questão (Figura 35, -2,563, 0,316, -1,194, respectivamente) aliado a seu deslocamento químico – CS – (Figura 35, CS= 37,30). Juntamente com esses dados, o arquivo estruturado apresenta, em sua segunda linha, os outros dados inseridos (classe, esqueleto, número de átomos e de carbonos, solvente e referência bibliográfica).

Prompt do MS-DOS - FOCOS

Auto

LIST STRUCTURED-FILE

REGISTER: 287
CLASS: 9 SKL: 4811 No ATO: 74 No CAR: 27 SOLV: 1 PER: 62 YEAR: 77 PG: 439

ATM	TYPE	STERE	No	BND	BND 1	BND 2	BND 3	BND 4	X	Y	Z	CS
1	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
2	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
3	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
4	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
5	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
6	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
7	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
8	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
9	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
10	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
11	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
12	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
13	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
14	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
15	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
16	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
17	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
18	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
19	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
20	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
21	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
22	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
23	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
24	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
25	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
26	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
27	1(2)	0	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30
28	3(14)	2	4	1	1	1	1	1	-2.563	0.316	-1.194	37.30

Press <ESC> to Quit

PROGRAM FOCOS (C18) - 01/08/2003 - 18:30

Figura 35. Arquivo estruturado gerado através do programa FOCOS

4.1.3.2. A criação do arquivo para simulação de dados de RMN ^{13}C

A partir dos arquivos estruturados e da opção 3 da Figura 34, podem-se criar os arquivos de focos para cada átomo. Como foco definimos todo o ambiente químico presente ao redor de um determinado átomo até uma distância estabelecida pelo usuário em Ångstrom (Å). Esse conceito foi originado a partir dos ELCOs^{140,141} desenvolvidos no gerador do sistema DARC. Inicialmente, o programa FOCOS calcula as distâncias entre um determinado carbono foco e todos os outros átomos da substância. A seguir, o programa organiza os átomos ao redor do carbono foco por ordem crescente de distância até 6,0Å, com um percentual de erro de 5% na distância. Esses valores foram padronizados a partir de inúmeros testes realizados com distâncias que variavam entre 4,0 a 8,0Å, sendo observado os melhores resultados a 6,0Å, que correspondem a cerca de cinco ligações C-C. Assim, foram criados os focos a 6,0Å, para cada átomo das substâncias.

Na Figura 36, é apresentada a lista completa dos focos para um determinado átomo. Nessa, pode-se observar que o átomo é do tipo 2 (-CH₂-), o deslocamento químico máximo, mínimo e médio para aquele conjunto de focos (Dq min, Dq máx, Dq médio), o número de vezes que aquele conjunto de focos está presente na base de dados (Num dq) e o número de vizinhos até 6,0Å que aquele respectivo átomo apresenta (Num viz). Abaixo, todos os grupamentos são apresentados em ordem crescente de distância em relação ao átomo estudado, juntamente com seus respectivos tipos e as estereoquímicas que apresentam.

Na Figura 37, é apresentada uma lista compacta dos focos presentes na base de dados. Nesta listagem são apresentados os tipos de focos (Type), os deslocamentos químicos médios, mínimos e máximos (CS Mean, CS Min e CS Max, respectivamente), a diferença entre os deslocamentos químicos (Delta CS), o número de vezes em que aquele foco se repete na base de dados (No data) e o número de grupamentos vizinhos que o respectivo foco apresenta (No groups), até 6,0Å de distância. O deslocamento químico médio (CS Mean) gerado para um determinado foco é utilizado pelo programa para a simulação dos dados de uma nova substância quando esta apresentar um átomo de carbono com o respectivo foco.

A base de dados de esteróides do programa FOCOS apresenta 800 substâncias, que originaram 6466 focos diferentes, os quais são utilizados pelo programa durante o processo de simulação de dados de RMN ^{13}C de uma amostra.

Figura 36. Lista completa do primeiro foco gerado através do programa FOCOS

Figura 37. Lista compacta dos focos presentes na base de dados

4.1.3.3. A previsão de dados de RMN ^{13}C

A partir do arquivo com os focos gerados e através da opção 4 da Figura 34, pode-se simular os dados de RMN ^{13}C de uma nova amostra. Para tanto, a estrutura deve ser inicialmente desenhada e submetida ao processo de obtenção da conformação de menor energia através do programa HyperChem, utilizando o mesmo procedimento e variáveis empregados durante o arquivamento das substâncias na base de dados. Após este processo, a substância é submetida ao simulador FOCOS, que descreverá todo o ambiente químico de cada carbono existente na estrutura-teste e realizará a busca destes no arquivo de focos gerados pelo programa, apresentando, ao final da análise, os dados de RMN ^{13}C para a substância.

Para demonstrar a utilização do simulador foi escolhido o esteróide empregado para exemplificar o sistema PESQUISA (Figura 18). Na Figura 38, pode-se visualizar os dados de RMN ^{13}C previstos (CS Mean) para cada um dos carbonos da estrutura. Com esses dados, são apresentados também os tipos de carbonos presentes na estrutura (Type), a estereoquímica de cada um desses carbonos (Stere), o deslocamento químico mínimo e máximo para cada um dos carbonos da estrutura (CS Min e CS Max) e o número de vezes em que o respectivo foco se apresenta na base de dados (No data). Quando o programa não encontra os focos para um determinado carbono, um deslocamento químico igual a $\delta -1,00$ é apresentado como resposta, o que não ocorreu para esse exemplo.

Uma análise por regressão linear múltipla foi realizada para os dados previstos pelo simulador FOCOS e os dados experimentais publicados na literatura, obtendo-se um $R = 0,99998$ com um $F = 748126,6$. O gráfico dos valores experimentais vs. previstos é apresentado na Figura 39.

Os dados de RMN ^{13}C do esteróide apresentado na Fig. 18 também foram estimados através de um programa comercial, o simulador da ACD Labs¹⁴². Esse programa simula dados de RMN ^{13}C baseando-se em relações lineares de aditividade e em consultas à base de dados espectroscópicos. Os dados da literatura, bem como os resultados obtidos com o programa da ACD Labs e o simulador FOCOS são apresentados e comparados na Tabela VII.

Figura 38. Dados de RMN ^{13}C previstos para o esteroide da Figura 18

Figura 39. Valores experimentais versus valores previstos pelo simulador FOCOS

Figura 18. Esteróide utilizado para testar o sistema PESQUISA

Tabela VII. Dados de RMN ^{13}C previstos por diferentes programas simuladores

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	207,7	207,5	0,2	202,0	5,7
2	129,1	129,0	0,1	128,8	0,1
3	141,0	141,3	0,3	141,0	0,0
4	36,4	36,8	0,4	35,2	1,2
5	76,9	76,6	0,3	78,1	1,2
6	75,3	75,9	0,6	76,1	0,8
7	29,2	30,0	0,8	31,6	2,4
8	29,2	29,5	0,3	29,9	0,7
9	47,1	47,0	0,1	52,2	5,1
10	53,7	53,6	0,1	51,2	2,5
11	68,4	68,3	0,1	66,4	2,0
12	51,1	50,6	0,5	49,6	1,5
13	43,1	43,0	0,1	43,1	0,0
14	55,8	56,3	0,5	55,8	0,0
15	23,8	24,4	0,6	24,4	0,6
16	28,3	28,3	0,0	28,3	0,0
17	55,9	56,3	0,4	55,2	0,7
18	13,6	13,9	0,3	13,9	0,3
19	14,6	14,3	0,3	17,1	2,5
20	36,3	36,4	0,1	34,9	1,4
21	18,6	18,9	0,3	18,9	0,3
22	33,5	33,9	0,4	34,0	0,5
23	30,6	30,9	0,3	30,7	0,1
24	39,1	39,3	0,2	38,4	0,7
25	31,5	31,5	0,0	31,7	0,2
26	20,5	20,6	0,1	20,6	0,1
27	17,6	17,7	0,1	18,2	0,6
28	15,5	15,5	0,0	15,6	0,1
29	169,9	169,8	0,1	171,1	1,2
30	21,3	21,4	0,1	21,5	0,2
Maior erro observado			0,80		5,70
Erro médio			0,26		1,09
Desvio padrão			0,20		1,40

V - O GERADOR DE ESTRUTURAS DE ESTERÓIDES

A grande maioria dos geradores estruturais desenvolvidos por inúmeros grupos de pesquisa realiza a geração exaustiva dos isômeros a partir de uma determinada fórmula molecular^{2,35,143}. Esse procedimento garante a construção da estrutura correta de uma substância, entretanto várias outras estruturas também são apresentadas, restando aos usuários discernir entre as prováveis propostas. O número de possíveis estruturas é relativamente reduzido quando fragmentos estruturais presentes e ausentes na substância, inferidos pelo espectroscopista através da análise prévia dos dados espectrais, são apresentados, na fase inicial, ao gerador.

Para estruturas de produtos naturais esse problema é ainda maior, devido ao número elevado de átomos e à complexidade estrutural presente nas substâncias. Um dos métodos encontrados para atenuar esse problema foi a utilização de módulos que examinam se as estruturas geradas incorporam esqueletos ou fragmentos de esqueletos de produtos naturais.

Um outro método de atuação dos geradores, que tem sido bastante explorado recentemente, é a geração exaustiva dos isômeros com posterior validação das estruturas geradas através da previsão dos dados de RMN ¹³C de cada estrutura e comparação com os dados experimentais^{37,38,41,144}. Entretanto, nos dois casos, esses módulos são usados ao final de cada análise, quando todas as prováveis estruturas já foram geradas e, conseqüentemente, foi gasto um elevado tempo computacional.

Para evitar e/ou solucionar esse problema, iniciamos a construção do gerador de estruturas a partir do conhecimento dos esqueletos de produtos naturais, mais especificamente, de esteróides, os quais foram estudados nesse trabalho.

O desenvolvimento do gerador de estruturas foi projetado através da sobreposição dos tipos de anéis existentes em um esteróide. O esqueleto-base de um esteróide, apresentado na Figura 40, é composto por 4 anéis, denominados de anéis A, B, C e D e, dependendo do esqueleto, pode existir ou não uma cadeia lateral interligada ao anel D.

Figura 40. Esqueleto base de um esteróide

Assim, foram estudados através da base de dados do SISTEMAT, todos os tipos de anéis e cadeias laterais que podem ocorrer em esteróides e coletadas as faixas de deslocamento químico que caracterizam aquele determinado tipo de substituição. Esses dados foram utilizados como o ponto de partida do gerador desenvolvido.

5.1. A busca de sinais característicos de RMN ^{13}C

5.1.1. O Programa TIPCARB

O programa TIPCARB^{66,69} é um aplicativo que serve para determinar quais tipos de átomos de carbono existem em uma determinada posição do esqueleto, isto é, o programa lista os tipos de átomos de carbono para cada esqueleto, baseado na numeração padronizada.

As informações apresentadas pelo TIPCARB são de grande importância pois elas guiam o usuário na busca de regras heurísticas, uma vez que, através destas, pode-se verificar quais são as posições mais ou menos oxidadas do esqueleto. Um exemplo da utilização do programa foi realizado para os esteróides de esqueleto cardanolídeo (Figura 41), onde os resultados são apresentados na Tabela VIII.

Figura 41. Esqueleto cardanolídeo

Através dos dados apresentados na Tabela VIII, pode-se verificar, por exemplo, que das 94 ocorrências do esqueleto cardanolídeo, em 91 casos os carbonos 20 e 22 apresentam uma ligação dupla e, em todos os casos, o carbono 23 apresenta-se como um carbono quaternário de ligação dupla. Tais dados nos permitem inferir quais são os

átomos ou subestruturas constantes de um determinado esqueleto. Esses dados são importantes na busca da caracterização de tipos ou subtipos de esqueletos, a qual é realizada através do sistema PICKUP.

Tabela VIII. Listagem obtida através do programa TIPCARB para o esqueleto cardanolídeo

Átomo	CH ₃	CH ₂	CH	C	CH ₂ =	CH=	C=	CH*	C*
1	0	93	1	0	0	0	0	0	0
2	0	77	15	0	0	0	2	0	0
3	0	0	87	0	0	0	7	0	0
4	0	90	1	0	0	3	0	0	0
5	0	0	62	19	0	0	13	0	0
6	0	69	15	0	0	10	0	0	0
7	0	88	1	0	0	0	5	0	0
8	0	0	90	2	0	0	2	0	0
9	0	0	93	0	0	0	1	0	0
10	0	0	0	94	0	0	0	0	0
11	0	71	8	0	0	1	14	0	0
12	0	84	5	0	0	0	5	0	0
13	0	0	0	94	0	0	0	0	0
14	0	0	27	60	0	0	7	0	0
15	0	87	2	0	0	5	0	0	0
16	0	74	4	0	0	16	0	0	0
17	0	0	73	5	0	0	16	0	0
18	91	3	0	0	0	0	0	0	0
19	90	2	0	0	0	2	0	0	0
20	0	0	1	2	0	0	91	0	0
21	5	89	0	0	0	0	0	0	0
22	0	3	0	0	0	91	0	0	0
23	0	0	0	0	0	0	94	0	0

CH* = CH aromático // C* = C aromático

5.1.2. O Sistema PICKUP

O sistema PICKUP^{69,145-147} realiza análises de RMN ¹³C com o objetivo de auxiliar na procura de faixas de deslocamentos químicos característicos dos tipos e/ou subtipos de esqueletos. Essas faixas de deslocamento químico obtidas podem ser consideradas como regras heurísticas, pois fornecem uma resposta pronta, a partir da observação de resultados típicos, ao problema questionado. O sistema é constituído de três programas descritos a seguir :

1. *O Programa PICKCR2*: esse programa cria arquivos com dados de deslocamentos químicos de RMN ^{13}C dos átomos de carbono dos compostos. Esses arquivos podem apresentar os dados por classe ou por tipos de esqueletos, acarretando na diminuição do tempo gasto durante a análise.

2. *O Programa SISPICK2* : esse programa, a partir das subestruturas e posições biogénéticas impostas pelo usuário, procura no banco de dados do esqueleto desejado o conjunto de substâncias que apresentam estas características e fornece os deslocamentos químicos máximos e mínimos de RMN ^{13}C , juntamente com as multiplicidades dos átomos de carbono requisitados.

As subestruturas são escolhidas com base na listagem apresentada pelo programa TIPCARB ou através do próprio conhecimento do químico acerca de uma classe de substâncias ou de um esqueleto específico. Para os esteróides, as subestruturas foram escolhidas com base na listagem dos tipos de anéis existentes nas estruturas, visando a construção do gerador de estruturas, as quais são codificadas a partir do tipo de subestruturas, apresentadas na Tabela IX, onde cada letra representa qualquer grupamento químico da Tabela VI.

3. *O Programa PICKRVSR* : esse programa, a partir dos deslocamentos químicos de RMN ^{13}C máximos e mínimos e multiplicidades obtidos anteriormente, testa sobre o banco com os dados de RMN ^{13}C de todos os compostos, se estas faixas obtidas são ou não, e em que grau, características de um determinado esqueleto, obtendo-se, desta forma, a porcentagem de reconhecimento desses deslocamentos para um determinado esqueleto.

Para exemplificar a utilização do sistema PICKUP na busca de faixas de deslocamentos químicos características, foi escolhida a subestrutura que compreende os carbonos 20, 21, 22 e 23, com ligação dupla entre 20(22) e uma lactona formada entre os carbonos 20:23, do esqueleto cardanolídeo. Esta subestrutura foi codificada segundo as subestruturas e grupamentos químicos apresentados, respectivamente, nas Tabelas IX e VI. A codificação da subestrutura testada é apresentada na Tabela X.

Tabela IX. Códigos de subestruturas utilizados pelo programa SISPICK2

Subestrutura	Código	Subestrutura	Código
A	01	A—B	02
A—B—C	03	A—B—C—D	05
04		06
07		08
A—B—C—D—E	09		10
11		12
13		14
15		16
17		

A subestrutura codificada foi então submetida ao programa SISPICK2 onde foram obtidas as faixas de deslocamento químico dos átomos de carbono da subestrutura no esqueleto cardanolídeo que, então, foram confrontadas, através do programa PICKRVSR, com todos os esteróides que apresentam dados de RMN ^{13}C (1405

substâncias), obtendo-se assim a porcentagem de reconhecimento destas faixas no banco de dados. Esses resultados são apresentados na Tabela XI.

Tabela X. Codificação da subestrutura do esqueleto cardanolídeo

Subestrutura	Codificação
Código da subestrutura : 12
	Código dos grupamentos : 15 02 07 06 07
	Código das ligações : 01 01 02 01

Ligações: simples sem configuração ou aromática: 01; dupla: 02; tripla: 03;
simples de configuração alfa: 04; simples de configuração beta: 05.

Tabela XI. Resultados obtidos para a subestrutura selecionada do esqueleto cardanolídeo

Átomo de Carbono	Faixas de deslocamento químico	Percentual de reconhecimento
20	178.6 – 158.6 s	100.0 %
21	76.8 - 71.5 t	
22	121.3 – 111.3 s	
23	177.6 – 170.6 s	

O procedimento descrito para o esqueleto cardanolídeo foi realizado para vários esqueletos da base de dados de esteróides e para os diferentes tipos de substituição que ocorrem nos anéis A, B, C e D e na cadeia lateral de esteróides. As faixas de deslocamento químico obtidas serão apresentadas no item resultados.

5.2. O Programa ESTGERA

O programa ESTGERA realiza a sobreposição de anéis de esteróides. Na Figura 42 é apresentada a tela inicial do programa ESTGERA. O programa encontra-se escrito em inglês para facilitar uma posterior comercialização. Através da opção 1, as faixas de deslocamento químico características dos mais variados tipos de substituição nos diferentes anéis e cadeia lateral de esteróides, obtidas através do sistema PICKUP, foram armazenadas no programa. A opção 2 permite a pesquisa de grupos substituintes presentes em um espectro de RMN ^{13}C e, através das opções 3 e 4, o usuário pode pesquisar os tipos de anéis existentes em uma substância e gerar suas estruturas, respectivamente.

Figura 42. Tela inicial do programa ESTGERA

Os dados de RMN ^{13}C do esteróide apresentado na Figura 18 (pág. 53) foram submetidos ao programa ESTGERA para a geração da provável estrutura.

Inicialmente, verificou-se a possível existência de grupos substituintes na estrutura da substância. O programa forneceu dois macronós como prováveis (Figura 43), sendo considerado correto o acetato, pois é o substituinte que apresenta a menor faixa de erro médio.

Figura 43. Prováveis grupos substituintes propostos pelo programa ESTGERA

A seguir, realizou-se a pesquisa nas faixas de deslocamento químico características, através da opção 3 da tela inicial do programa (Figura 42), e o programa selecionou os possíveis tipos de fragmentos (anéis e cadeias laterais) existentes na estrutura, os quais são apresentados na Figura 44. A partir da opção 4 (Figura 42) foram geradas todas as prováveis estruturas para a substância (Figuras 45-47).

Uma análise rápida das três propostas estruturais geradas pelo programa ESTGERA, permite descartar a primeira estrutura imediatamente, uma vez que os dados de RMN ^{13}C da substância apresentam dois dubletos em 75,3 e 68,4δ, que indicam a presença de hidroxilas ou grupamentos -OR, que não são encontrados na primeira estrutura fornecida. A diferença existente entre as duas últimas estruturas, está na estereoquímica da hidroxila ligada em C-11, restando então, ao usuário, decidir qual das estruturas, a 11 α -OH ou 11 β -OH, corresponde à correta estrutura da substância.

Figura 44. Fragmentos propostos pelo programa ESTGERA

Figura 45. Primeira proposta estrutural gerada pelo programa ESTGERA

Figura 46. Segunda proposta estrutural gerada pelo programa ESTGERA

Figura 47. Terceira proposta estrutural gerada pelo programa ESTGERA

5.3. Validação de propostas estruturais

A partir das propostas estruturais apresentadas, utilizou-se a previsão de deslocamentos químicos para validar ou rejeitar as estruturas candidatas. Para tanto, as três estruturas foram submetidas ao programa FOCOS, que realizou a simulação dos dados de RMN ^{13}C para cada uma das substâncias. Os dados simulados foram confrontados com os dados experimentais de modo que as estruturas candidatas fossem classificadas. Frequentemente, realiza-se uma regressão linear múltipla entre os deslocamentos químicos previstos e os experimentais para classificar as substâncias, sendo escolhida a substância que apresenta a maior correlação entre os dados. Através dessa análise, foram obtidas correlações de 0,8322, 0,9999 e 0,9996 para as estruturas I, II e III, respectivamente, assegurando a proposta II como a correta estrutura da substância.

Entretanto, essa metodologia não considera os dados não previstos pelo programa. Para tanto, nesse trabalho, foi desenvolvido um novo índice para a classificação das propostas estruturais, denominado de Grau de Validade (GV). Esse índice é calculado a partir do número de acertos e erros entre os deslocamentos químicos estimados e os experimentais, considerando-se os intervalos de confiança (95%) e os átomos para os quais o deslocamento químico não foi previsto. Assim, para uma estrutura contendo N_C átomos de carbono temos: $NC = NV + NF + NP$, onde NV é o número de sinais que se assemelham com os valores previstos, NF expressa o número de erros e NP o número de sinais que não foram previstos. Desta forma, para cada estrutura o índice GV foi calculado a partir da equação: $GV = (NV - NF) / NC$. Portanto, quando o $GV = 1$ significa uma validade de 100,0% da estrutura, isto é, que cada deslocamento químico previsto a partir da estrutura candidata confronta com um sinal do espectro da substância desconhecida. Na Tabela XII são apresentados os resultados da avaliação do sistema de validação de estruturas.

Através da análise dos dados da Tabela XII, pode-se verificar que, dentre as três propostas estruturais apresentadas pelo programa ESTGERA, somente a estrutura II, que é a estrutura correta da substância, seria classificada por apresentar o maior grau de validação, demonstrando assim a importância da integração entre o gerador ESTGERA e o simulador FOCOS. Recentemente, essa metodologia foi testada¹⁴⁸ para validar as propostas estruturais de esteróides apresentadas pelo gerador de estruturas MOLGEN¹⁴³.

Nesse caso, entre as 18 propostas fornecidas pelo programa, apenas quatro foram selecionadas por apresentarem o maior grau de validade.

Tabela XII. Avaliação do sistema de validação de estruturas

Dados experimentais		Proposta estrutural I		Proposta estrutural II		Proposta estrutural III	
Carbono	Literatura	Dados previstos		Dados previstos		Dados previstos	
1	207,7	—	NP	207,5	V	208,3	V
2	129,1	—	NP	129,0	V	129,9	V
3	141,0	—	NP	141,3	V	142,0	V
4	36,4	—	NP	36,8	V	36,9	V
5	76,9	—	NP	76,6	V	76,6	V
6	75,3	—	NP	75,9	V	75,9	V
7	29,2	29,8	V	30,0	V	30,0	V
8	29,2	29,8	V	29,5	V	30,3	V
9	47,1	47,5	V	47,0	V	48,3	V
10	53,7	50,9	F	53,6	V	52,9	V
11	68,4	31,5	F	68,3	V	72,9	F
12	51,1	33,4	F	50,6	V	48,4	F
13	43,1	42,0	V	43,0	V	42,1	V
14	55,8	54,9	V	56,3	V	54,9	V
15	23,8	27,6	F	24,4	V	24,0	V
16	28,3	28,3	V	28,3	V	28,3	V
17	55,9	56,3	V	56,3	V	56,3	V
18	13,6	14,0	V	13,9	V	13,6	V
19	14,6	15,0	V	14,3	V	11,2	F
20	36,3	36,4	V	36,4	V	36,4	V
21	18,6	18,9	V	18,9	V	18,9	V
22	33,5	33,9	V	33,9	V	33,9	V
23	30,6	30,9	V	30,9	V	30,9	V
24	39,1	39,3	V	39,3	V	39,3	V
25	31,5	31,5	V	31,5	V	31,5	V
26	20,5	20,6	V	20,6	V	20,6	V
27	17,6	17,7	V	17,7	V	17,7	V
28	15,5	14,4	F	15,5	V	15,5	V
29	169,9	169,8	V	169,8	V	169,8	V
30	21,3	21,4	V	21,4	V	21,4	V
		NV = 19		NV = 30		NV = 27	
		NF = 5		NF = 0		NF = 3	
		GV = 0.47		GV = 1.00		GV = 0.80	

Na Figura 48, é apresentado um fluxograma de análise de dados para uma substância desconhecida através dos programas do sistema especialista SISTEMAT.

Figura 48. Fluxograma de análise de dados dos programas do SISTEMAT

VI - RESULTADOS

Para avaliar a eficiência dos programas aplicativos do SISTEMAT no processo de determinação estrutural e identificação de esqueletos de substâncias, bem como do simulador de dados de RMN ^{13}C , foram utilizados os dados referentes a cinquenta esteróides, publicados na literatura a partir do ano de 2000, sendo consideradas apenas as substâncias ditas inéditas ou que ainda não estão presentes na base de dados.

Os resultados obtidos para cada uma das substâncias são apresentados a seguir. Nas tabelas de probabilidade da substância pertencer a um esqueleto proposto pelos programas aplicativos, somente são apresentados os cinco primeiros esqueletos mais prováveis, portanto, quando os programas propuseram mais que cinco esqueletos, a somatória das probabilidades constantes nas tabelas não serão iguais a 100%. O resultado global apresentado é uma média ponderada das probabilidades fornecidas pelos programas, tendo sido atribuídos os pesos 0,9, 0,7, 0,5 e 1,0 aos programas SISCONST, C13MACH, SISOCBOT e REGRAS, respectivamente.

As subestruturas propostas pelo programa SISCONST são apresentadas e estão representadas como aparecem na tela do programa, subentendendo-se a presença dos átomos de hidrogênio. Aos átomos de carbono de cada subestrutura estão associados os deslocamentos químicos do espectro em análise, cuja atribuição é feita pela comparação com os deslocamentos químicos das subestruturas compatíveis com o espectro questionado, obtidas a partir das substâncias presentes no banco de dados. Nos testes realizados com o programa SISCONST foram solicitados, em geral, a metade do número de átomos presentes na estrutura da substância, como número mínimo de átomos, e uma faixa de erro de 1,0 δ . Nos casos em que o programa não encontra subestruturas com deslocamentos químicos nesta faixa, esta é automaticamente aumentada para 2,0 e 3,0 δ . Em alguns testes, são apresentadas subestruturas com um número de átomos inferior ao estipulado, as quais foram fornecidas pelo programa SISCONST, após a retirada dos sinais atribuídos aos átomos da primeira subestrutura proposta.

No programa C13MACH, a probabilidade de esqueleto foi obtida com base nas dez primeiras substâncias com maior índice de semelhança, dentro de uma faixa de erro de 1,0 δ , listadas pelo programa. Também são apresentados os prováveis macronós, propostos pelo programa MACRONO, para as substâncias que possuem grupos

substituintes, sendo a análise dos dados espectrais pelos demais programas aplicativos realizada somente após a retirada dos sinais de RMN ^{13}C atribuídos aos substituintes.

Através do programa SISOCBOT são indicados os percentuais de ocorrência dos esqueletos na família e gênero da planta dos quais a substância foi isolada. Nos casos em que o gênero ainda não havia sido estudado, a pesquisa não foi realizada.

O programa REGRAS indica o esqueleto e o respectivo percentual de reconhecimento baseado na desfuncionalização do espectro de RMN ^{13}C (Tabela IV, pg.64) e comparação com os tipos de átomos de carbono apresentados para cada esqueleto na Tabela III (pg. 61). Quando existe mais de um provável esqueleto com o mesmo número de carbonos quaternários, metínicos, metilênicos e metílicos, o programa confronta o espectro de RMN ^{13}C em análise com as faixas de deslocamentos químicos características, obtidas a partir do sistema PICKUP, a fim de propor o esqueleto da substância. O mesmo procedimento é realizado pelo programa para listar os prováveis tipos de anéis presentes na substância questionada.

Na obtenção, através do sistema PICKUP, das faixas de deslocamento químico características de esteróides, a metodologia empregada foi o estudo dos tipos de anéis presentes nas substâncias, como apresentado na seção V. Nas Tabelas XIII à XVII são apresentadas, respectivamente, as faixas de deslocamento químico características de vários tipos de substituição nos anéis A, B, C, D e cadeia lateral de esteróides. Essas faixas de deslocamento químico foram avaliadas através do gerador ESTGERA para a geração das estruturas das substâncias empregadas nos testes dos programas aplicativos. Os tipos de anéis selecionados pelo programa ESTGERA são apresentados na Tabela XVIII.

Os esteróides usados para avaliar os programas aplicativos do SISTEMAT foram também utilizados para testar a eficiência do simulador de dados de RMN ^{13}C . Essas substâncias (1-50) foram desenhadas no programa HyperChem, segundo a metodologia descrita para inserir substâncias na base de dados e, em seguida, submetidas ao programa FOCOS para simulação de seus dados de RMN ^{13}C e posterior comparação com os dados experimentais da literatura. Também avaliamos o desempenho do simulador tridimensional frente a um simulador comercial, o ACD NMR Predictor. Os dados simulados através dos dois programas são apresentados após os resultados obtidos com os programas aplicativos do SISTEMAT.

Tabela XIII. Faixas de deslocamento químico características de esteróides – Anel A

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
<div>Anel A</div> <div> </div>				
[5 α -H]	1	40,7 – 37,7	t	94,6
	2	22,6 – 21,5	t	
	3	27,1 – 25,2	t	
	4	29,7 – 20,3	t	
	5	58,7 – 39,2	d	
	10	41,7 – 36,0	s	
	19	15,8 – 11,1	q	
[5 β -H]	1	40,9 – 36,7	t	100,0
	2	22,3 – 21,3	t	
	3	27,5 – 27,0	t	
	4	27,7 – 27,1	t	
	5	45,5 – 43,7	d	
	10	36,9 – 35,0	s	
	19	27,7 – 24,1	q	
[1 α -OH]	1	71,7 – 71,4	d	100,0
	2	29,0 – 28,8	t	
	3	20,4 – 20,1	t	
	4	28,7 – 28,4	t	
	5	39,1 – 38,7	d	
	10	40,3 – 40,0	s	
	19	12,9 – 12,6	q	
[1 β -OH]	1	79,0 – 78,8	d	100,0
	2	33,4 – 33,2	t	
	3	24,8 – 24,6	t	
	4	28,8 – 28,5	t	
	5	46,4 – 46,0	d	
	10	42,6 – 42,3	s	
	19	6,5 – 6,8	q	
[2 β -OH]	1	45,2 – 45,0	t	100,0
	2	68,1 – 67,8	d	
	3	34,0 – 33,7	t	
	4	23,9 – 23,6	t	
	5	47,5 – 47,3	d	
	10	36,1 – 35,8	s	
	19	14,8 – 14,6	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[3 α -OH]	1	33,0 – 32,0	t	100,0
	2	29,7 – 28,2	t	
	3	66,5 – 65,5	d	
	4	36,9 – 28,2	t	
	5	48,0 – 39,0	d	
	10	39,7 – 35,0	s	
	19	14,8 – 11,1	q	
[3 α -OH; 5 β -H]	1	37,2 – 34,5	t	100,0
	2	31,2 – 29,7	t	
	3	72,3 – 71,0	d	
	4	39,9 – 30,5	t	
	5	48,7 – 41,5	d	
	10	36,0 – 33,5	s	
	19	25,2 – 22,7	q	
[3 β -OH]	1	41,4 – 28,7	t	93,0
	2	32,5 – 27,2	t	
	3	76,4 – 69,5	d	
	4	39,2 – 30,0	t	
	5	56,9 – 36,7	d	
	10	47,0 – 34,0	s	
	19	19,2 – 11,5	q	
[3 β -OH; 5 β -H]	1	33,2 – 29,2	t	97,0
	2	29,2 – 25,0	t	
	3	72,9 – 65,5	d	
	4	37,5 – 30,6	t	
	5	37,9 – 35,9	d	
	10	40,2 – 34,7	s	
	19	28,6 – 23,0	q	
[3 α -OR]	1	32,9 – 32,7	t	100,0
	2	26,4 – 26,0	t	
	3	70,1 – 69,9	d	
	4	32,9 – 32,5	t	
	5	40,2 – 40,0	d	
	10	36,0 – 35,7	s	
	19	11,4 – 11,0	q	
[3 α -OR; 5 β -H]	1	35,4 – 35,0	t	100,0
	2	31,6 – 26,3	t	
	3	74,5 – 74,1	d	
	4	35,2 – 32,2	t	
	5	43,0 – 41,2	d	
	10	35,7 – 34,5	s	
	19	23,5 – 23,0	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[3 β -OR]	1	39,4	– 28,5	79,8
	2	29,6	– 25,1	
	3	79,6	– 70,4	
	4	36,0	– 26,2	
	5	56,5	– 35,5	
	10	40,9	– 33,9	
	19	23,8	– 11,1	
[3 β -OR; 5 β -H]	1	36,2	– 30,2	83,3
	2	25,5	– 24,1	
	3	72,0	– 69,0	
	4	32,0	– 30,2	
	5	37,7	– 32,9	
	10	41,0	– 34,4	
	19	30,8	– 23,0	
[4 α -OH]	1	38,1	– 37,7	98,5
	2	20,5	– 20,3	
	3	36,5	– 36,1	
	4	70,6	– 70,2	
	5	54,3	– 53,9	
	10	37,7	– 37,5	
	19	13,8	– 13,4	
[4 β -OH]	1	38,7	– 38,5	93,7
	2	17,3	– 16,9	
	3	34,0	– 33,6	
	4	72,6	– 72,2	
	5	50,3	– 50,0	
	10	36,3	– 36,0	
	19	14,8	– 14,5	
[1 β ,3 α -OH]	1	74,2	– 73,8	100,0
	2	40,4	– 39,7	
	3	66,9	– 66,1	
	4	36,1	– 35,6	
	5	38,2	– 37,4	
	10	42,8	– 42,2	
	19	6,0	– 5,3	
[1 β ,3 β -OH]	1	76,1	– 75,0	100,0
	2	42,4	– 33,2	
	3	69,6	– 67,5	
	4	37,5	– 34,7	
	5	39,4	– 32,2	
	10	41,4	– 41,2	
	19	19,7	– 6,5	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[2β,3β-OH]	1	45,9 – 43,4	t	100,0
	2	70,1 – 69,6	d	
	3	72,3 – 71,0	d	
	4	28,7 – 25,0	t	
	5	54,5 – 40,2	d	
	10	44,5 – 42,7	s	
	19	20,1 – 15,6	q	
[2β,3β-OH; 5β-H]	1	39,0 – 36,2	t	100,0
	2	68,0 – 66,9	d	
	3	70,0 – 66,9	d	
	4	35,0 – 30,7	t	
	5	50,2 – 40,2	d	
	10	43,5 – 39,4	s	
	19	24,3 – 21,6	q	
[2β,3β-OR]	1	41,0 – 40,4	t	100,0
	2	69,9 – 69,0	d	
	3	72,5 – 72,1	d	
	4	29,3 – 21,0	t	
	5	53,5 – 45,4	d	
	10	35,7 – 22,8	s	
	19	15,3 – 14,3	q	
[2β,3β-OR; 5β-H]	1	38,2 – 33,5	t	85,7
	2	69,3 – 67,0	d	
	3	69,6 – 66,5	d	
	4	34,0 – 28,7	t	
	5	51,4 – 37,2	d	
	10	38,5 – 37,5	s	
	19	23,8 – 23,6	q	
[2α,3β-OH; 5β-H]	1	37,2 – 37,0	t	100,0
	2	70,5 – 70,2	d	
	3	71,7 – 71,4	d	
	4	30,4 – 30,0	t	
	5	37,6 – 37,2	d	
	10	35,8 – 35,4	s	
	19	24,5 – 24,1	q	
[2α,3β-OR]	1	43,4 – 42,2	t	100,0
	2	71,8 – 71,3	d	
	3	74,5 – 74,0	d	
	4	32,5 – 27,5	t	
	5	45,5 – 43,5	d	
	10	37,4 – 37,0	s	
	19	15,8 – 12,5	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[2 α .3 α .4 β -OH]	1	42,0 – 41,5	t	100,0
	2	66,5 – 66,2	d	
	3	75,0 – 74,8	d	
	4	77,6 – 77,0	d	
	5	44,4 – 43,9	d	
	10	37,9 – 37,3	s	
	19	16,5 – 16,0	q	
[3 β .4 β -OH]	1	41,0 – 38,9	t	50,0
	2	26,8 – 25,5	t	
	3	73,5 – 72,1	d	
	4	77,6 – 68,1	d	
	5	57,0 – 45,2	d	
	10	37,9 – 36,5	s	
	19	19,2 – 16,7	q	
[3 α .5 β -OH]	1	31,0 – 29,6	t	100,0
	2	31,1 – 29,6	t	
	3	68,6 – 67,1	d	
	4	42,7 – 41,2	t	
	5	76,0 – 74,9	s	
	10	41,2 – 39,7	s	
	19	17,0 – 16,0	q	
[3 β .5 α -OH]	1	36,9 – 26,7	t	76,5
	2	34,4 – 30,1	t	
	3	68,4 – 66,4	d	
	4	45,9 – 34,9	t	
	5	80,8 – 73,5	s	
	10	47,0 – 37,0	s	
	19	25,0 – 15,6	q	
[3 β .5 β -OH]	1	26,7 – 24,7	t	66,7
	2	28,7 – 27,8	t	
	3	69,4 – 65,8	d	
	4	43,9 – 36,9	t	
	5	83,6 – 74,5	s	
	10	45,2 – 40,7	s	
	19	18,2 – 16,7	q	
[3 β -OR; 5 α -OH]	1	33,4 – 29,3	t	100,0
	2	27,0 – 26,5	t	
	3	71,6 – 70,4	d	
	4	41,5 – 32,2	t	
	5	77,1 – 73,3	s	
	10	41,5 – 38,0	s	
	19	24,5 – 15,5	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[3β,5α-OR]	1	32,7 – 30,7	t	100,0
	2	26,6 – 26,2	t	
	3	70,3 – 70,0	d	
	4	31,5 – 30,2	t	
	5	88,5 – 86,6	s	
	10	41,5 – 39,7	s	
	19	19,0 – 16,1	q	
[3β,19-OH]	1	33,0 – 32,5	t	100,0
	2	32,9 – 32,5	t	
	3	71,8 – 70,6	d	
	4	39,7 – 39,2	t	
	5	46,5 – 45,2	d	
	10	40,6 – 40,2	s	
	19	60,7 – 59,2	t	
[3β,19-OH; 5β-H]	1	24,3 – 23,0	t	100,0
	2	28,2 – 27,5	t	
	3	67,5 – 66,5	d	
	4	34,2 – 33,2	t	
	5	30,1 – 28,2	d	
	10	40,5 – 39,2	s	
	19	66,1 – 65,5	t	
[3α,4α,5α-OH]	1	24,0 – 23,5	t	100,0
	2	25,5 – 25,0	t	
	3	70,7 – 70,4	d	
	4	77,1 – 76,8	d	
	5	71,2 – 70,9	s	
	10	40,5 – 40,3	s	
	19	16,0 – 15,6	q	
[3β,4β,5α-OH]	1	33,2 – 32,7	t	100,0
	2	27,1 – 26,3	t	
	3	69,3 – 67,5	d	
	4	72,9 – 72,5	d	
	5	78,5 – 78,0	s	
	10	40,4 – 39,7	s	
	19	17,6 – 17,2	q	
[3β,5β,19-OH]	1	19,0 – 18,7	t	100,0
	2	27,7 – 27,3	t	
	3	67,8 – 67,3	d	
	4	37,0 – 36,5	t	
	5	78,1 – 77,8	s	
	10	42,3 – 41,7	s	
	19	65,5 – 65,0	t	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento
[2 α ,3 α -OH; 2 β :19-OXI]	1	43,7 – 39,9	t
	2	107,8 – 105,8	s
	3	74,1 – 73,1	d
	4	38,4 – 37,5	t
	5	42,9 – 39,7	d
	10	48,4 – 47,5	s
	19	67,0 – 66,0	t
[2 α ,3 α -OR; 2 β :19-OXI]	1	42,4 – 38,5	t
	2	108,7 – 108,4	s
	3	72,3 – 71,6	d
	4	38,2 – 33,7	t
	5	42,2 – 39,0	d
	10	48,0 – 47,5	s
	19	68,0 – 67,6	t
[3 β -OH; 5 α ,6 α -EP]	1	34,5 – 30,8	t
	2	31,0 – 28,7	t
	3	68,5 – 67,1	d
	4	40,0 – 38,7	t
	5	75,6 – 65,6	s
	10	38,0 – 35,0	s
	19	22,8 – 15,8	q
[3 β -OH; 5 β ,6 β -EP]	1	37,5 – 37,0	t
	2	31,7 – 31,4	t
	3	69,3 – 69,0	d
	4	42,2 – 41,8	t
	5	63,2 – 62,8	s
	10	35,5 – 35,0	s
	19	17,3 – 17,0	q
[3 β -OR; 5 α ,6 α -EP]	1	32,2 – 29,8	t
	2	27,2 – 27,0	t
	3	71,3 – 71,0	d
	4	36,2 – 35,2	t
	5	65,1 – 63,9	s
	10	38,4 – 35,0	s
	19	22,5 – 15,8	q
[1EN; 3 β -OH]	1	136,6 – 136,1	d
	2	132,5 – 132,0	d
	3	70,7 – 70,2	d
	4	44,0 – 43,5	t
	5	41,7 – 39,9	d
	10	37,2 – 29,3	s
	19	15,3 – 15,0	q

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[2EN]	1	40,2 – 39,4	t	100,0
	2	125,9 – 124,9	d	
	3	126,0 – 124,5	d	
	4	30,3 – 21,1	t	
	5	53,4 – 41,5	d	
	10	42,5 – 34,5	s	
	19	13,5 – 11,6	q	
[2,4EN]	1	36,0 – 35,7	t	100,0
	2	123,3 – 123,0	d	
	3	128,5 – 128,0	d	
	4	133,0 – 132,5	d	
	5	140,8 – 140,5	s	
	10	37,9 – 37,5	s	
	19	17,9 – 17,6	q	
[2EN; 4 α ,5 α -EP]	1	35,0 – 34,7	t	100,0
	2	122,5 – 122,1	d	
	3	132,4 – 132,1	d	
	4	52,4 – 52,1	d	
	5	69,5 – 69,2	s	
	10	36,8 – 36,5	s	
	19	14,5 – 14,0	q	
[3,5EN]	1	34,2 – 32,7	t	100,0
	2	23,7 – 23,3	t	
	3	136,8 – 129,8	d	
	4	127,6 – 124,5	d	
	5	161,3 – 141,6	s	
	10	36,5 – 35,5	s	
	19	18,8 – 16,6	q	
[4EN]	1	37,9 – 35,5	t	100,0
	2	19,3 – 17,8	t	
	3	25,7 – 25,5	t	
	4	132,3 – 119,0	d	
	5	146,0 – 144,8	s	
	10	37,7 – 37,0	s	
	19	20,2 – 19,3	q	
[3 β -OH; 4EN]	1	37,7 – 35,7	t	95,0
	2	29,8 – 29,0	t	
	3	68,5 – 68,0	d	
	4	124,5 – 120,6	d	
	5	151,3 – 151,0	s	
	10	37,6 – 37,2	s	
	19	22,3 – 19,7	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[3β-OR; 4EN]	1	37,2	–	35,2	t	87,7
	2	25,0	–	24,7	t	
	3	71,2	–	70,6	d	
	4	120,1	–	116,5	d	
	5	152,8	–	152,3	s	
	10	37,3	–	36,9	s	
	19	22,1	–	19,5	q	
[2α,3β-OH; 4EN]	1	44,8	–	44,5	t	100,0
	2	71,7	–	71,1	d	
	3	74,6	–	74,2	d	
	4	121,7	–	121,4	d	
	5	146,5	–	146,2	s	
	10	40,5	–	40,1	s	
	19	20,3	–	20,0	q	
[5EN]	1	40,0	–	38,7	t	100,0
	2	24,0	–	22,0	t	
	3	28,1	–	26,7	t	
	4	33,0	–	32,7	t	
	5	169,1	–	143,5	s	
	10	39,2	–	37,5	s	
	19	19,5	–	17,2	q	
[3β-OH; 5EN]	1	39,0	–	31,7	t	90,9
	2	32,7	–	30,7	t	
	3	72,6	–	70,0	d	
	4	44,4	–	35,2	t	
	5	167,3	–	132,8	s	
	10	48,4	–	36,0	s	
	19	22,8	–	16,2	q	
[3β-OR; 5EN]	1	38,9	–	27,7	t	77,3
	2	31,7	–	26,8	t	
	3	80,5	–	71,8	d	
	4	42,2	–	35,7	t	
	5	165,1	–	137,1	s	
	10	38,5	–	36,2	s	
	19	19,6	–	16,1	q	
[1β,3α-OH; 5EN]	1	78,0	–	77,6	d	100,0
	2	36,0	–	35,4	t	
	3	68,3	–	67,8	d	
	4	42,3	–	41,8	t	
	5	138,4	–	137,8	s	
	10	43,0	–	42,6	s	
	19	13,1	–	12,9	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[3β,4β-OH; 5EN]	1	37,0	–	36,5	t	100,0
	2	25,5	–	25,0	t	
	3	77,1	–	76,6	d	
	4	72,5	–	72,3	t	
	5	142,7	–	142,5	s	
	10	36,5	–	36,0	s	
	19	21,0	–	19,8	q	
[3β,19-OH; 5EN]	1	33,6	–	33,4	t	100,0
	2	32,0	–	31,5	t	
	3	71,5	–	71,0	d	
	4	42,5	–	42,1	t	
	5	136,3	–	135,8	s	
	10	42,0	–	41,6	s	
	19	63,1	–	62,6	t	
[3β-OR; 19-OH; 5EN]	1	33,7	–	33,2	t	100,0
	2	31,0	–	28,0	t	
	3	73,7	–	73,1	d	
	4	38,5	–	38,1	t	
	5	135,0	–	134,3	s	
	10	41,9	–	41,3	s	
	19	63,0	–	62,5	t	
[1EN; 3OXO]	1	159,3	–	157,8	d	100,0
	2	127,5	–	127,4	d	
	3	200,1	–	198,3	s	
	4	41,4	–	38,0	t	
	5	44,5	–	44,2	d	
	10	59,7	–	35,2	s	
	19	17,2	–	13,1	q	
[1,4EN; 3OXO]	1	159,5	–	149,6	d	100,0
	2	130,3	–	126,6	d	
	3	188,3	–	185,1	s	
	4	127,4	–	119,5	d	
	5	170,6	–	163,0	s	
	10	50,0	–	42,0	s	
	19	26,6	–	18,0	q	
[1,4EN; 2-OH; 3OXO]	1	124,5	–	123,6	d	100,0
	2	146,3	–	145,8	s	
	3	181,8	–	181,3	s	
	4	121,3	–	120,9	d	
	5	173,3	–	172,1	s	
	10	44,5	–	43,8	s	
	19	20,2	–	19,4	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[1,4EN; 4-Cl; 3OXO]	1	156,6 – 155,5	d	100,0
	2	126,1 – 125,1	d	
	3	178,6 – 178,3	s	
	4	130,3 – 126,5	s	
	5	162,6 – 158,8	s	
	10	47,9 – 46,2	s	
	19	20,7 – 19,2	q	
[1,4EN; 3OXO; 19-OR]	1	151,6 – 151,0	d	100,0
	2	130,6 – 130,2	d	
	3	186,0 – 185,5	s	
	4	127,0 – 126,3	d	
	5	163,9 – 163,4	s	
	10	47,9 – 47,4	s	
	19	63,6 – 63,3	t	
[1OXO; 2EN; 5 α -OH]	1	204,2 – 204,0	s	100,0
	2	128,6 – 128,4	d	
	3	141,6 – 141,3	d	
	4	36,0 – 35,7	t	
	5	77,5 – 77,1	s	
	10	54,2 – 52,2	s	
	19	15,1 – 14,0	q	
[1OXO; 2EN; 4 β ,5 β -OH]	1	201,3 – 201,0	s	100,0
	2	127,4 – 127,0	d	
	3	147,0 – 146,6	d	
	4	68,0 – 67,5	d	
	5	80,0 – 79,5	s	
	10	57,0 – 56,6	s	
	19	10,7 – 10,4	q	
[1OXO; 2EN; 4 β -OH; 5 β :6 β EP]	1	202,6 – 202,3	s	100,0
	2	131,5 – 131,3	d	
	3	143,5 – 142,8	d	
	4	69,5 – 69,5	d	
	5	63,4 – 63,2	s	
	10	47,9 – 47,5	s	
	19	17,2 – 15,8	q	
[1OXO; 2EN; 5 β :6 β EP]	1	203,6 – 201,8	s	100,0
	2	130,1 – 128,8	d	
	3	144,8 – 140,3	d	
	4	33,2 – 32,7	t	
	5	62,2 – 61,5	s	
	10	50,0 – 47,4	s	
	19	15,1 – 13,5	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[1OXO; 2,4EN]	1	206,1 – 191,6	s	100,0
	2	126,3 – 116,4	d	
	3	140,6 – 140,0	d	
	4	126,0 – 118,3	d	
	5	161,0 – 148,1	s	
	10	55,2 – 53,2	s	
	19	20,2 – 18,7	q	
[1OXO; 2,5EN]	1	204,5 – 202,8	s	100,0
	2	128,0 – 127,3	d	
	3	146,0 – 144,8	d	
	4	33,5 – 33,2	t	
	5	136,1 – 135,1	s	
	10	51,4 – 50,0	s	
	19	19,0 – 18,2	q	
[3OXO; 4EN]	1	39,9 – 25,5	t	100,0
	2	36,2 – 31,8	t	
	3	202,8 – 173,1	s	
	4	129,3 – 119,8	d	
	5	175,8 – 130,3	s	
	10	46,2 – 35,7	s	
	19	22,8 – 16,2	q	
[1 α -OR; 3OXO; 4EN]	1	75,4 – 73,8	d	100,0
	2	40,2 – 39,4	t	
	3	198,0 – 194,8	s	
	4	126,4 – 126,0	d	
	5	168,6 – 163,6	s	
	10	44,2 – 41,7	s	
	19	18,3 – 18,0	q	
[2 α ,19-OR; 3OXO; 4EN]	1	40,0 – 39,6	d	100,0
	2	71,1 – 70,7	t	
	3	193,8 – 193,5	s	
	4	125,3 – 124,9	d	
	5	165,0 – 164,3	s	
	10	44,2 – 42,4	s	
	19	66,6 – 66,3	q	
[3 β -OH; 4EN; 19OXO]	1	30,0 – 29,5	t	100,0
	2	29,7 – 29,3	t	
	3	66,9 – 66,5	d	
	4	130,1 – 129,7	d	
	5	138,5 – 138,0	s	
	10	54,5 – 54,0	d	
	19	205,0 – 204,6	d	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento
[3OXO; 4EN; 19-OH]	1	33,6 – 33,3	t		100,0
	2	35,2 – 34,6	t		
	3	200,3 – 199,6	s		
	4	127,3 – 126,0	d		
	5	168,5 – 165,6	s		
	10	44,2 – 43,6	s		
	19	66,0 – 65,0	t		
[3OXO; 4EN; 19-OR]	1	32,7 – 26,5	t		100,0
	2	34,5 – 33,2	t		
	3	199,0 – 198,5	s		
	4	126,9 – 114,8	d		
	5	171,6 – 164,8	s		
	10	45,9 – 41,7	s		
	19	75,5 – 66,5	t		
[1OXO; 5EN]	1	215,0 – 214,5	s		100,0
	2	38,7 – 38,3	t		
	3	27,5 – 27,1	t		
	4	31,5 – 31,0	t		
	5	141,5 – 141,1	s		
	10	54,0 – 53,5	s		
	19	19,5 – 19,0	q		
[4OXO; 5EN]	1	36,5 – 36,1	t		100,0
	2	19,3 – 19,0	t		
	3	40,3 – 39,9	t		
	4	203,4 – 203,0	s		
	5	145,5 – 145,1	s		
	10	38,7 – 38,4	s		
	19	21,5 – 21,1	q		
[1OXO; 2 α OH; 5 β :6 β EP]	1	209,9 – 209,4	s		100,0
	2	44,4 – 44,0	t		
	3	64,3 – 63,3	d		
	4	39,6 – 39,2	t		
	5	62,5 – 62,1	s		
	10	52,7 – 52,3	s		
	19	11,8 – 11,4	q		
[1OXO; 3 β OR; 5 β :6 β EP]	1	208,5 – 208,0	s		100,0
	2	42,0 – 41,4	t		
	3	66,0 – 65,3	d		
	4	36,3 – 35,7	t		
	5	60,9 – 60,5	s		
	10	52,2 – 51,7	s		
	19	11,3 – 11,0	q		

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[1OXO; 3 α :6 α OXI; 4 β OR; 5 β OH]	1	214,0 – 209,1	s	100,0
	2	44,2 – 42,2	t	
	3	77,5 – 76,3	d	
	4	75,5 – 73,3	d	
	5	80,1 – 78,5	s	
	10	55,7 – 54,9	s	
	19	15,4 – 15,0	q	
[1OXO; 5 β :6 β EP]	1	212,3 – 212,0	s	100,0
	2	30,7 – 30,4	t	
	3	30,7 – 30,5	t	
	4	36,0 – 35,8	t	
	5	65,3 – 65,1	s	
	10	51,7 – 51,5	s	
	19	14,3 – 14,1	q	
[1OXO; 5 β :6 β EP; 19OH]	1	210,8 – 210,5	s	100,0
	2	31,5 – 31,1	t	
	3	31,3 – 31,0	t	
	4	35,7 – 35,3	t	
	5	64,3 – 63,8	s	
	10	56,0 – 55,6	s	
	19	63,0 – 62,5	t	
[3OXO]	1	39,0 – 37,0	t	72,5
	2	38,2 – 28,1	t	
	3	219,8 – 208,6	s	
	4	46,5 – 37,0	t	
	5	57,5 – 40,2	d	
	10	41,2 – 34,4	s	
	19	18,1 – 10,6	q	
[3OXO; 5 β -H]	1	37,2 – 36,2	t	100,0
	2	37,4 – 36,4	t	
	3	213,0 – 211,5	s	
	4	42,4 – 41,7	t	
	5	44,7 – 43,4	d	
	10	35,4 – 34,4	s	
	19	22,7 – 22,2	q	
[3OXO; 4 β -OR]	1	39,0 – 38,6	t	100,0
	2	34,7 – 34,3	t	
	3	206,2 – 205,9	s	
	4	72,0 – 71,7	d	
	5	52,9 – 52,6	d	
	10	36,9 – 36,6	s	
	19	14,9 – 14,6	q	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[3OXO; 5 α -OH]	1	32,9 – 32,6	t	100,0
	2	38,0 – 37,6	t	
	3	211,3 – 211,0	s	
	4	52,0 – 51,5	t	
	5	77,7 – 77,3	s	
	10	39,5 – 39,1	s	
	19	16,0 – 15,7	q	
[3OXO; 5 α -OR]	1	32,9 – 32,6	t	100,0
	2	38,0 – 37,7	t	
	3	210,8 – 210,5	s	
	4	45,0 – 44,8	t	
	5	81,5 – 81,1	s	
	10	40,3 – 39,9	s	
	19	16,3 – 16,0	q	
[4OXO]	1	37,7 – 37,0	t	97,7
	2	22,9 – 22,5	t	
	3	41,4 – 41,0	d	
	4	213,5 – 212,6	t	
	5	59,3 – 59,0	s	
	10	42,8 – 42,5	s	
	19	14,1 – 13,8	t	
[3 β ,5 β -OH; 19OXO]	1	25,0 – 24,5		100,0
	2	27,5 – 27,0		
	3	67,3 – 66,8		
	4	38,2 – 37,8		
	5	75,5 – 75,0		
	10	56,0 – 55,5		
	19	196,0 – 195,5		
[3 β ,5 β -OR; 19OXO]	1	23,3 – 22,8		100,0
	2	24,1 – 23,6		
	3	68,1 – 68,0		
	4	31,7 – 31,7		
	5	83,5 – 83,0		
	10	53,5 – 53,2		
	19	205,6 – 205,5		
[3Me; 4EN]	1	37,4 – 37,0	t	100,0
	2	29,8 – 29,5	t	
	3	148,8 – 148,5	s	
	4	122,5 – 122,3	d	
	5	143,8 – 143,5	s	
	10	37,5 – 37,3	s	
	19	18,5 – 18,2	q	
	Me	107,8 – 107,5	t	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[3 β -OH; 4 α -Me]	1	37,0 – 35,0	t	100,0
	2	31,2 – 30,8	t	
	3	76,5 – 76,0	d	
	4	40,4 – 39,2	d	
	5	50,2 – 46,5	d	
	10	37,0 – 34,7	s	
	19	18,2 – 14,1	q	
	Me	15,3 – 15,0	q	
[3 β -OR; 4 α -Me]	1	36,5 – 34,2	t	100,0
	2	27,3 – 27,0	t	
	3	78,5 – 78,1	d	
	4	36,9 – 36,7	d	
	5	46,5 – 45,5	d	
	10	36,2 – 34,5	s	
	19	20,5 – 13,8	q	
	Me	15,1 – 14,8	q	
[3-OXO; 4 α -Me]	1	39,4 – 37,0	t	100,0
	2	38,3 – 37,9	t	
	3	213,6 – 213,2	s	
	4	45,5 – 45,0	d	
	5	50,2 – 49,9	d	
	10	36,5 – 35,0	s	
	19	17,5 – 13,6	q	
	Me	11,6 – 11,2	q	
[3-OXO; 5 α -Me]	1	32,9 – 32,6	t	100,0
	2	37,9 – 37,5	t	
	3	212,3 – 212,0	s	
	4	52,5 – 52,1	t	
	5	40,0 – 39,6	s	
	10	37,5 – 37,2	s	
	19	14,3 – 14,0	q	
	Me	21,0 – 20,7	q	
[1,3,5-EN; 3-OH; 19NOR]	1	131,3 – 126,5	d	100,0
	2	113,8 – 112,5	d	
	3	156,1 – 153,1	s	
	4	116,0 – 115,1	d	
	5	139,5 – 138,1	s	
	10	132,6 – 127,8	s	
[1,3,5-EN; 3-OR; 19NOR]	1	127,3 – 126,0	d	100,0
	2	119,5 – 111,1	d	
	3	158,8 – 149,5	s	
	4	122,3 – 113,5	d	
	5	138,6 – 137,3	s	
	10	138,6 – 131,8	s	

Tabela XIII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[3-OXO; 4EN; 19NOR]	1	27,7 – 26,0	t	100,0
	2	36,7 – 35,2	t	
	3	199,8 – 198,0	s	
	4	126,6 – 124,6	d	
	5	167,1 – 154,5	s	
	10	42,9 – 37,5	d	
[1,3,5-EN; 1-Me; 19NOR]	1	138,3 – 125,3	s	100,0
	2	133,8 – 125,1	d	
	3	129,1 – 126,0	d	
	4	135,1 – 125,3	d	
	5	137,1 – 135,8	s	
	10	144,1 – 136,3	s	
	Me	25,0 – 22,2	q	
[1,3,5-EN; 4-Me; 19NOR]	1	123,1 – 122,9	d	100,0
	2	132,3 – 125,1	d	
	3	130,6 – 127,1	d	
	4	141,3 – 134,8	s	
	5	136,3 – 131,0	s	
	10	147,8 – 139,6	s	
	Me	23,3 – 19,7	q

[3 α ,5 α -Ciclo]	1	33,9 – 31,2	t	76,2
	2	26,6 – 24,0	t	
	3	36,0 – 19,2	d	
	4	14,6 – 7,1	t	
	5	46,7 – 35,2	s	
	10	46,7 – 36,7	s	
[2OXO]	19	20,8 – 17,7	q	100,0
	2	176,6 – 175,6		
	3	29,1 – 28,6		
	4	30,2 – 26,1		
	5	89,0 – 85,6		
	10	42,5 – 41,6		
	19	14,3 – 7,0

Tabela XIV. Faixas de deslocamento químico características de esteróides – Anel B

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
Anel B

[5 α -H]	5	59,2 – 30,7	d	83,3
	6	31,8 – 20,5	t	
	7	36,9 – 20,7	t	
	8	43,7 – 31,0	d	
	9	67,0 – 34,2	d	
	10	59,7 – 29,3	s	
	19	23,8 – 5,5	q	
[5 β -H]	5	57,2 – 32,2	d	100,0
	6	28,6 – 20,1	t	
	7	27,6 – 20,2	t	
	8	43,2 – 31,7	d	
	9	52,0 – 29,5	d	
	10	42,0 – 34,0	s	
	19	27,7 – 12,5	q	
[5 α -OH]	5	78,5 – 71,0	s	100,0
	6	35,5 – 24,5	t	
	7	28,7 – 25,3	t	
	8	35,5 – 34,2	d	
	9	46,2 – 45,5	d	
	10	43,9 – 36,5	s	
[5 β -OH]	5	76,5 – 73,0	s	100,0
	6	37,4 – 32,0	t	
	7	24,8 – 23,2	t	
	8	42,7 – 38,0	d	
	9	44,4 – 32,5	d	
	10	43,7 – 39,7	s	
	19	17,6 – 16,0	q	
[5 α -OR]	5	87,6 – 81,3	s	100,0
	6	28,2 – 26,8	t	
	7	26,7 – 26,1	t	
	8	35,0 – 34,7	d	
	9	46,2 – 45,0	d	
	10	40,2 – 40,0	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[6 α -OH]	5	53,7 – 51,5	d	100,0
	6	70,0 – 69,3	d	
	7	41,9 – 41,5	t	
	8	34,5 – 34,0	d	
	9	54,2 – 53,7	d	
	10	36,9 – 36,2	s	
	19	13,5 – 13,3	q	
[6 β -OH]	5	49,7 – 47,2	d	100,0
	6	72,5 – 71,9	d	
	7	41,9 – 39,5	t	
	8	31,2 – 30,3	d	
	9	55,7 – 54,2	d	
	10	36,5 – 35,4	s	
	19	16,2 – 15,8	q	
[6 α -OR]	5	52,7 – 44,2	d	100,0
	6	73,4 – 66,9	d	
	7	38,2 – 35,5	t	
	8	34,0 – 33,5	d	
	9	54,5 – 45,5	d	
	10	36,7 – 36,2	s	
	19	15,0 – 13,3	q	
[6 β -OR]	5	48,7 – 45,5	d	100,0
	6	74,1 – 71,9	d	
	7	36,7 – 36,0	t	
	8	31,2 – 29,8	d	
	9	54,7 – 53,5	d	
	10	37,0 – 35,4	s	
	19	15,8 – 15,1	q	
[7 α -OH]	5	42,2 – 41,9	d	100,0
	6	35,5 – 35,0	t	
	7	68,9 – 67,8	d	
	8	40,0 – 39,6	d	
	9	33,2 – 26,7	d	
	10	35,5 – 35,0	s	
	19	23,2 – 22,7	q	
[7 β -OH]	5	41,9 – 39,2	d	66,7
	6	38,4 – 36,7	t	
	7	74,6 – 68,1	d	
	8	43,4 – 39,7	d	
	9	53,5 – 46,4	d	
	10	41,2 – 35,0	s	
	19	12,5 – 6,5	q	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[7β-OR]	5	41,5 – 38,7	d	80,0
	6	33,7 – 32,7	t	
	7	76,4 – 71,5	d	
	8	39,5 – 38,4	d	
	9	54,0 – 47,4	d	
	10	39,9 – 34,9	s	
	19	12,3 – 7,6	q	
[5α,6α-OH]	5	78,8 – 76,5	s	100,0
	6	72,0 – 70,5	d	
	7	36,7 – 34,5	t	
	8	34,5 – 33,7	d	
	9	44,7 – 43,7	d	
	10	41,9 – 39,2	s	
[5α,6β-OH]	5	77,5 – 75,5	s	90,0
	6	76,5 – 74,3	d	
	7	36,0 – 26,5	t	
	8	35,4 – 30,2	d	
	9	46,5 – 35,2	d	
	10	54,2 – 38,4	s	
[5α-OH; 6α-OR]	5	76,6 – 75,5	s	100,0
	6	74,9 – 74,1	d	
	7	32,2 – 30,6	t	
	8	34,7 – 33,2	d	
	9	51,2 – 44,2	d	
	10	41,9 – 39,7	s	
[5α-OH; 6β-OR]	5	75,8 – 74,3	s	100,0
	6	85,6 – 76,4	d	
	7	31,7 – 29,1	t	
	8	31,7 – 30,5	d	
	9	51,7 – 44,9	d	
	10	40,4 – 38,0	s	
[5β-OH; 6α-OR]	5	80,1 – 78,5	s	100,0
	6	77,3 – 71,1	d	
	7	29,7 – 26,3	t	
	8	35,7 – 30,2	d	
	9	34,9 – 29,7	d	
	10	55,7 – 54,9	s	
[5β,19-OH]	5	83,3 – 78,0	s	100,0
	6	35,4 – 25,7	t	
	7	27,6 – 24,0	t	
	8	40,7 – 34,0	d	
	9	43,5 – 39,0	d	
	10	43,9 – 42,0	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[6 α ,7 β -OH]	5	47,7 – 42,2	d	100,0
	6	75,3 – 73,3	d	
	7	80,5 – 75,3	d	
	8	41,7 – 41,0	d	
	9	44,7 – 32,2	d	
	10	35,7 – 35,0	s	
[6 α ,8 β -OH]	5	57,0 – 52,2	d	100,0
	6	67,5 – 64,0	d	
	7	46,7 – 43,2	t	
	8	79,4 – 75,9	s	
	9	57,4 – 47,5	d	
	10	37,7 – 36,2	s	
[5 α ,6 α ,8 β -OH]	5	78,5 – 78,0	s	100,0
	6	66,5 – 65,8	d	
	7	40,4 – 40,0	t	
	8	79,1 – 78,8	s	
	9	41,4 – 40,9	d	
	10	40,0 – 39,7	s	
[5 α ,6 β ,8 β -OH]	5	76,0 – 75,8	s	100,0
	6	78,0 – 77,7	d	
	7	42,0 – 41,7	t	
	8	76,8 – 76,5	s	
	9	48,9 – 48,6	d	
	10	39,3 – 39,0	s	
[6 α ,7 α ,8 β -OH]	5	47,9 – 44,5	d	100,0
	6	68,9 – 66,0	d	
	7	76,6 – 76,4	d	
	8	77,7 – 77,5	s	
	9	52,0 – 51,2	d	
	10	37,9 – 37,7	s	
[6 β ,7 α ,8 β -OH]	5	45,2 – 42,7	d	100,0
	6	80,0 – 74,0	d	
	7	78,4 – 73,4	d	
	8	78,1 – 77,8	s	
	9	52,0 – 51,4	d	
	10	36,7 – 36,3	s	
[6 α -OR: 7 α ,8 β -OH]	5	44,2 – 40,5	d	100,0
	6	71,3 – 69,6	d	
	7	73,0 – 72,5	d	
	8	77,3 – 77,1	s	
	9	49,2 – 48,2	d	
	10	37,4 – 36,9	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[5 α ,6 α .7 β ,8 β -OH]	5	78,5 – 78,0	s	100,0
	6	71,0 – 70,7	d	
	7	73,6 – 73,2	d	
	8	79,9 – 79,5	s	
	9	41,0 – 40,4	d	
	10	40,0 – 39,4	s	
[5 α ,7 β ,8 β -OH: 6 α -OR]	5	78,6 – 78,3	s	100,0
	6	79,9 – 79,5	d	
	7	72,9 – 72,6	d	
	8	79,9 – 79,6	s	
	9	40,6 – 40,3	d	
	10	40,5 – 40,2	s	
[4,5EP]	5	70,2 – 70,0	s	100,0
	6	29,7 – 29,1	t	
	7	30,3 – 28,8	t	
	8	35,4 – 35,0	d	
	9	50,7 – 46,4	d	
	10	37,0 – 36,7	s	
[5 α ,6 α EP]	5	68,5 – 64,0	s	100,0
	6	59,7 – 57,0	d	
	7	30,8 – 27,7	t	
	8	30,5 – 29,5	d	
	9	43,0 – 42,5	d	
	10	38,5 – 35,0	s	
[5 β ,6 β EP]	5	66,1 – 60,7	s	83,3
	6	64,4 – 58,7	d	
	7	32,5 – 21,5	t	
	8	40,2 – 28,1	d	
	9	51,2 – 35,7	d	
	10	55,9 – 35,2	s	
[4EN]	5	172,1 – 167,5	s	95,0
	6	33,7 – 31,6	t	
	7	32,7 – 25,6	t	
	8	38,9 – 31,0	d	
	9	63,2 – 46,2	d	
	10	42,0 – 35,7	s	
	19	20,5 – 16,7	q	
[4,6EN]	5	163,0 – 162,6	s	100,0
	6	129,0 – 128,5	d	
	7	138,5 – 138,1	d	
	8	37,7 – 36,8	d	
	9	48,9 – 48,4	d	
	10	36,4 – 35,7	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento
[4EN; 6 α -OH]	5	171,1	– 160,8	s	75,0
	6	71,9	– 68,1	d	
	7	42,7	– 39,7	t	
	8	35,5	– 34,2	d	
	9	54,9	– 52,2	d	
	10	47,9	– 39,0	s	
[4EN; 6 β -OH]	5	170,8	– 158,8	s	100,0
	6	73,8	– 68,5	d	
	7	39,7	– 34,5	t	
	8	33,2	– 29,5	d	
	9	59,7	– 51,5	d	
	10	46,4	– 38,7	s	
[2,4EN; 6 β -OH]	5	161,0	– 160,0	s	100,0
	6	74,1	– 72,8	d	
	7	37,2	– 29,7	t	
	8	37,7	– 34,9	d	
	9	44,7	– 43,7	d	
	10	55,2	– 53,2	s	
[4EN; 6 β -Br]	5	166,0	– 163,1	s	100,0
	6	52,2	– 49,9	d	
	7	41,0	– 40,2	t	
	8	30,5	– 29,6	d	
	9	54,9	– 52,2	d	
	10	44,9	– 38,2	s	
[4EN; 7 α -OH]	5	170,8	– 163,6	s	100,0
	6	42,2	– 40,9	t	
	7	70,5	– 67,0	d	
	8	40,9	– 37,7	d	
	9	45,7	– 36,7	d	
	10	45,2	– 38,5	s	
[4EN; 7 α -OR]	5	164,1	– 163,0	s	100,0
	6	37,2	– 36,9	t	
	7	72,0	– 70,6	d	
	8	40,4	– 34,5	d	
	9	46,4	– 38,2	d	
	10	43,2	– 42,5	s	
[4EN; 7-CH ₂ R]	5	169,6	– 168,8	s	100,0
	6	38,9	– 34,5	t	
	7	41,9	– 38,5	d	
	8	39,7	– 35,4	d	
	9	54,5	– 46,5	d	
	10	38,9	– 38,4	s	
	-CH ₂	38,0	27,8	t	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[4EN; 9α-OH]	5	168,3 – 168,1	s	100,0
	6	32,0 – 31,4	t	
	7	25,2 – 24,8	t	
	8	37,5 – 37,1	d	
	9	76,7 – 76,0	s	
	10	44,6 – 44,2	s	
[4EN; 9α,11αEP]	5	164,0 – 163,6	s	75,0
	6	32,6 – 32,4	t	
	7	29,9 – 29,3	t	
	8	35,9 – 35,3	d	
	9	68,2 – 67,8	s	
	10	45,2 – 45,0	s	
	11	54,7 – 54,5	d	
[5EN]	5	145,3 – 134,8	s	82,2
	6	132,5 – 119,0	d	
	7	34,0 – 26,0	t	
	8	38,2 – 27,3	d	
	9	53,9 – 39,5	d	
	10	53,9 – 35,5	s	
[5,7EN]	5	141,0 – 140,6	s	100,0
	6	119,9 – 119,5	d	
	7	117,3 – 116,9	d	
	8	141,3 – 140,8	s	
	9	47,0 – 46,5	d	
	10	37,8 – 37,2	s	
[5EN; 7β-OH]	5	143,8 – 142,1	s	100,0
	6	126,3 – 125,5	d	
	7	73,6 – 72,8	d	
	8	40,7 – 39,5	d	
	9	48,2 – 46,9	d	
	10	37,4 – 36,5	s	
[5EN; 8-OH]	5	140,8 – 137,1	s	100,0
	6	120,0 – 117,5	d	
	7	37,0 – 34,0	t	
	8	74,9 – 73,0	s	
	9	44,4 – 43,0	d	
	10	37,0 – 36,2	s	
[7EN]	5	50,2 – 35,5	d	100,0
	6	38,0 – 26,5	t	
	7	122,3 – 117,0	d	
	8	139,6 – 134,8	s	
	9	57,9 – 48,9	d	
	10	35,0 – 33,9	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[5 α ,9 α -OH; 6 α -OR; 7EN]	5	76,2	–	75,8	s	100,0
	6	79,9	–	79,7	d	
	7	119,5	–	119,0	d	
	8	144,7	–	144,5	s	
	9	78,0	–	77,8	s	
	10	42,6	–	42,4	s	
[7EN; 9 α ,11 α EP]	5	36,8	–	36,6	d	100,0
	6	29,9	–	29,3	t	
	7	125,3	–	124,9	d	
	8	134,8	–	134,5	s	
	9	64,5	–	64,0	s	
	10	34,6	–	34,3	s	
[7,9(11)EN]	5	45,5	–	39,0	d	100,0
	6	29,7	–	26,3	t	
	7	119,0	–	118,0	d	
	8	143,0	–	135,6	s	
	9	143,3	–	143,1	s	
	10	36,2	–	35,5	s	
[5 α -OH; 7,9(11)EN]	5	73,5	–	73,1	s	100,0
	6	37,5	–	37,0	t	
	7	115,9	–	115,6	d	
	8	136,5	–	136,0	s	
	9	140,3	–	140,0	s	
	10	41,2	–	39,9	s	
[8EN]	5	49,9	–	47,0	d	100,0
	6	22,1	–	20,7	t	
	7	25,6	–	25,2	t	
	8	133,6	–	132,3	s	
	9	135,6	–	134,6	s	
	10	36,5	–	36,2	s	
[8(14)EN]	5	50,2	–	44,0	d	100,0
	6	28,7	–	26,0	t	
	7	29,7	–	29,0	t	
	8	128,6	–	126,0	s	
	9	49,7	–	49,0	d	
	10	37,0	–	35,7	s	
[5 α ,6 α EP; 7 α -OR; 8EN]	5	64,0	–	63,8	s	100,0
	6	59,5	–	59,3	d	
	7	70,5	–	70,1	d	
	8	137,5	–	137,0	s	
	9	123,2	–	123,0	s	
	10	38,6	–	38,3	s	

BIBLIOTE

BIBLIOTECA
INSTITUTO DE QUÍMICA
Universidade de São Paulo

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[9(11)EN]	5	43,5	–	42,7	d	100,0
	6	28,7	–	27,7	t	
	7	33,5	–	27,8	t	
	8	41,7	–	36,0	d	
	9	172,1	–	145,8	s	
	10	39,5	–	38,0	s	
[5OR; 6 α ,7 α EP; 9(11)EN]	5	85,8	–	85,6	s	100,0
	6	55,0	–	54,7	d	
	7	54,5	–	53,7	d	
	8	36,5	–	35,4	d	
	9	139,8	–	132,3	s	
	10	42,5	–	41,4	d	
[4EN; 6OXO]	5	158,3	–	140,6	s	100,0
	6	203,1	–	200,0	s	
	7	46,0	–	45,5	t	
	8	33,7	–	32,7	d	
	9	51,7	–	50,2	d	
	10	39,5	–	37,7	s	
[4EN; 6=NOH]	5	162,4	–	162,0	s	100,0
	6	156,2	–	156,0	s	
	7	29,9	–	29,5	t	
	8	32,8	–	32,4	d	
	9	51,5	–	51,1	d	
	10	39,0	–	38,6	s	
[5EN; 7OXO]	5	169,1	–	161,1	s	100,0
	6	124,3	–	124,0	d	
	7	202,0	–	200,8	s	
	8	46,4	–	44,2	d	
	9	48,7	–	45,7	d	
	10	39,2	–	36,2	s	
[6OXO; 7EN]	5	54,5	–	48,0	d	93,3
	6	208,1	–	204,1	s	
	7	123,3	–	122,5	d	
	8	166,8	–	164,3	s	
	9	51,2	–	47,2	d	
	10	39,7	–	22,7	s	
[6OXO; 7EN; 5 β -H]	5	51,4	–	50,5	d	100,0
	6	204,5	–	175,6	s	
	7	122,0	–	119,3	d	
	8	166,6	–	162,8	s	
	9	35,9	–	33,5	d	
	10	38,7	–	37,7	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[5 α -OH; 6OXO; 7EN]	5	76,9 – 76,5	s	100,0
	6	199,8 – 199,5	s	
	7	119,6 – 119,3	d	
	8	166,3 – 166,0	s	
	9	43,9 – 43,5	d	
	10	40,8 – 40,5	s	
[5 β -OH; 6OXO; 7EN]	5	79,9 – 79,7	s	100,0
	6	201,0 – 200,8	s	
	7	120,0 – 119,6	d	
	8	167,0 – 166,8	s	
	9	38,5 – 38,1	d	
	10	44,9 – 44,6	s	
[5 α ,6 α -OH; 7EN; 9OXO]	5	80,9 – 80,6	s	100,0
	6	71,6 – 71,3	d	
	7	144,2 – 143,9	d	
	8	137,6 – 137,5	s	
	9	206,1 – 206,0	s	
	10	47,1 – 46,8	s	
[6OXO]	5	58,7 – 53,4	d	100,0
	6	212,3 – 207,3	s	
	7	47,0 – 37,0	t	
	8	41,0 – 36,2	d	
	9	55,0 – 46,7	d	
	10	42,5 – 36,5	s	
[6OXO; 5 β -H]	5	40,4 – 40,1	d	100,0
	6	215,3 – 214,8	s	
	7	42,7 – 42,5	t	
	8	36,9 – 36,6	d	
	9	53,5 – 53,2	d	
	10	40,2 – 39,2	s	
[4 α ,5 α EP; 6OXO]	5	69,3 – 53,7	s	100,0
	6	207,6 – 205,6	s	
	7	45,7 – 44,9	t	
	8	37,0 – 36,5	d	
	9	49,4 – 45,4	d	
	10	38,5 – 36,7	s	
[4 β ,5 β EP; 6OXO]	5	67,7 – 67,4	s	100,0
	6	207,2 – 206,8	s	
	7	46,4 – 46,0	t	
	8	35,0 – 34,7	d	
	9	46,9 – 46,5	d	
	10	38,3 – 38,0	s	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento
[7OXO]	5	49,0 – 45,7 d	100,0
	6	46,5 – 45,4 t	
	7	211,3 – 207,6 s	
	8	50,0 – 48,7 d	
	9	55,9 – 54,7 d	
	10	36,7 – 36,4 s	
[7OXO; 5 β -H]	5	41,7 – 30,0 d	100,0
	6	37,2 – 31,2 t	
	7	211,3 – 211,1 s	
	8	57,5 – 56,7 d	
	9	59,5 – 57,4 d	
	10	55,9 – 41,2 s	
[5 β -OR; 19OXO]	5	83,5 – 83,0 s	100,0
	6	29,3 – 28,8 t	
	7	24,5 – 22,0 t	
	8	41,7 – 40,7 d	
	9	38,7 – 38,0 d	
	10	53,5 – 53,2 s	
	19	205,6 – 205,5 d	
[6OXO;6:7OXI; 6,7SECO]	5	44,2 – 40,2 d	100,0
	6	179,5 – 178,8 s	
	7	72,6 – 71,8 t	
	8	40,0 – 36,9 d	
	9	59,9 – 53,9 d	
	10	44,2 – 28,5 s	
[6 α -Me]	5	52,5 – 50,9 d	100,0
	6	31,8 – 30,8 d	
	7	41,7 – 41,2 t	
	8	35,0 – 34,7 d	
	9	54,2 – 53,7 d	
	10	35,7 – 35,5 s	
	Me	20,2 – 19,7 q	
[6 β -Me]	5	48,5 – 46,2 d	100,0
	6	33,7 – 33,2 d	
	7	41,2 – 39,2 t	
	8	30,7 – 30,5 d	
	9	54,7 – 54,2 d	
	10	36,0 – 35,7 s	
	Me	15,6 – 14,8 q	

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento
[4EN; 6 α -Me]	5	174,3 – 151,1	s
	6	33,4 – 32,4	t
	7	41,9 – 41,0	t
	8	35,7 – 35,2	d
	9	54,5 – 53,9	d
	10	38,7 – 37,2	s
	Me	18,8 – 18,2	q
[4EN; 6 β -Me]	5	175,8 – 151,1	s
	6	38,4 – 37,4	t
	7	38,7 – 37,4	t
	8	30,7 – 30,2	d
	9	54,5 – 53,2	d
	10	38,2 – 37,0	s
	Me	23,1 – 22,8	q
[4EN; 19NOR]	5	167,1 – 154,5	s
	6	36,5 – 34,0	t
	7	32,0 – 29,2	t
	8	41,0 – 34,2	d
	9	59,4 – 49,2	d
	10	42,9 – 37,5	d
[5(10)EN; 19NOR]	5	139,5 – 129,0	s
	6	30,7 – 28,3	t
	7	27,3 – 26,2	t
	8	42,0 – 37,5	d
	9	50,9 – 43,4	d
	10	132,1 – 127,8	s
[5(10)EN; 6OXO]	5	137,1 – 131,0	s
	6	199,5 – 198,3	s
	7	45,7 – 41,4	t
	8	38,6 – 38,2	d
	9	46,4 – 43,4	d
	10	147,8 – 144,1	s

[6 β -OH]	5	39,2 – 38,9	s
	6	73,9 – 73,3	d
	7	38,0 – 33,7	t
	8	29,8 – 32,7	d
	9	48,4 – 41,2	d
	10	43,5 – 42,9	s

Tabela XIV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[6β-OR]	5	36,4 – 35,0	s	100,0
	6	84,0 – 82,1	d	
	7	36,0 – 34,9	t	
	8	31,7 – 30,2	d	
	9	50,0 – 48,0	d	
	10	44,5 – 42,7	s	
[6OXO; 7EN]	5	44,9 – 44,5	s	100,0
	6	196,8 – 196,5	d	
	7	123,6 – 123,2	t	
	8	163,8 – 163,3	d	
	9	41,9 – 41,6	d	
	10	44,5 – 44,0	s	
[6OXO]	5	46,7 – 46,0	s	100,0
	6	209,8 – 207,3	d	
	7	46,5 – 38,9	t	
	8	37,5 – 34,2	d	
	9	48,5 – 40,0	d	
	10	46,7 – 42,0	s	

Tabela XV. Faixas de deslocamento químico características de esteróides – Anel C

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
<div>Anel C</div>
[14 α -H]	8	53,4 – 27,3	d	100,0
	9	59,9 – 33,2	d	
	11	37,7 – 18,0	t	
	12	42,7 – 22,7	t	
	13	49,2 – 36,9	s	
	14	66,0 – 45,0	d	
	18	21,8 – 10,8	q	
[14 β -H]	8	29,7 – 29,2	d	100,0
	9	43,4 – 42,9	d	
	11	20,7 – 20,0	t	
	12	31,8 – 24,8	t	
	13	47,9 – 40,2	s	
	14	49,2 – 45,2	d	
	18	24,2 – 17,1	q	
[8 β -OH]	8	78,1 – 74,8	s	100,0
	9	57,4 – 48,2	d	
	11	19,7 – 17,1	t	
	12	43,2 – 40,2	t	
	13	45,5 – 43,9	s	
	14	66,5 – 53,7	d	
	18	16,8 – 15,6	q	
[9 α -OH]	8	37,5 – 37,2	d	100,0
	9	76,4 – 76,2	s	
	11	26,7 – 26,5	t	
	12	34,0 – 32,0	t	
	13	43,7 – 42,7	s	
	14	49,4 – 44,0	d	
	18	12,5 – 10,1	q	
[11 α -OH]	8	45,2 – 29,5	d	100,0
	9	61,4 – 48,2	d	
	11	69,4 – 67,4	d	
	12	51,5 – 46,7	t	
	13	45,0 – 40,7	s	
	14	56,0 – 49,5	d	
	18	18,3 – 13,3	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[11β-OH]	8	34,2 – 31,0	d	100,0
	9	59,2 – 46,4	d	
	11	70,3 – 66,1	d	
	12	47,7 – 38,7	t	
	13	47,9 – 39,9	s	
	14	54,7 – 50,0	d	
	18	20,2 – 15,1	q	
[11α-OR]	8	35,4 – 28,1	d	95,0
	9	57,0 – 46,7	d	
	11	71,6 – 69,3	d	
	12	45,5 – 43,2	t	
	13	43,7 – 40,9	s	
	14	55,4 – 49,2	d	
	18	18,1 – 12,1	q	
[12α-OH]	8	41,7 – 35,0	d	100,0
	9	48,2 – 26,2	d	
	11	31,6 – 28,0	t	
	12	73,4 – 72,5	d	
	13	47,5 – 45,0	s	
	14	48,4 – 41,0	d	
	18	18,7 – 12,5	q	
[12β-OH]	8	44,2 – 34,0	d	100,0
	9	53,9 – 47,4	d	
	11	31,3 – 29,8	t	
	12	80,6 – 77,9	d	
	13	46,9 – 45,9	s	
	14	54,9 – 49,2	d	
	18	11,8 – 10,3	q	
[12α-OR]	8	39,0 – 34,9	d	100,0
	9	38,2 – 36,4	d	
	11	27,2 – 26,7	t	
	12	75,8 – 74,5	d	
	13	45,2 – 44,7	s	
	14	50,0 – 42,7	d	
	18	12,8 – 12,1	q	
[14α-OH]	8	41,0 – 32,7	d	100,0
	9	46,7 – 33,0	d	
	11	23,3 – 19,1	t	
	12	38,7 – 25,1	t	
	13	55,5 – 47,0	s	
	14	86,5 – 80,6	s	
	18	21,3 – 14,8	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[14 β -OH]	8	43,7 – 34,7	d	100,0
	9	51,4 – 35,2	d	
	11	24,0 – 17,7	t	
	12	41,4 – 21,1	t	
	13	52,5 – 46,0	s	
	14	89,1 – 80,9	s	
	18	19,2 – 12,8	q	
[8 β .14 α -OH]	8	79,9 – 78,4	s	100,0
	9	49,5 – 40,5	d	
	11	18,7 – 17,6	t	
	12	40,4 – 35,5	t	
	13	49,7 – 46,9	s	
	14	85,1 – 83,3	s	
	18	17,3 – 16,7	q	
[9 α .14 α -OH]	8	41,3 – 40,9	d	100,0
	9	81,2 – 80,8	s	
	11	30,0 – 29,5	t	
	12	30,1 – 29,6	t	
	13	48,7 – 48,0	s	
	14	88,6 – 88,0	s	
	18	18,4 – 18,0	q	
[9 α -F: 11 β -OH]	8	34,7 – 30,7	d	100,0
	9	101,1 – 100,0	s	
	11	70,8 – 69,0	d	
	12	43,0 – 35,2	t	
	13	47,7 – 43,5	s	
	14	48,2 – 43,4	d	
	18	16,2 – 15,1	q	
[11 β .14 β -OH]	8	37,3 – 36,7	d	100,0
	9	54,1 – 53,4	d	
	11	66,8 – 66,1	d	
	12	48,6 – 48,0	t	
	13	50,4 – 49,9	s	
	14	85,9 – 85,2	s	
	18	19,7 – 19,4	q	
[12 α .14 β -OH]	8	42,3 – 41,7	d	100,0
	9	29,7 – 29,1	d	
	11	29,9 – 29,5	t	
	12	76,8 – 76,3	d	
	13	52,3 – 51,8	s	
	14	85,9 – 85,3	s	
	18	17,4 – 17,0	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[12 β ,14 β -OH]	8	41,4 – 41,2	d	100,0
	9	32,7 – 32,4	d	
	11	30,2 – 30,0	t	
	12	75,0 – 74,8	d	
	13	56,4 – 55,5	s	
	14	85,8 – 85,5	s	
	18	9,3 – 9,0	q	
[12 β -OR; 14 β -OH]	8	41,2 – 41,0	d	98,3
	9	33,5 – 32,4	d	
	11	28,1 – 26,5	t	
	12	85,0 – 77,4	d	
	13	54,7 – 54,2	s	
	14	85,7 – 85,4	s	
	18	10,4 – 10,0	q	
[8,12 β ,14 β -OH]	8	74,1 – 73,0	s	100,0
	9	43,7 – 43,0	d	
	11	28,0 – 26,8	t	
	12	71,6 – 67,6	d	
	13	58,7 – 56,4	s	
	14	88,5 – 86,3	s	
	18	13,8 – 8,6	q	
[8,14 β -OH; 12 β -OR]	8	74,9 – 74,4	s	97,5
	9	44,4 – 43,9	d	
	11	24,3 – 24,2	t	
	12	73,9 – 71,0	d	
	13	55,5 – 54,0	s	
	14	86,9 – 86,3	s	
	18	15,1 – 12,0	q	
[9 α ,11 α EP]	8	35,7 – 35,4	d	100,0
	9	68,1 – 67,8	s	
	11	54,7 – 54,5	d	
	12	30,3 – 30,0	t	
	13	46,4 – 46,0	s	
	14	41,5 – 41,2	d	
	18	17,1 – 16,7	q	
[14 β ,15 β EP]	8	42,7 – 39,2	d	100,0
	9	33,5 – 32,5	d	
	11	21,5 – 20,6	t	
	12	40,0 – 39,4	t	
	13	45,3 – 45,0	s	
	14	74,6 – 72,5	s	
	18	17,2 – 16,7	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[7EN]	8	143,8 – 138,8	s	100,0
	9	49,7 – 43,0	d	
	11	21,7 – 21,1	t	
	12	44,2 – 39,0	t	
	13	43,5 – 42,7	s	
	14	55,2 – 54,2	d	
	18	12,1 – 11,6	q	
[7EN]	8	166,8 – 163,6	s	100,0
	9	51,2 – 35,9	d	
	11	22,8 – 21,8	t	
	12	39,5 – 38,7	t	
	13	45,5 – 43,4	s	
	14	56,5 – 55,4	d	
	18	12,8 – 12,1	q	
[7EN; 9 α -OH]	8	144,6 – 143,3	s	100,0
	9	78,5 – 77,8	s	
	11	29,0 – 28,7	t	
	12	36,5 – 36,0	t	
	13	45,2 – 44,9	s	
	14	52,0 – 51,5	d	
	18	12,3 – 11,8	q	
[7EN; 14 α -OH]	8	164,5 – 162,8	s	100,0
	9	47,2 – 33,7	d	
	11	21,0 – 20,7	t	
	12	31,7 – 30,7	t	
	13	48,0 – 46,7	s	
	14	83,9 – 82,3	s	
	18	17,1 – 15,3	q	
[7,9(11)EN]	8	136,3 – 135,6	s	100,0
	9	143,3 – 140,1	s	
	11	122,8 – 118,8	d	
	12	42,2 – 42,0	t	
	13	42,7 – 42,2	s	
	14	51,5 – 51,2	d	
	18	11,5 – 11,2	q	
[7EN; 9 α ,11 α EP]	8	134,8 – 134,5	s	100,0
	9	64,5 – 64,0	s	
	11	55,2 – 54,7	d	
	12	40,7 – 40,4	t	
	13	43,0 – 42,7	s	
	14	46,9 – 46,6	d	
	18	13,9 – 13,5	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[8EN]	8	138,0 – 137,6	s	100,0
	9	132,4 – 132,0	s	
	11	26,0 – 25,5	t	
	12	37,6 – 37,0	t	
	13	42,0 – 41,4	s	
	14	52,0 – 51,6	d	
	18	11,5 – 11,0	q	
[8EN]	8	156,1 – 134,5	s	100,0
	9	131,1 – 123,0	s	
	11	29,7 – 23,1	t	
	12	36,2 – 35,5	t	
	13	44,0 – 42,2	s	
	14	49,5 – 44,2	d	
	18	16,7 – 11,1	q	
[8(14)EN]	8	129,3 – 126,0	s	100,0
	9	49,7 – 35,7	d	
	11	20,0 – 19,7	t	
	12	37,5 – 36,2	t	
	13	44,0 – 42,7	s	
	14	142,3 – 139,1	s	
	18	19,3 – 18,1	q	
[9(11)EN]	8	36,7 – 36,0	d	100,0
	9	147,5 – 141,8	s	
	11	120,4 – 116,1	d	
	12	42,0 – 32,0	t	
	13	46,7 – 39,2	s	
	14	54,2 – 47,9	d	
	18	15,8 – 14,3	q	
[14EN]	8	36,7 – 30,2	d	76,9
	9	53,2 – 29,7	d	
	11	24,1 – 21,3	t	
	12	42,9 – 31,7	t	
	13	56,0 – 48,0	s	
	14	154,1 – 146,3	s	
	18	20,1 – 17,1	q	
[7EN; 11OXO]	8	135,0 – 134,3	s	100,0
	9	58,0 – 57,6	d	
	11	184,9 – 184,1	s	
	12	58,7 – 58,3	t	
	13	43,4 – 42,7	s	
	14	52,5 – 52,0	d	
	18	17,1 – 16,8	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[8EN; 11OXO]	8	139,0 – 138,4	s	100,0
	9	154,4 – 154,0	s	
	11	204,0 – 203,5	s	
	12	58,9 – 58,5	t	
	13	48,3 – 47,7	s	
	14	52,7 – 52,3	d	
	18	12,4 – 12,0	q	
[9(11)EN; 12OXO]	8	37,2 – 36,7	d	100,0
	9	172,3 – 171,8	s	
	11	120,0 – 119,5	d	
	12	205,7 – 205,1	s	
	13	51,3 – 50,6	s	
	14	53,9 – 53,3	d	
	18	15,3 – 15,0	q	
[9(11)EN; 12OXO; 14 β OH]	8	41,5 – 41,0	d	100,0
	9	165,7 – 165,2	s	
	11	120,8 – 120,1	d	
	12	202,4 – 201,8	s	
	13	61,7 – 60,9	s	
	14	83,9 – 83,3	s	
	18	15,2 – 14,9	q	
[11OXO]	8	57,5 – 35,2	d	72,0
	9	67,0 – 53,5	d	
	11	211,3 – 206,8	s	
	12	57,5 – 36,9	t	
	13	50,5 – 42,7	s	
	14	56,5 – 49,5	d	
	18	18,2 – 12,1	q	
[11OXO; 12 β -OH]	8	37,2 – 37,0	d	100,0
	9	62,9 – 62,6	d	
	11	210,6 – 210,5	s	
	12	84,5 – 84,4	d	
	13	50,5 – 50,0	s	
	14	53,0 – 52,9	d	
	18	11,6 – 11,3	q	
[11OXO; 14 β -OH]	8	43,2 – 42,5	d	100,0
	9	50,2 – 48,2	d	
	11	211,8 – 208,6	s	
	12	55,5 – 51,7	t	
	13	57,0 – 53,4	s	
	14	83,5 – 82,8	s	
	18	20,0 – 17,6	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[12OXO]	8	48,9 – 32,7	d	100,0
	9	56,5 – 51,7	d	
	11	40,5 – 36,7	t	
	12	215,3 – 208,3	s	
	13	55,5 – 54,7	s	
	14	55,7 – 48,0	d	
	18	17,7 – 15,8	q	
[12OXO; 14β-OH]	8	41,4 – 40,5	d	100,0
	9	37,5 – 33,2	d	
	11	37,4 – 37,2	t	
	12	212,0 – 211,1	s	
	13	64,8 – 64,0	s	
	14	86,5 – 85,9	s	
	18	16,7 – 16,5	q	
[11α,14βOH; 12OXO]	8	38,3 – 37,6	d	100,0
	9	44,5 – 44,0	d	
	11	73,0 – 72,1	d	
	12	213,0 – 212,4	s	
	13	61,7 – 61,2	s	
	14	84,2 – 83,8	s	
	18	16,9 – 16,5	q	
[14α-OH; 18OXO; 18-OR]	8	37,9 – 37,6	d	100,0
	9	44,7 – 39,5	d	
	11	23,0 – 22,7	t	
	12	36,2 – 35,4	t	
	13	60,5 – 60,0	s	
	14	83,5 – 83,0	s	
	18	177,6 – 177,3	s	
[17β:18β-Ciclo]	8	36,7 – 36,4	d	100,0
	9	50,2 – 49,9	d	
	11	26,1 – 25,5	t	
	12	30,2 – 29,2	t	
	13	43,0 – 35,2	s	
	14	50,9 – 50,2	d	
	18	19,1 – 14,6	t	
[12β-OR; 12α-R]	8	32,5 – 29,2	d	100,0
	9	41,9 – 40,0	d	
	11	36,4 – 30,0	t	
	12	101,9 – 82,0	s	
	13	51,2 – 47,7	s	
	14	45,9 – 44,0	d	
	18	12,6 – 9,8	q	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento	
[7,9(11)EN; 14 α -Me]	8	143,2	– 142,9	s	100,0
	9	143,3	– 143,0	s	
	11	117,2	– 117,5	d	
	12	38,0	– 37,7	t	
	13	43,9	– 43,5	s	
	14	50,5	– 50,2	s	
	18	15,8	– 15,5	q	
	Me	25,7	– 25,5	q	
[8EN; 14 α -Me]	8	133,6	– 132,3	s	100,0
	9	135,6	– 134,6	s	
	11	21,9	– 21,7	t	
	12	31,1	– 31,0	t	
	13	44,6	– 44,4	s	
	14	49,9	– 49,7	s	
	18	15,8	– 15,6	q	
	Me	24,5	– 24,2	q	
[11 β -OH; 18-Me]	8	34,5	– 34,2	d	100,0
	9	54,0	– 53,9	d	
	11	66,5	– 66,3	d	
	12	39,5	– 34,0	t	
	13	50,2	– 43,7	s	
	14	51,9	– 51,6	d	
	18	20,2	– 19,5	t	
	Me	11,3	– 8,3	q

[18-Me]	8	42,7	– 33,7	d	100,0
	9	54,5	– 43,5	d	
	11	147,8	– 147,0	s	
	12	40,9	– 40,2	t	
	13	50,9	– 50,2	s	
	14	52,9	– 52,0	d	
	18	20,0	– 19,0	t	
	Me	9,3	– 9,0	q	
	CH ₂	109,1	– 108,3	t	

Tabela XV. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento

[9OXO; 11-OH]	8	40,2 – 35,0	d	100,0
	9	215,1 – 209,3	s	
	11	61,2 – 58,9	t	
	12	41,2 – 35,4	t	
	13	46,4 – 45,4	s	
	14	47,9 – 43,5	d	
	18	17,7 – 15,8	q	
[9OXO; 11-OH; 11-OR]	8	41,5 – 41,2	d	100,0
	9	215,3 – 214,8	s	
	11	94,5 – 94,1	d	
	12	45,8 – 45,6	t	
	13	45,0 – 44,7	s	
	14	45,2 – 45,0	d	
	18	12,9 – 12,6	q	
[7EN; 9OXO; 11-OH]	8	137,7 – 137,4	s	100,0
	9	206,2 – 205,9	s	
	11	59,4 – 59,0	t	
	12	42,1 – 42,0	t	
	13	50,6 – 50,2	s	
	14	43,6 – 43,3	d	
	18	18,0 – 17,7	q

[14 α -OR; 14,15-SECO]	8	34,0 – 33,5	d	100,0
	9	45,7 – 45,5	d	
	11	19,5 – 19,3	t	
	12	29,2 – 28,5	t	
	13	41,7 – 41,2	s	
	14	76,5 – 75,6	d	
	18	23,2 – 20,5	q	

Tabela XVI. Faixas de deslocamento químico características de esteróides – Anel D

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento
<div>Anel D</div> <div> </div>					
[14 α -H]	13	54,9 – 39,7	s	100,0	
	14	56,4 – 46,4	d		
	15	27,7 – 24,7	t		
	16	21,5 – 19,5	t		
	17	42,0 – 31,8	t		
[15 α -OH]	13	42,0 – 41,5	s	100,0	
	14	62,2 – 61,6	d		
	15	75,8 – 75,1	d		
	16	33,0 – 32,4	t		
	17	38,5 – 38,0	t		
[16 β -OH]	13	40,2 – 39,4	s	100,0	
	14	54,2 – 53,7	d		
	15	37,5 – 37,2	t		
	16	71,9 – 71,3	d		
	17	52,2 – 51,5	t		
[17 α -OH]	13	46,2 – 42,2	s	100,0	
	14	48,9 – 46,0	d		
	15	24,8 – 21,1	t		
	16	32,5 – 29,2	t		
	17	83,4 – 79,6	d		
[17 β -OH]	13	45,0 – 42,5	s	100,0	
	14	52,9 – 44,0	d		
	15	24,2 – 22,5	t		
	16	31,3 – 26,7	t		
	17	84,5 – 79,8	d		
[17 α -OR]	13	45,5 – 44,5	s	100,0	
	14	50,3 – 50,0	d		
	15	24,9 – 24,6	t		
	16	30,8 – 30,2	t		
	17	82,2 – 81,5	d		
[17 β -OR]	13	44,2 – 42,4	s	93,3	
	14	51,7 – 48,5	d		
	15	26,0 – 23,0	t		
	16	28,2 – 27,2	t		
	17	90,8 – 81,4	d		

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[14 α -OH; 17 β -OR]	13	47,5 – 46,9	s	100,0
	14	81,7 – 81,2	s	
	15	27,5 – 27,0	t	
	16	20,8 – 20,1	t	
	17	81,9 – 81,3	d	
[17 α ,17 β -OR]	13	46,2 – 45,7	s	100,0
	14	50,7 – 49,4	d	
	15	22,7 – 22,3	t	
	16	34,2 – 34,0	t	
	17	119,5 – 119,1	s	
[15EN; 17OXO]	13	51,0 – 50,6	s	100,0
	14	57,5 – 57,2	d	
	15	158,8 – 158,5	d	
	16	132,3 – 131,8	d	
	17	213,3 – 213,0	s	
[15OXO]	13	39,5 – 39,0	s	100,0
	14	63,7 – 63,1	d	
	15	216,4 – 215,8	s	
	16	35,2 – 34,6	t	
	17	35,8 – 35,0	t	
[16OXO]	13	39,2 – 38,7	s	100,0
	14	51,9 – 50,7	d	
	15	39,7 – 39,2	t	
	16	218,3 – 217,1	s	
	17	55,9 – 55,5	t	
[16OXO; 17 β -OH]	13	42,7 – 42,3	s	100,0
	14	45,6 – 45,1	d	
	15	36,5 – 36,0	t	
	16	217,8 – 217,1	s	
	17	86,8 – 86,3	d	
[17OXO]	13	54,2 – 46,9	s	100,0
	14	53,2 – 45,7	d	
	15	25,7 – 20,0	t	
	16	37,2 – 34,5	t	
	17	222,8 – 214,5	s	
[14 α -OH; 17OXO]	13	54,5 – 52,5	s	100,0
	14	82,1 – 80,6	s	
	15	33,7 – 30,2	t	
	16	30,5 – 24,5	t	
	17	219,3 – 218,0	s	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[16 β -OR; 17OXO]	13	47,5 – 47,0	s	100,0
	14	46,4 – 46,1	d	
	15	31,7 – 29,5	t	
	16	75,3 – 75,0	d	
	17	220,8 – 215,3	s	
[16-CH ₂ OR; 17 α -OR]	13	44,5 – 44,2	s	100,0
	14	51,2 – 49,9	d	
	15	29,7 – 29,2	t	
	16	45,7 – 45,5	d	
	17	83,6 – 83,4	d	
	CH ₂	66,8 – 66,4	t	
[16-CH ₂ OR; 17 β -OR]	13	44,2 – 43,0	s	100,0
	14	49,9 – 48,5	d	
	15	29,6 – 27,3	t	
	16	40,2 – 37,5	d	
	17	83,8 – 81,5	d	
[17 α -Me; 17 β -OH]	13	46,0 – 43,5	s	100,0
	14	53,0 – 44,5	d	
	15	23,3 – 22,7	t	
	16	39,0 – 31,3	t	
	17	81,6 – 79,9	s	
	Me	26,2 – 25,6	q	
[17OXO; 18-Me]	13	54,2 – 50,2	s	100,0
	14	51,7 – 49,9	d	
	15	20,8 – 20,1	t	
	16	36,4 – 35,0	t	
	17	218,1 – 214,5	s	
	18	19,5 – 18,8	t	
	Me	8,3 – 7,0	q

[---]	13	46,7 – 41,2	s	100,0
	14	57,7 – 54,4	d	
	15	24,8 – 24,0	t	
	16	28,5 – 27,7	t	
	17	53,7 – 52,0	d	
	20	23,2 – 23,0	t	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento
[20 β -OR]	13	42,4	– 42,1	s	100,0
	14	56,3	– 56,0	d	
	15	24,4	– 24,0	t	
	16	25,6	– 25,2	t	
	17	55,4	– 55,0	d	
	20	81,4	– 72,8	d	
[14 β ,17 β ,20-OH]	13	57,2	– 50,7	s	100,0
	14	87,5	– 86,5	s	
	15	33,4	– 32,2	t	
	16	33,4	– 31,7	t	
	17	88,0	– 87,8	s	
	20	71,6	– 69,4	d	
[16 β -OR; 17(20)EN]	13	43,3	– 42,8	s	100,0
	14	51,0	– 50,6	d	
	15	33,5	– 33,0	t	
	16	73,1	– 72,4	d	
	17	148,9	– 148,1	s	
	20	118,5	– 118,0	d	
	21	13,5	– 13,0	q	
[16EN; 20OXO]	13	46,5	– 45,5	s	100,0
	14	56,2	– 55,5	d	
	15	32,2	– 31,6	t	
	16	155,8	– 153,1	d	
	17	143,5	– 141,8	s	
	20	209,1	– 195,6	s	
[16OXO]	13	42,2	– 41,7	s	100,0
	14	50,7	– 50,2	d	
	15	38,5	– 37,7	t	
	16	219,5	– 218,1	s	
	17	65,4	– 63,4	d	
	20	18,1	– 17,6	t	
	21	13,6	– 13,3	q	
[20OXO]	13	58,0	– 42,7	s	80,0
	14	57,0	– 49,4	d	
	15	24,7	– 23,7	t	
	16	23,6	– 21,7	t	
	17	64,3	– 50,7	d	
	20	210,0	– 195,5	s	
[14 α -OH; 20OXO]	13	48,7	– 48,0	s	91,7
	14	88,1	– 83,6	s	
	15	32,5	– 30,2	t	
	16	22,2	– 20,1	t	
	17	61,0	– 59,2	d	
	20	213,6	– 208,6	s	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[14β-OH; 20OXO]	13	54,4 – 53,9	s	100,0
	14	86,6 – 86,1	s	
	15	35,7 – 35,2	t	
	16	24,9 – 24,0	t	
	17	58,0 – 57,3	d	
	20	217,7 – 217,2	s	
[14β-OH; 20OXO; 20α]	13	56,4 – 55,0	s	100,0
	14	86,9 – 86,3	s	
	15	33,6 – 33,3	t	
	16	21,5 – 20,7	t	
	17	60,0 – 59,7	d	
	20	210,0 – 209,3	s	
[17α-OH; 20OXO]	13	49,0 – 45,5	s	100,0
	14	51,5 – 44,7	d	
	15	24,1 – 22,6	t	
	16	34,5 – 31,8	t	
	17	89,6 – 88,1	s	
	20	213,1 – 204,6	s	
[14β,17β-OH; 20OXO]	13	58,7 – 51,5	s	100,0
	14	89,1 – 88,5	s	
	15	33,6 – 33,4	t	
	16	32,0 – 27,2	t	
	17	93,3 – 91,0	s	
	20	208,1 – 207,6	s	
[16αMe; 17α-OH;20OXO]	13	47,9 – 47,1	s	100,0
	14	43,6 – 43,1	d	
	15	32,4 – 32,0	t	
	16	35,3 – 34,7	d	
	17	90,5 – 90,0	s	
	20	211,4 – 211,0	s	
	Me	17,9 – 17,5	q	
[17β-OH; 18-Me; 20IN]	13	50,9 – 50,2	s	100,0
	14	52,9 – 52,0	d	
	15	21,9 – 21,5	t	
	16	40,0 – 39,7	t	
	17	81,8 – 80,9	s	
	18	20,0 – 19,0	t	
	20	88,0 – 87,6	s	
	21	74,1 – 73,8	d	
	Me	9,3 – 9,0	q	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[17 α -OH; 20OXO; 20-OR 21NOR]	13	47,9	–	46,2	s	100,0
	14	50,2	–	50,0	d	
	15	24,1	–	23,5	t	
	16	32,9	–	32,5	t	
	17	85,6	–	84,9	s	
	20	175,0	–	174,0	s	
[17 α .20-OR; 20OXO; 21NOR]	13	50,5	–	46,0	s	100,0
	14	51,2	–	41,4	d	
	15	24,6	–	22,7	t	
	16	31,7	–	29,7	t	
	17	92,4	–	90,5	s	
	20	170,1	–	167,1	s	
[14,15SECO;14 α ,20 β -OR; 15OXO; 15:20OXI; 16EN]	13	41,7	–	41,2	s	100,0
	14	76,5	–	75,6	d	
	15	169,1	–	168,8	s	
	16	120,4	–	118,5	d	
	17	174,8	–	172,1	d	
	20	110,1	–	109,9	s	
	21	24,1	–	23,7	q

[17 β :18 β -Ciclo; 20OXO]	13	43,0	–	38,4	s	100,0
	14	51,9	–	50,2	d	
	15	24,7	–	24,0	t	
	16	29,3	–	29,0	t	
	17	46,9	–	40,9	s	
	18	19,8	–	19,1	t	
	20	207,8	–	207,5	s

CL: Cadeia Lateral						
[---]	13	50,5	–	42,0	s	100,0
	14	59,9	–	41,0	d	
	15	27,7	–	22,3	t	
	16	40,0	–	21,6	t	
	17	58,0	–	47,0	d	
	20	41,7	–	33,5	d	
	22	45,2	–	25,7	t	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[14 α -OH]	13	47,2	– 46,2	100,0
	14	83,9	– 82,3	
	15	31,3	– 31,0	
	16	27,2	– 25,7	
	17	51,0	– 47,2	
	20	42,2	– 35,2	
[20 β -OR]	13	43,7	– 43,2	100,0
	14	56,7	– 56,5	
	15	24,3	– 23,3	
	16	22,5	– 21,6	
	17	55,5	– 53,4	
	20	77,4	– 75,0	
[14 α ,15 α -OH]	13	49,7	– 47,7	100,0
	14	85,1	– 83,3	
	15	70,0	– 62,0	
	16	35,7	– 34,7	
	17	51,7	– 50,7	
	20	40,0	– 35,0	
[14 α -OH; 15 α -OR]	13	46,9	– 46,6	100,0
	14	84,5	– 84,1	
	15	76,5	– 76,1	
	16	34,2	– 33,6	
	17	50,8	– 50,5	
	20	34,8	– 34,1	
[14 β ,17 β -OH; 20 α]	13	51,0	– 50,2	100,0
	14	87,5	– 87,3	
	15	31,6	– 30,0	
	16	37,0	– 35,7	
	17	88,8	– 88,4	
	20	41,7	– 41,5	
[14 α ,20 β -OH]	13	48,9	– 47,0	100,0
	14	84,0	– 83,0	
	15	32,4	– 30,0	
	16	22,0	– 20,2	
	17	50,2	– 49,0	
	20	76,9	– 76,1	
[15 α ,16 β -OH]	13	45,5	– 44,2	100,0
	14	64,5	– 59,4	
	15	85,0	– 79,3	
	16	83,0	– 82,3	
	17	61,5	– 59,5	
	20	32,5	– 30,5	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento	
[15α-OR; 16β-OH]	13	44,2	– 43,2	s	100,0
	14	59,5	– 53,7	d	
	15	91,5	– 83,9	d	
	16	80,0	– 78,5	d	
	17	60,0	– 58,7	d	
	20	30,7	– 29,0	d	
[14α,17β,20β-OH; 20α]	13	55,5	– 54,5	s	100,0
	14	83,9	– 82,3	s	
	15	31,1	– 30,1	t	
	16	37,9	– 37,4	t	
	17	88,6	– 87,8	s	
	20	80,0	– 78,3	s	
[16β,22OXI; 22-OR]	13	55,5	– 39,2	s	100,0
	14	56,9	– 48,0	d	
	15	34,9	– 29,7	t	
	16	83,0	– 78,6	d	
	17	65,3	– 50,9	d	
	20	43,0	– 37,4	d	
	22	120,4	– 108,5	s	
[18,20OXI]	13	59,3	– 59,0	s	100,0
	14	84,2	– 83,7	d	
	15	35,3	– 35,0	t	
	16	25,4	– 25,0	t	
	17	57,5	– 55,5	d	
	18	72,8	– 71,7	t	
	20	88,6	– 87,0	s	
[16β,17β-EP]	13	42,5	– 42,1	s	100,0
	14	56,3	– 56,0	d	
	15	28,1	– 27,7	t	
	16	63,7	– 63,4	d	
	17	75,2	– 75,0	s	
	20	37,4	– 37,0	d	
[20,22-EP]	13	43,2	– 42,4	s	100,0
	14	56,7	– 56,2	d	
	15	24,7	– 23,8	t	
	16	26,8	– 26,2	t	
	17	56,7	– 50,5	d	
	20	61,0	– 60,0	s	
[14EN; 17β,20β-OH; 20α]	13	56,0	– 55,5	s	100,0
	14	151,1	– 150,6	s	
	15	115,3	– 114,8	d	
	16	35,7	– 35,3	t	
	17	87,3	– 86,9	s	
	20	77,9	– 77,0	s	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento
[14 α -OH; 15 α -OR; 16EN]	13	52,3	–	52,0	100,0
	14	82,3	–	81,3	
	15	83,6	–	83,4	
	16	121,0	–	120,4	
	17	162,3	–	160,8	
	20	36,0	–	35,2	
[20(22)EN]	13	43,6	–	43,2	100,0
	14	56,5	–	56,0	
	15	24,5	–	24,2	
	16	26,0	–	24,6	
	17	59,0	–	58,7	
	20	134,1	–	133,7	
[20(22)EN]	13	45,0	–	44,2	100,0
	14	57,0	–	49,5	
	15	27,0	–	24,2	
	16	26,7	–	24,8	
	17	60,5	–	49,5	
	20	173,1	–	159,8	
[14 β -OH; 20EN]	13	61,9	–	47,5	100,0
	14	85,3	–	83,0	
	15	33,2	–	31,0	
	16	28,8	–	27,5	
	17	51,5	–	40,4	
	20	124,0	–	120,9	
	21	150,5	–	148,1	
[14 β -OH; 20EN]	13	64,8	–	49,2	75,0
	14	86,5	–	81,0	
	15	34,7	–	28,5	
	16	28,8	–	24,7	
	17	52,2	–	39,7	
	20	178,6	–	171,1	
[14 β -OH; 20EN; 20 α]	13	50,5	–	48,7	70,0
	14	87,5	–	84,5	
	15	33,0	–	31,5	
	16	27,2	–	24,3	
	17	51,4	–	48,2	
	20	175,8	–	171,0	
[17 α -OH; 20EN]	13	49,2	–	48,5	100,0
	14	50,9	–	45,0	
	15	26,6	–	23,7	
	16	38,0	–	37,0	
	17	84,3	–	81,6	
	20	175,5	–	175,0	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[14β-OH; 15β-OR; 20EN]	13	48,3	–	47,8	s	100,0
	14	82,0	–	81,6	s	
	15	75,9	–	75,2	d	
	16	34,3	–	33,8	t	
	17	48,9	–	48,4	d	
	20	173,0	–	172,3	s	
[14β,15β-EP; 20EN]	13	45,2	–	45,0	s	100,0
	14	74,7	–	74,4	s	
	15	59,9	–	59,7	d	
	16	32,5	–	32,2	t	
	17	47,7	–	47,5	d	
	20	122,3	–	122,0	s	
[14β-OH; 16β-OR; 20EN]	13	50,7	–	50,5	s	100,0
	14	84,0	–	83,8	s	
	15	39,5	–	39,2	t	
	16	75,0	–	74,6	d	
	17	56,7	–	56,5	d	
	20	171,7	–	171,4	s	
[22EN]	13	50,5	–	41,7	s	96,6
	14	57,4	–	43,5	d	
	15	30,1	–	22,3	t	
	16	29,5	–	22,7	t	
	17	57,7	–	51,4	d	
	20	41,7	–	34,2	d	
	22	142,1	–	132,0	d	
[8(14),20EN]	13	44,2	–	43,8	s	100,0
	14	139,8	–	139,1	s	
	15	26,2	–	25,7	t	
	16	27,7	–	27,0	t	
	17	52,0	–	51,6	d	
	20	171,6	–	171,2	s	
[14,20EN]	13	48,7	–	48,4	s	100,0
	14	154,1	–	152,8	s	
	15	118,0	–	116,5	d	
	16	33,9	–	32,9	t	
	17	62,2	–	52,5	d	
	20	171,3	–	159,6	s	
[14β-OH; 16,20EN]	13	57,0	–	52,5	s	100,0
	14	85,6	–	82,8	s	
	15	42,0	–	38,7	t	
	16	135,8	–	133,8	d	
	17	161,1	–	141,3	s	
	20	172,8	–	158,3	s	

Tabela XVI. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento
[16OXO]	13	43,2 – 42,7 s	100,0
	14	45,0 – 44,6 d	
	15	38,3 – 38,0 t	
	16	216,9 – 216,6 s	
	17	63,1 – 62,8 d	
	20	36,8 – 36,5 d	
[14 α -OH; 18,20OXI; 18OXO]	13	60,5 – 60,0 s	100,0
	14	83,5 – 83,0 s	
	15	34,9 – 34,5 t	
	16	25,7 – 24,8 t	
	17	57,7 – 55,0 d	
	18	177,6 – 177,3 s	
	20	83,9 – 83,4 s	
[14 α -Me]	13	44,5 – 43,7 s	100,0
	14	50,4 – 37,0 s	
	15	34,0 – 30,7 t	
	16	28,2 – 27,6 t	
	17	51,0 – 50,4 d	
	20	36,5 – 35,9 d	
	Me	25,6 – 18,2 q	

Tabela XVII. Faixas de deslocamento químico características de esteróides – Cadeia Lateral

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
--------------	------	--------------------------------	--	---------------------

Cadeia Lateral

[---]	20	19,8 – 18,5	q	
	21	37,4 – 34,5	d	
	22	37,2 – 35,9	t	
	23	25,6 – 23,6	t	
	24	40,7 – 39,0	t	
	25	29,6 – 27,6	d	
	26	23,8 – 22,3	q	
	27	23,5 – 22,2	q	
[21-OR]	20	42,5 – 40,0	d	100,0
	21	70,2 – 68,9	t	
	22	32,0 – 29,7	t	
	23	26,3 – 23,9	t	
	24	41,0 – 39,7	t	
	25	30,9 – 28,6	d	
	26	23,6 – 22,7	q	
	27	23,6 – 22,7	q	
[22-OH]	20	42,5 – 40,2	d	100,0
	21	12,5 – 11,6	q	
	22	74,0 – 73,7	d	
	23	33,2 – 27,5	t	
	24	36,0 – 35,7	t	
	25	27,9 – 27,6	d	
	26	22,9 – 22,5	q	
	27	22,7 – 22,4	q	
[24-OH]	20	38,3 – 36,5	d	100,0
	21	20,0 – 18,6	q	
	22	34,7 – 32,7	t	
	23	33,0 – 30,9	t	
	24	78,6 – 77,3	d	
	25	35,2 – 34,0	d	
	26	20,0 – 19,3	q	
	27	18,0 – 17,4	q	
[26-OH]	20	32,5 – 30,5	d	100,0
	21	18,6 – 18,2	q	
	22	37,4 – 36,9	t	
	23	24,8 – 24,3	t	
	24	35,0 – 34,5	t	
	25	37,0 – 36,7	d	
	26	68,5 – 67,5	t	
	27	17,5 – 17,2	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[26-OR]	20	36,9 – 29,0	d	100,0
	21	19,1 – 17,6	q	
	22	37,0 – 35,4	t	
	23	24,2 – 23,3	t	
	24	34,5 – 33,7	t	
	25	34,2 – 32,4	d	
	26	73,8 – 69,4	t	
	27	17,0 – 16,8	q	
[20β,22α-OH]	20	77,4 – 76,1	s	77,8
	21	21,2 – 19,2	q	
	22	76,8 – 76,4	d	
	23	30,2 – 29,1	t	
	24	37,2 – 36,0	t	
	25	28,2 – 27,7	d	
	26	23,2 – 21,8	q	
	27	22,5 – 21,2	q	
[20β,22β-OH]	20	76,8 – 76,3	s	60,0
	21	21,4 – 21,0	q	
	22	76,7 – 76,2	d	
	23	30,3 – 30,0	t	
	24	37,2 – 36,7	t	
	25	28,4 – 28,0	d	
	26	23,2 – 22,2	q	
	27	22,5 – 22,2	q	
[20α,22β-OH; 21β]	20	78,5 – 78,1	s	100,0
	21	21,7 – 21,5	q	
	22	76,5 – 76,2	d	
	23	29,8 – 29,5	t	
	24	36,2 – 35,9	t	
	25	28,5 – 28,1	d	
	26	22,7 – 22,5	q	
	27	22,7 – 22,5	q	
[22β-OR; 25-OH]	20	42,2 – 39,0	s	100,0
	21	13,6 – 13,3	q	
	22	76,4 – 73,8	d	
	23	25,3 – 22,3	t	
	24	42,7 – 40,5	t	
	25	69,5 – 68,0	s	
	26	29,8 – 29,5	q	
	27	29,7 – 29,0	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[20β,22α,25-OH]	20	77,0 – 76,6	s	100,0
	21	21,9 – 21,5	q	
	22	77,7 – 77,3	d	
	23	27,7 – 27,3	t	
	24	42,6 – 42,2	t	
	25	69,7 – 69,4	s	
	26	30,1 – 29,8	q	
	27	30,3 – 30,0	q	
[20β-OH; 22α,25-OR]	20	77,6 – 77,0	s	100,0
	21	21,3 – 21,0	q	
	22	79,6 – 79,3	d	
	23	24,8 – 24,5	t	
	24	40,2 – 37,5	t	
	25	81,6 – 70,5	s	
	26	30,2 – 26,2	q	
	27	29,2 – 25,8	q	
[20β,25-OH; 22β-OR]	20	76,8 – 76,0	s	100,0
	21	21,2 – 21,0	q	
	22	80,1 – 77,4	d	
	23	29,7 – 26,0	t	
	24	42,4 – 41,5	t	
	25	69,5 – 69,1	s	
	26	30,0 – 29,7	q	
	27	29,2 – 27,2	q	
[16β:22-OXI]	20	38,0 – 37,7	d	100,0
	21	19,2 – 19,0	q	
	22	90,6 – 90,3	d	
	23	30,0 – 29,6	t	
	24	31,3 – 30,0	t	
	25	36,0 – 28,2	d	
	26	22,6 – 16,7	q	
[16β:22,22α:25β-OXI; 26-OH]	20	38,5 – 37,4	d	100,0
	21	14,8 – 14,1	q	
	22	120,1 – 119,5	s	
	23	32,7 – 26,7	t	
	24	34,0 – 32,9	t	
	25	87,1 – 84,0	s	
	26	69,5 – 64,9	t	
[16β:22,22α:26-OXI]	20	43,0 – 40,7	d	100,0
	21	14,5 – 13,1	q	
	22	109,9 – 108,5	s	
	23	31,0 – 25,7	t	
	24	29,2 – 25,7	t	
	25	30,6 – 26,0	d	
	26	67,1 – 65,0	t	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[16β:22,22α:26-OXI; 23α-Br]	20	38,9	–	38,0	d	100,0
	21	14,6	–	12,6	q	
	22	109,0	–	108,8	s	
	23	49,9	–	49,5	d	
	24	40,6	–	40,2	t	
	25	33,7	–	33,4	d	
	26	65,9	–	65,4	t	
[16β:22,22α:26-OXI; 24-OR]	20	42,5	–	41,0	d	100,0
	21	15,3	–	14,3	q	
	22	112,1	–	111,0	s	
	23	42,2	–	40,2	t	
	24	82,3	–	80,3	d	
	25	40,0	–	37,6	d	
	26	66,3	–	64,8	t	
	27	15,0	–	13,2	q	
[16β:22,22α:26-OXI; 25-OH]	20	41,5	–	41,0	d	100,0
	21	14,6	–	14,3	q	
	22	109,1	–	108,7	s	
	23	29,6	–	24,7	t	
	24	34,9	–	32,7	t	
	25	67,0	–	66,5	s	
	26	69,0	–	68,8	t	
	27	27,0	–	24,0	q	
[22,25-OXI]	20	39,4	–	38,1	d	100,0
	21	13,3	–	12,2	q	
	22	80,8	–	80,0	d	
	23	26,0	–	24,7	t	
	24	40,0	–	38,0	t	
	25	80,1	–	78,9	s	
	26	29,2	–	27,8	q	
	27	29,2	–	27,8	q	
[20,22-EP]	20	61,0	–	60,0	s	100,0
	21	22,7	–	17,2	q	
	22	66,5	–	59,7	d	
	23	24,7	–	21,8	t	
	24	36,4	–	35,5	t	
	25	28,2	–	27,9	d	
	26	22,9	–	22,5	q	
	27	22,9	–	22,5	q	
[24OXO; 24-OR 25,26,27NOR]	20	35,7	–	35,0	d	100,0
	21	18,3	–	17,2	q	
	22	31,3	–	31,0	t	
	23	31,3	–	31,0	t	
	24	175,3	–	174,6	s	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[20(22)EN]	20	137,8 – 133,8	s	100,0
	21	19,1 – 17,7	q	
	22	126,0 – 125,3	d	
	23	26,0 – 24,8	t	
	24	39,2 – 38,9	t	
	25	28,0 – 27,7	d	
	26	22,9 – 22,4	q	
	27	23,0 – 22,5	q	
[22EN]	20	40,4 – 39,2	d	100,0
	21	21,3 – 20,8	q	
	22	138,1 – 137,6	d	
	23	127,0 – 126,3	d	
	24	42,0 – 41,7	t	
	25	28,8 – 28,5	d	
	26	23,1 – 22,2	q	
	27	22,7 – 22,2	q	
[22EN; 20 β -OH]	20	76,1 – 74,7	s	100,0
	21	30,3 – 29,0	q	
	22	139,8 – 138,0	d	
	23	127,0 – 125,1	d	
	24	42,7 – 41,3	t	
	25	29,4 – 28,3	d	
	26	23,1 – 22,2	q	
	27	23,1 – 22,2	q	
[22EN; 24 α -OH]	20	41,4 – 40,2	d	100,0
	21	21,5 – 20,8	q	
	22	139,8 – 138,8	d	
	23	131,0 – 129,6	d	
	24	79,4 – 78,0	d	
	25	36,4 – 34,8	d	
	26	19,1 – 18,2	q	
	27	19,1 – 18,2	q	
[24EN; 26-OH]	20	35,2 – 34,7	d	100,0
	21	18,7 – 18,0	q	
	22	37,5 – 36,5	t	
	23	25,0 – 24,3	t	
	24	125,5 – 124,9	d	
	25	135,8 – 134,8	s	
	26	68,3 – 67,5	t	
	27	14,0 – 13,3	q	
[20(22)EN; 23OXO; 23OR 24,25,26,27-NOR]	20	160,1 – 159,6	s	100,0
	21	20,7 – 20,3	q	
	22	116,3 – 115,9	d	
	23	167,0 – 166,8	s	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[16 β :22,22 α :27-OXI; 25EN]	20	41,6 – 41,4	d	100,0
	21	14,7 – 14,4	q	
	22	109,2 – 108,8	s	
	23	28,6 – 28,3	t	
	24	32,8 – 32,5	t	
	25	143,5 – 143,2	s	
	26	108,5 – 108,2	t	
	27	65,0 – 64,7	t	
[23-Me]	20	33,7 – 33,5	d	100,0
	21	19,5 – 19,0	q	
	22	45,2 – 44,0	t	
	23	25,4 – 25,1	d	
	24	48,4 – 45,7	t	
	25	27,7 – 27,2	d	
	26	23,0 – 21,3	q	
	27	22,7 – 21,6	q	
[24 α -OH; 27NOR]	20	35,6 – 35,4	d	100,0
	21	18,9 – 18,6	q	
	22	32,7 – 32,5	t	
	23	34,6 – 34,3	t	
	24	73,1 – 72,8	d	
	25	30,7 – 30,5	t	
	26	10,3 – 10,1	q	
[24-OH; 25,26,27NOR]	20	36,0 – 35,6	d	100,0
	21	18,6 – 18,0	q	
	22	32,3 – 32,0	t	
	23	29,7 – 29,4	t	
	24	63,7 – 62,9	t

[---]	20	36,9 – 35,9	d	100,0
	21	19,1 – 18,7	q	
	22	34,0 – 33,5	t	
	23	30,8 – 30,1	t	
	24	41,7 – 39,0	d	
	25	32,5 – 31,3	d	
	26	20,6 – 17,6	q	
	27	20,5 – 17,6	q	
	28	15,5 – 14,3	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[22 α -OH]	20	43,0 – 42,0	d	100,0
	21	13,1 – 12,2	q	
	22	72,4 – 71,3	d	
	23	35,8 – 33,3	t	
	24	36,7 – 34,2	d	
	25	31,5 – 28,1	d	
	26	17,6 – 15,7	q	
	27	21,9 – 20,6	q	
	28	16,2 – 15,0	q	
[28 β]	20	36,4 – 35,9	d	100,0
	21	18,7 – 18,5	q	
	22	33,9 – 33,7	t	
	23	30,6 – 30,3	t	
	24	39,0 – 38,6	d	
	25	32,6 – 32,2	d	
	26	20,2 – 18,2	q	
	27	20,2 – 18,2	q	
	28	15,5 – 15,2	q	
[22,24 β ,25-OH]	20	44,5 – 44,2	d	100,0
	21	12,7 – 12,5	q	
	22	71,0 – 70,7	d	
	23	38,5 – 36,7	t	
	24	76,2 – 75,8	s	
	25	77,5 – 77,1	s	
	26	25,3 – 25,0	q	
	27	25,3 – 25,0	q	
	28	22,7 – 22,5	q	
[22,24 α ,25-OH; 28 β]	20	43,4 – 43,1	d	100,0
	21	13,1 – 12,8	q	
	22	71,5 – 71,2	d	
	23	34,1 – 33,6	t	
	24	76,3 – 75,9	s	
	25	78,5 – 78,1	s	
	26	25,0 – 24,6	q	
	27	25,0 – 24,6	q	
	28	22,4 – 22,0	q	
[22,24,25-OR]	20	50,5 – 36,5	d	100,0
	21	14,5 – 10,3	q	
	22	71,6 – 69,0	d	
	23	35,0 – 30,1	t	
	24	83,5 – 82,0	s	
	25	83,1 – 81,0	s	
	26	20,1 – 19,7	q	
	27	20,6 – 20,1	q	
	28	25,2 – 24,7	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[17 α :22OXI; 24OH;25OR]	17	90,6 – 89,3	s	100,0
	20	49,9 – 48,9	d	
	21	12,0 – 11,7	q	
	22	78,7 – 78,4	d	
	23	28,8 – 28,5	t	
	24	76,0 – 75,8	s	
	25	93,5 – 93,2	s	
	26	20,8 – 20,5	q	
	27	22,2 – 21,5	q	
	28	23,3 – 23,0	q	
[17 α :22OXI; 24OR;25OH]	17	90,0 – 88,5	s	100,0
	20	49,0 – 48,7	d	
	21	11,7 – 11,5	q	
	22	78,9 – 78,6	d	
	23	28,4 – 28,1	t	
	24	94,3 – 94,1	s	
	25	73,6 – 73,3	s	
	26	25,3 – 24,2	q	
	27	25,3 – 24,2	q	
	28	19,2 – 18,7	q	
[22:25OXI; 24 β -OR]	20	39,7 – 37,7	d	100,0
	21	12,5 – 12,2	q	
	22	77,8 – 77,0	d	
	23	38,9 – 37,7	t	
	24	89,1 – 84,1	s	
	25	84,1 – 81,0	s	
	26	22,8 – 22,2	q	
	27	21,7 – 19,2	q	
	28	27,7 – 25,8	q	
[22EN]	20	40,7 – 39,2	d	100,0
	21	21,5 – 19,6	q	
	22	136,3 – 132,0	d	
	23	135,6 – 130,0	d	
	24	43,9 – 42,5	d	
	25	34,0 – 31,1	d	
	26	20,0 – 17,6	q	
	27	21,2 – 17,2	q	
	28	21,2 – 17,2	q	
[22EN; 28 β]	20	40,4 – 40,1	d	66,7
	21	21,2 – 20,8	q	
	22	136,1 – 135,5	d	
	23	132,3 – 131,8	d	
	24	43,0 – 42,7	d	
	25	33,3 – 33,0	d	
	26	20,5 – 19,7	q	
	27	20,5 – 20,0	q	
	28	18,1 – 17,6	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[22EN; 25,28-OH]	20	40,2	–	40,0	d	100,0
	21	20,7	–	20,5	q	
	22	142,1	–	140,0	d	
	23	126,3	–	124,6	d	
	24	55,6	–	55,3	d	
	25	73,0	–	72,6	s	
	26	26,3	–	26,0	q	
	27	30,1	–	29,3	q	
	28	64,2	–	64,0	t	
[22OR; 23:26OXI; 24EN 26OXO]	20	35,3	–	35,0	d	100,0
	21	12,1	–	11,8	q	
	22	69,0	–	68,6	d	
	23	82,5	–	82,2	d	
	24	157,1	–	156,1	s	
	25	124,5	–	123,6	s	
	26	175,0	–	174,6	s	
	27	8,3	–	8,0	q	
	28	12,0	–	11,8	q	
[24(28)EN; 26-OH]	20	36,5	–	35,7	d	100,0
	21	19,5	–	18,3	q	
	22	36,0	–	34,7	t	
	23	33,2	–	32,0	t	
	24	154,1	–	152,9	s	
	25	43,8	–	42,9	d	
	26	68,1	–	67,2	t	
	27	18,2	–	16,8	q	
	28	109,6	–	108,9	t	
[24-Me; 24(28)EN]	20	36,5	–	35,4	d	100,0
	21	18,7	–	18,3	q	
	22	36,0	–	34,4	t	
	23	31,2	–	29,7	t	
	24	156,8	–	156,3	s	
	25	34,0	–	33,5	d	
	26	22,1	–	21,7	q	
	27	22,1	–	21,7	q	
	28	106,0	–	105,6	t	
[25EN]	20	36,2	–	35,8	d	100,0
	21	18,5	–	18,1	q	
	22	34,1	–	33,9	t	
	23	31,9	–	31,6	t	
	24	41,9	–	41,6	d	
	25	150,5	–	150,1	s	
	26	109,5	–	109,2	t	
	27	18,9	–	18,6	q	
	28	20,3	–	20,0	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento	
[22,24-Me]	20	33,7 – 31,7	d	80,0
	21	13,0 – 12,1	q	
	22	41,9 – 35,2	d	
	23	40,0 – 33,2	t	
	24	36,9 – 35,7	d	
	25	31,6 – 28,8	d	
	26	21,3 – 17,8	q	
	27	20,1 – 17,7	q

[---]	20	178,6 – 158,3	s	98,8
	21	76,8 – 71,5	t	
	22	121,3 – 111,3	d	
	23	177,6 – 170,6	s

[---]	20	36,2 – 35,9	d	100,0
	21	19,1 – 18,7	q	
	22	34,0 – 33,7	t	
	23	29,2 – 26,1	t	
	24	50,2 – 45,7	d	
	25	29,1 – 26,2	d	
	26	19,0 – 18,7	q	
	27	19,7 – 19,4	q	
	28	23,2 – 23,0	t	
	29	12,3 – 11,8	q	
[22 α -OH]	20	43,0 – 42,1	d	100,0
	21	12,9 – 11,3	q	
	22	71,6 – 70,0	d	
	23	30,2 – 29,1	t	
	24	41,8 – 41,0	d	
	25	29,1 – 28,4	d	
	26	18,0 – 16,5	q	
	27	20,7 – 19,1	q	
	28	24,3 – 22,0	t	
	29	12,1 – 10,5	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[29-OH]	20	36,9 – 36,6	d	100,0
	21	19,1 – 18,8	q	
	22	34,2 – 33,9	t	
	23	27,9 – 27,5	t	
	24	41,2 – 41,0	d	
	25	29,9 – 29,6	d	
	26	18,8 – 18,2	q	
	27	19,7 – 19,2	q	
	28	34,3 – 34,0	t	
	29	62,3 – 62,0	t	
[22 α ,23 α -OH]	20	43,7 – 42,0	d	100,0
	21	14,6 – 13,8	q	
	22	73,0 – 72,0	d	
	23	71,6 – 70,5	d	
	24	50,9 – 49,4	d	
	25	27,8 – 26,7	d	
	26	22,2 – 21,6	q	
	27	17,8 – 17,3	q	
	28	19,7 – 18,2	t	
	29	14,8 – 14,3	q	
[22 β ,23 β -OH]	20	48,2 – 36,7	d	100,0
	21	12,5 – 11,8	q	
	22	75,5 – 74,0	d	
	23	73,5 – 72,0	d	
	24	51,9 – 46,2	d	
	25	30,3 – 23,8	d	
	26	23,2 – 19,2	q	
	27	21,0 – 18,1	q	
	28	20,1 – 18,7	t	
	29	15,0 – 13,3	q	
[22EN; 24-Et]	20	40,5 – 40,2	d	100,0
	21	21,3 – 21,0	q	
	22	138,5 – 137,8	d	
	23	129,5 – 129,3	d	
	24	51,4 – 51,0	d	
	25	32,0 – 31,7	d	
	26	19,5 – 19,0	q	
	27	21,2 – 18,8	q	
	28	25,5 – 25,2	t	
	29	12,3 – 12,0	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico			% de Reconhecimento	
[22EN; 24-Et; 29-OH]	20	40,8	–	40,6	d	100,0
	21	21,3	–	21,0	q	
	22	138,8	–	138,5	d	
	23	129,5	–	129,1	d	
	24	46,5	–	46,2	d	
	25	32,5	–	32,3	d	
	26	20,7	–	19,0	q	
	27	20,7	–	19,1	q	
	28	36,6	–	36,3	t	
	29	62,2	–	61,9	t	
[25EN; 28α]	20	35,9	–	35,5	d	100,0
	21	18,6	–	18,3	q	
	22	34,0	–	33,7	t	
	23	29,7	–	29,3	t	
	24	49,6	–	49,4	d	
	25	147,8	–	147,5	s	
	26	111,5	–	111,2	t	
	27	18,0	–	17,6	q	
	28	26,7	–	26,3	t	
	29	12,3	–	12,0	q	
[22EN;28α;29OXO;29OR]	20	40,4	–	40,0	d	100,0
	21	21,3	–	21,0	q	
	22	138,8	–	138,5	d	
	23	128,0	–	127,6	d	
	24	45,7	–	45,4	d	
	25	31,9	–	31,5	d	
	26	20,5	–	20,2	q	
	27	19,0	–	18,6	q	
	28	38,6	–	38,2	t	
	29	173,0	–	172,6	s	
[24α-Et; 29OXO; 29-OR]	20	36,6	–	36,3	d	100,0
	21	19,0	–	18,6	q	
	22	33,7	–	33,4	t	
	23	27,9	–	27,6	t	
	24	41,6	–	41,3	d	
	25	30,0	–	29,7	d	
	26	19,7	–	18,6	q	
	27	19,7	–	18,6	q	
	28	36,5	–	36,2	t	
	29	173,0	–	172,6	s	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico	% de Reconhecimento
--------------	------	--------------------------------	---------------------

[---]	20	34,0 – 33,2	d	100,0
	21	23,1 – 22,0	q	
	22	31,3 – 30,0	d	
	23	26,0 – 24,9	s	
	24	51,2 – 50,1	d	
	25	32,7 – 32,1	d	
	26	22,5 – 21,1	q	
	27	22,5 – 21,1	q	
	28	15,8 – 15,2	q	
	29	21,9 – 21,3	t	
	30	14,1 – 13,2	q	

[24 α ,25 α -EP; 26 α -OH]	20	41,7 – 41,5	d	100,0
	21	10,0 – 9,6	q	
	22	64,4 – 64,0	d	
	23	32,5 – 31,8	t	
	24	66,9 – 63,5	s	
	25	63,5 – 63,2	s	
	26	91,5 – 91,3	d	
	27	16,7 – 16,4	q	
	28	18,9 – 18,6	q	

[20 β -OH]	20	76,0 – 75,1	s	100,0
	21	21,0 – 20,2	q	
	22	80,9 – 80,0	d	
	23	32,3 – 31,0	t	
	24	32,2 – 30,8	d	
	25	40,7 – 39,3	d	
	26	177,1 – 175,8	s	
	27	15,0 – 14,0	q	
	28	21,9 – 21,0	q	

Tabela XVII. Continuação

Subesqueleto	No C	Faixas de deslocamento químico		% de Reconhecimento
[24EN]	20	42,5 – 35,2	d	100,0
	21	17,6 – 9,5	q	
	22	83,6 – 77,9	d	
	23	34,5 – 29,7	t	
	24	157,6 – 148,3	s	
	25	125,8 – 121,3	s	
	26	169,0 – 165,3	s	
	27	12,5 – 12,1	q	
	28	20,7 – 20,2	q	
[20β-OH; 24EN]	20	79,9 – 77,0	s	100,0
	21	22,2 – 19,3	q	
	22	83,6 – 80,3	d	
	23	34,5 – 32,7	t	
	24	154,1 – 148,5	s	
	25	121,6 – 121,4	s	
	26	169,0 – 165,6	s	
	27	12,6 – 12,1	q	
	28	20,7 – 20,2	q

[---]	20	124,0 – 116,1	s	100,0
	21	151,3 – 148,1	d	
	22	150,1 – 145,3	d	
	23	115,9 – 112,8	d	
	24	164,1 – 161,5	s	

1

Esqueleto: Vitanolídeo - Solanaceae – *Withania coagulans*¹⁴⁹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 204,5(s), 127,9(d), 145,2(d), 33,4(t), 135,9(s), 124,7(d), 31,6(t), 39,7(d), 40,1(d), 50,1(s), 21,9(t), 23,4(t), 49,6(s), 54,7(s), 29,6(t), 42,9(t), 56,6(d), 13,6(q), 18,9(q), 75,2(s), 20,5(q), 81,0(d), 30,6(t), 149,0(s), 122,0(s), 166,2(s), 12,7(q), 21,7(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	79,7	31,3	100,0	100,0	78,6
Colestano	10,1	20,0	---	---	7,5
17-Metilandrostando	---	29,5	---	---	6,7
Androstando	---	19,2	---	---	4,3
9,11-Secocolestano	5,6	---	---	---	1,6

Esqueleto: Estigmastano - Lamiaceae - *Ajuga salicifolia*¹⁵⁰

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 38,3(t), 32,1(t), 71,5(d), 38,6(t), 41,5(d), 30,9(t), 119,0(d), 140,6(s), 50,8(d), 35,3(s), 22,3(t), 34,0(t), 49,6(s), 48,7(d), 33,5(t), 81,3(d), 89,5(s), 13,4(q), 13,5(q), 40,3(d), 62,6(t), 79,0(d), 77,9(d), 55,3(d), 86,2(s), 18,7(q), 69,7(t), 23,1(t), 14,0(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Estigmastano	47,4	9,9	100,0	50,0	48,3
Poriferastano	5,2	---	---	50,0	17,6
Pregnano	4,5	49,5	---	---	12,5
Colestano	14,6	10,7	---	---	6,7
Cardanolídeo	---	19,9	---	---	4,5

3

Esqueleto: Gorgostano - Coral - *Capnella lacertiliensis*¹⁵¹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 36,9(t), 31,2(t), 71,2(d), 41,8(t), 146,0(s), 123,9(d), 64,9(d), 36,5(d), 41,5(d), 37,5(s), 27,0(t), 80,8(d), 46,1(s), 57,4(d), 23,6(t), 27,9(t), 48,3(d), 9,0(q), 18,1(q), 33,6(d), 22,2(q), 30,6(d), 25,3(s), 50,6(d), 32,2(d), 21,5(q), 22,2(q), 15,4(q), 21,5(t), 13,8(q); Ac:170,7(s), 21,8(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Gorgostano	46,9	29,7	---	100,0	62,7
Colestano	7,6	19,7	---	---	7,9
9,11-Secogorgostano	2,5	10,9	---	---	3,8
Ergostano	10,8	---	---	---	3,7
6-Metilcolestano	1,3	10,2	---	---	3,2

Macronós Propostos: Acetoil: 170,7, 21,5 – Erro médio: 1,100

4

Esqueleto: Ergostano - Coral - *Capnella lacertiliensis*¹⁵¹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 26,3(t), 31,4(t), 71,3(d), 41,3(t), 141,5(s), 120,4(d), 31,7(t), 27,2(d), 53,2(d), 36,9(s), 70,4(d), 82,7(d), 45,9(s), 57,2(d), 41,1(t), 70,7(d), 55,1(d), 10,7(q), 22,6(q), 39,6(d), 13,5(q), 23,6(d), 36,5(t), 35,0(d), 32,7(d), 18,0(q), 20,1(q), 15,2(q); Ac: 170,0(s), 21,7(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	47,5	49,6	---	50,0	49,0
Campestando	1,1	---	---	50,0	19,6
Androstano	---	30,4	---	---	8,2
Colestano	5,9	10,9	---	---	5,0
Estigmastano	8,6	---	---	---	3,0

Macronós Propostos: Acetoil: 170,7, 21,7 – Erro médio: 0,850

Esqueleto: Vitanolídeo - Solanaceae - *Withania somnifera*¹⁵²

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 210,2(s), 42,6(t), 70,4(d), 73,8(d), 63,8(s), 56,6(d), 29,8(t), 31,0(d), 42,5(d), 49,1(s), 20,7(t), 27,1(t), 42,6(s), 56,1(d), 24,2(t), 39,1(t), 51,8(d), 10,7(q), 13,6(q), 39,0(d), 12,5(q), 78,9(d), 29,6(t), 159,1(s), 122,5(s), 167,4(s), 61,6(t), 19,5(q); Glc: 102,7(d), 75,6(d), 76,9(d), 72,4(d), 76,8(d), 62,4(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	39,0	48,8	100,0	100,0	70,7
Ergostano	19,2	21,5	---	---	10,4
Colestano	18,0	10,1	---	---	7,5
Campestando	3,6	9,8	---	---	3,3
23,28-Cicloestigmastano	---	9,8	---	---	2,2

Macronós Propostos: Glicosil: 102,7, 76,9, 76,8, 75,6, 72,4, 62,4 – Erro médio: 0,650

Esqueleto: Colestano - Estrela-do-mar - *Pteraster pulvillus*¹⁵³

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 33,6(t), 23,4(t), 78,6(d), 74,1(d), 45,4(d), 26,5(t), 33,6(t), 36,9(d), 56,7(d), 36,6(s), 21,3(t), 40,4(t), 43,5(s), 57,9(d), 25,1(t), 28,6(t), 51,9(d), 12,7(q), 14,8(q), 41,4(d), 69,4(t), 30,8(t), 24,4(t), 40,7(t), 29,1(d), 23,0(q), 23,2(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	49,6	80,5	100,0	100,0	81,0
3,5-Ciclocolestano	---	10,3	---	---	2,3
27-Norcolestano	---	9,2	---	---	2,1
Estigmastano	6,5	---	---	---	1,9
Colano	5,9	---	---	---	1,7

7

Esqueleto: Bufanolídeo - Ranunculaceae - *Helleborus orientalis*¹⁵⁴

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 18,7(t), 25,6(t), 73,1(d), 35,4(t), 74,0(s), 37,0(t), 25,0(t), 42,3(d), 39,6(d), 55,3(s), 22,8(t), 41,0(t), 49,4(s), 84,3(s), 42,3(t), 72,4(d), 58,9(d), 17,1(q), 208,4(d), 119,2(s), 150,7(d), 151,3(d), 112,6(d), 162,2(s); Rha: 100,6(d), 72,5(d), 72,9(d), 70,8(d), 73,6(d), 18,6(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colano	51,0	10,3	---	100,0	58,9
Cardanolídeo	18,9	20,0	---	---	11,9
Colestano	5,6	30,6	---	---	10,22
Pregnano	6,1	19,2	---	---	7,3
Vitanolídeo	13,6	---	---	---	4,7

Macronós Propostos: Rhamnosil: 100,6, 73,6, 72,9, 72,5, 70,8, 18,6 – Erro médio: 1,100

Fucosil: 100,6, 73,6, 72,9, 72,5, 70,8, 18,6 – Erro médio: 0,917

8

Esqueleto: Colestano - Liliaceae - *Tupistra chinensis*¹⁵⁵

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 74,6(d), 40,0(t), 68,4(d), 37,1(t), 75,9(s), 35,6(t), 30,3(t), 34,9(d), 46,4(d), 44,6(s), 22,2(t), 40,7(t), 41,3(s), 56,9(d), 32,8(t), 82,2(d), 63,6(d), 17,2(q), 14,4(q), 42,6(d), 15,7(q), 110,2(s), 33,8(t), 29,6(t), 144,9(s), 65,7(t), 109,6(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	73,8	49,8	100,0	100,0	79,0
Cardanolídeo	5,4	29,1	---	---	8,1
Vitanolídeo	---	20,1	---	---	4,5
Pregnano	---	10,1	---	---	2,3
Ergostano	5,3	---	---	---	1,5

9

Esqueleto: Pregnano - Liliaceae - *Tupistra chinensis*¹⁵⁵

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (C₅D₅N): 75,1(d), 40,8(t), 66,5(d), 41,1(t), 140,5(s), 124,3(d), 31,9(t), 31,5(d), 51,8(d), 46,0(s), 24,3(t), 35,8(t), 44,7(s), 56,8(d), 32,5(t), 144,5(d), 155,6(s), 16,2(q), 13,2(q), 196,3(s), 27,0(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Pregnano	43,8	50,6	---	100,0	56,4
Colestano	---	29,3	100,0	---	22,7
Cardanolídeo	16,6	20,2	---	---	9,4
Androstano	23,6	---	---	---	6,9
17-Metilandrostando	5,1	---	---	---	1,5

10

Esqueleto: Colestano - Caryophyllaceae - *Silene italica* ssp. *nemoralis*¹⁵⁶

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 37,6(t), 69,7(d), 69,8(d), 36,3(t), 51,4(d), 206,7(s), 123,8(d), 162,6(s), 76,6(s), 43,0(s), 30,9(t), 29,6(t), 49,1(s), 87,2(s), 31,6(t), 21,4(t), 50,4(d), 18,5(q), 29,6(q), 78,0(s), 21,2(q), 78,6(d), 27,5(t), 42,5(t), 71,4(s), 29,1(q), 29,9(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	80,8	60,6	100,0	100,0	85,5
Cardanolídeo	---	10,1	---	---	4,1
Pregnano	6,2	9,9	---	---	2,2
Androstano	---	9,7	---	---	2,2
17-Metilandrostando	---	9,7	---	---	2,2

11

Esqueleto: Ergostano - Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 39,7(t), 26,2(t), 73,7(d), 69,1(d), 57,3(d), 64,8(d), 49,7(t), 77,4(s), 58,4(d), 38,1(s), 19,2(t), 43,3(t), 44,4(s), 62,7(d), 71,1(d), 42,5(t), 57,8(d), 16,5(q), 17,0(q), 36,0(d), 18,9(q), 35,5(t), 32,7(t), 153,8(s), 43,3(d), 67,5(t), 17,2(q), 109,4(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
8,9-Secoergostano	---	---	---	100,0	38,5
Ergostano	66,8	28,9	---	---	30,9
Colestano	7,9	21,7	---	---	8,6
27-Norcolestano	---	30,3	---	---	8,1
Estigmastano	9,7	---	---	---	3,3

12

Esqueleto: Colestano - Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 38,6(t), 32,0(t), 72,0(d), 33,0(t), 53,1(d), 70,0(d), 41,8(t), 31,6(d), 55,7(d), 37,5(s), 22,2(t), 42,5(t), 43,4(s), 62,2(d), 70,6(d), 42,8(t), 57,5(d), 15,3(q), 13,8(q), 41,1(d), 21,2(q), 139,2(d), 130,1(d), 78,9(d), 35,3(d), 18,8(q), 18,7(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	55,8	10,1	---	---	35,8
Pregnano	---	49,8	---	---	21,8
Androstano	---	30,5	---	---	13,4
Estigmastano	7,7	---	---	---	4,3
Gorgostano	---	9,7	---	---	4,2

13

Esqueleto: Colestano - Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 38,6(t), 32,0(t), 72,0(d), 33,0(t), 53,1(d), 70,0(d), 41,8(t), 31,6(d), 55,7(d), 37,5(s), 22,2(t), 42,7(t), 43,4(s), 62,2(d), 70,6(d), 42,3(t), 57,8(d), 15,2(q), 13,8(q), 37,1(d), 19,4(q), 33,5(t), 31,7(t), 78,1(d), 34,5(d), 19,5(q), 17,5(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	57,6	20,1	---	---	41,2
Pregnano	---	60,4	---	---	26,4
Ergostano	3,9	9,6	---	---	6,4
6-Metilcolestano	8,3	---	---	---	4,6
Androstano	---	10,0	---	---	4,4

14

Esqueleto: 23-Metilelgestano – Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 39,4(t), 31,5(t), 72,2(d), 32,4(t), 53,9(d), 67,6(d), 49,5(t), 77,5(d), 57,4(d), 38,0(s), 19,7(t), 43,3(t), 44,2(s), 62,7(d), 71,2(d), 42,8(t), 58,5(d), 16,8(q), 14,1(q), 35,3(d), 20,7(q), 132,9(d), 136,2(s), 45,9(d), 39,2(d), 67,2(t), 14,3(q), 16,4(q), 14,0(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	53,6	79,4	---	---	64,9
Estigmastano	11,5	---	---	---	6,4
Pregnano	---	11,4	---	---	5,0
Androstano	---	9,2	---	---	4,0
Ergostano	7,0	---	---	---	4,0

15

Esqueleto: Estigmastano - Verbenaceae - *Verbena littoralis*¹⁵⁸

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 36,9(t), 24,4(t), 72,2(d), 77,1(d), 146,6(s), 128,6(d), 64,6(d), 37,8(d), 42,5(d), 36,8(s), 23,9(t), 39,2(t), 42,2(s), 49,0(d), 19,9(t), 27,5(t), 53,0(d), 10,7(q), 18,6(q), 42,9(d), 11,6(q), 70,5(d), 29,4(t), 41,1(d), 28,8(d), 16,9(q), 19,4(q), 22,8(t), 10,8(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Estigmastano	44,7	---	---	50,0	34,7
Poriferastano	3,1	---	---	50,0	20,3
Colestano	24,3	29,4	---	---	16,3
Androstanos	---	41,0	---	---	11,0
Cardanolídeo	---	10,1	---	---	2,7

16

Esqueleto: 23-Metilergostano - Coral - *Isis hippuris*¹⁵⁹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 38,1(t), 38,0(t), 211,8(s), 44,5(t), 46,5(d), 28,6(t), 31,5(t), 33,8(d), 53,4(d), 35,7(s), 20,9(t), 39,3(t), 43,6(s), 50,9(d), 38,8(t), 218,8(s), 67,4(d), 14,4(q), 11,4(q), 77,1(s), 20,6(q), 81,1(d), 32,3(d), 44,5(d), 85,7(s), 23,6(q), 25,1(q), 9,5(q), 10,7(q); Ac: 171,4(s), 22,7(q); 170,3(s), 21,1(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	62,9	18,7	---	---	43,5
Pregnano	---	52,4	---	---	22,9
1-Metilandrostando	---	9,9	---	---	4,3
Cardanolídeo	---	9,6	---	---	4,2
Androstando	---	9,4	---	---	4,1

Macronós Propostos: Acetoil: 170,3, 21,1; 171,4, 20,6 – Erros médios: 0,700; 1,400

Isopropionil: 67,4, 23,6, 22,7 – Erro médio: 0,767

17

Esqueleto: Ergostano - Coral - *Sinularia gibberosa*¹⁶⁰

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 70,7(d), 37,5(t), 67,7(d), 37,3(t), 83,7(s), 211,5(s), 41,2(d), 37,4(d), 43,3(d), 49,1(s), 23,6(t), 39,7(t), 42,4(s), 56,8(d), 24,2(t), 27,7(t), 56,0(d), 11,9(q), 13,4(q), 35,6(d), 18,5(q), 34,5(t), 30,9(t), 156,7(s), 33,8(d), 21,8(q), 22,0(q), 106,0(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	20,8	11,2	100,0	100,0	57,0
Colestano	47,9	49,1	---	---	25,0
3-Metil-colestano	5,9	9,6	---	---	3,9
4-Metilergostano	1,3	10,4	---	---	2,7
6-Metil-colestano	1,5	9,7	---	---	2,6

18

Esqueleto: Ergostano - Fungo - *Gymnascella dankaliensis*¹⁶¹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (CDCl₃): 32,6(t), 29,6(t), 65,6(d), 121,9(d), 146,2(s), 66,5(d), 124,1(d), 136,0(s), 48,2(d), 38,2(s), 21,3(t), 40,1(t), 45,3(s), 72,7(s), 67,9(d), 29,6(t), 53,6(d), 15,9(q), 22,3(q), 39,0(d), 23,4(q), 135,5(d), 132,9(d), 44,0(d), 33,9(d), 20,1(q), 20,4(q), 18,1(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	41,7	19,7	100,0	100,0	64,9
Colestano	24,6	18,6	---	---	11,3
Androstano	---	22,1	---	---	5,0
Cardanolídeo	---	20,0	---	---	4,5
Campestando	9,8	---	---	---	2,9

19

Esqueleto: Vitanolídeo - Solanaceae – *Physalis philadelphica*¹⁶²

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 202,1(s), 129,3(d), 142,4(d), 69,1(d), 143,0(s), 131,4(d), 64,4(d), 37,2(d), 35,0(d), 50,0(s), 21,7(t), 39,7(t), 42,9(s), 50,0(d), 23,8(t), 22,6(t), 55,4(d), 13,4(q), 21,7(q), 75,5(s), 20,7(q), 80,4(d), 31,2(t), 31,3(d), 40,5(d), 176,0(s), 14,2(q), 21,4(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	35,8	30,9	100,0	100,0	65,8
Colestano	32,8	9,7	---	---	11,7
Pregnano	---	20,0	---	---	4,5
Campestando	10,6	---	---	---	3,1
Cardanolídeo	---	10,2	---	---	2,3

Esqueleto: Colestano - Liliaceae – *Tupistra wattii*¹⁶³

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 73,4(d), 33,0(t), 68,2(d), 34,4(t), 31,3(d), 26,7(t), 26,6(t), 35,7(d), 42,2(d), 40,4(s), 21,1(t), 40,7(t), 40,4(s), 56,5(d), 32,1(t), 81,6(d), 62,2(d), 16,8(q), 19,4(q), 42,2(d), 14,9(q), 111,8(s), 40,7(t), 81,6(d), 38,2(d), 65,1(t), 13,6(q); Glc-(1→6)-Glc: 106,1(d), 75,1(d), 78,5(d), 71,3(d), 76,7(d), 70,0(t), 105,6(d), 75,5(d), 78,4(d), 71,6(d), 78,5(d), 62,8(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	84,4	39,7	100,0	100,0	81,8
Androstano	---	20,1	---	---	4,6
Pregnano	---	19,6	---	---	4,4
Cardanolídeo	1,9	11,0	---	---	3,0
3,5-Cicloergostano	---	9,6	---	---	2,2

Macronós Propostos: Gentiobiosil: 106,1, 105,6, 78,5, 78,5, 78,4, 76,7, 75,5, 75,1, 71,6, 71,3, 70,0, 62,8 – Erro médio: 0,471

21

Esqueleto: Ergostano - Fungo – *Byssoschlamys nivea*¹⁶⁴

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 35,7(t), 26,4(t), 72,2(d), 25,5(t), 54,4(d), 209,3(s), 78,8(d), 42,4(d), 55,2(d), 40,6(s), 68,3(d), 48,8(t), 42,9(s), 56,0(d), 36,4(t), 74,6(d), 60,1(d), 15,4(q), 15,5(q), 34,3(d), 21,1(q), 134,9(d), 133,0(d), 43,2(d), 33,0(d), 19,6(q), 20,1(q), 18,0(q); Ac: 170,6(s), 21,3(q); 170,0(s), 21,5(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	22,4	90,4	---	100,0	70,5
Colestano	36,3	---	---	---	12,6
Estigmastano	14,4	---	---	---	5,0
Androstano	---	9,6	---	---	2,6
3,5-Cicloestigmastano	5,7	---	---	---	2,0

Macronós Propostos: Acetoil: 170,0, 21,1; 170,6, 21,3 – Erros médios: 0,550; 0,950

Isobutirol: 170,6, 34,3, 19,6, 18,0 – Erro médio: 1,325

22

Esqueleto: Vitanolídeo - Solanaceae – *Withania somnifera*¹⁶⁵

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (DMSO- d_6): 204,2(s), 127,7(d), 142,1(d), 31,9(t), 78,0(s), 73,4(d), 30,7(t), 35,4(d), 35,8(d), 54,2(s), 22,2(t), 33,6(t), 51,3(s), 82,5(s), 28,9(t), 34,4(t), 87,5(s), 20,8(q), 9,8(q), 76,3(s), 19,3(q), 81,3(d), 33,1(t), 150,8(s), 120,1(s), 166,1(s), 12,1(q), 20,3(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	98,7	71,0	100,0	100,0	93,1
Androstano	---	19,1	---	---	4,3
Ergostano	---	9,9	---	---	2,2
12,23-Cicloergostano	1,3	11,0	---	---	0,4

23

Esqueleto: Colestano - Alga – *Hypnea musciformis*¹⁶⁶

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 38,2(t), 37,8(t), 211,3(s), 37,6(t), 57,5(d), 209,2(s), 46,4(t), 37,5(d), 53,6(d), 41,5(s), 21,8(t), 40,0(t), 43,5(s), 56,9(d), 23,8(t), 29,5(t), 60,1(d), 13,9(q), 12,6(q), 75,2(s), 29,5(q), 139,3(d), 125,6(d), 42,1(t), 28,7(d), 22,5(q), 22,5(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	72,1	69,4	100,0	100,0	85,0
Ergostano	6,8	11,8	---	---	4,7
22-Metil-colestano	---	18,8	---	---	4,2
Estigmastano	3,6	---	---	---	1,0
6-Metil-colestano	3,4	---	---	---	1,0

Esqueleto: Colestano - Alga – *Hypnea musciformis*¹⁶⁶

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 37,1(t), 34,3(t), 200,1(s), 126,4(d), 168,0(s), 73,3(d), 38,5(t), 29,7(d), 53,6(d), 37,9(s), 21,0(t), 39,4(t), 42,2(s), 56,2(d), 24,8(t), 28,5(t), 55,9(d), 12,0(q), 19,5(q), 32,8(d), 18,7(q), 37,3(t), 24,5(t), 39,3(t), 28,1(d), 22,5(q), 22,5(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	57,7	49,6	100,0	100,0	76,3
Androstano	---	20,1	---	---	4,5
6-Metil-colestano	13,9	---	---	---	4,0
Pregnano	9,1	10,6	---	---	2,4
17-Metilandrostandano	---	10,0	---	---	2,3

25

Esqueleto: Bufanolídeo - Apocynaceae – *Catharanthus roseus*¹⁶⁷

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 30,6(t), 27,0(t), 74,2(d), 30,3(t), 36,5(d), 26,2(t), 21,2(t), 35,5(d), 39,8(d), 33,4(s), 21,0(t), 39,6(t), 45,3(s), 72,5(s), 59,8(d), 75,4(d), 50,5(d), 17,2(q), 23,6(q), 116,7(s), 152,6(d), 148,9(d), 113,9(d), 161,4(s); Glc: 103,2(d), 75,4(d), 78,8(d), 71,9(d), 78,4(d), 63,0(t); Ac: 169,9(s), 20,2(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colano	31,5	43,2	---	100,0	61,0
Colestano	22,4	---	---	---	7,7
Cardanolídeo	6,5	18,7	---	---	7,3
Androstano	---	19,3	---	---	5,2
Pregnano	11,6	---	---	---	4,0

Macronós Propostos: Acetoil: 169,9, 20,2 – Erro médio: 0,850

Glicosil: 103,2, 78,8, 78,4, 74,2, 71,9, 63,0 – Erro médio: 0,817

26

Esqueleto: Colano – Lampréia – *Petromyzon marinus*¹⁶⁸

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (DMSO- d_6): 22,7(t), 44,0(t), 210,6(s), 28,8(t), 41,4(d), 36,8(t), 65,4(d), 46,0(d), 38,9(d), 34,9(s), 27,1(t), 70,8(d), 45,8(s), 38,2(d), 30,8(t), 30,7(t), 39,5(d), 10,0(q), 12,3(q), 35,0(d), 16,9(q), 37,9(t), 37,7(t), 174,8(s)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colano	30,4	40,4	---	100,0	59,9
Pregnano	24,3	19,9	---	---	13,8
Colestano	16,8	10,0	---	---	8,5
Androstano	---	29,8	---	---	8,0
Cardanolídeo	11,0	---	---	---	3,8

27

Esqueleto: Estigmastano - Lamiaceae - *Ajuga salicifolia*¹⁶⁹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 38,5(t), 30,4(t), 79,0(d), 35,2(t), 41,4(d), 30,9(t), 118,9(d), 140,7(s), 50,8(d), 35,5(s), 22,2(t), 34,0(t), 49,5(s), 49,0(d), 33,6(t), 82,2(d), 87,8(s), 13,6(q), 13,5(q), 35,3(d), 12,8(q), 81,8(d), 78,5(d), 56,2(d), 84,1(s), 18,0(q), 72,0(t), 22,9(t), 13,8(q); Glc: 102,3(d), 75,1(d), 78,1(d), 71,7(d), 77,9(d), 62,8(t); Ac: 172,7(s), 20,7(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Estigmastano	58,2	9,9	100,0	50,0	51,4
Poriferastano	3,9	---	---	50,0	17,3
Androstano	---	60,7	---	---	13,7
Colestano	4,1	9,8	---	---	3,4
Ergostano	10,3	---	---	---	3,0

Macronós Propostos: Glicosil: 102,3, 78,1, 77,9, 75,1, 71,7, 62,8 – Erro médio: 0,567

Acetoil: 172,7, 20,7 – Erro médio: 2,000

28

Esqueleto: Colestano - Zoanthidae – *Zoanthus sp.*¹⁷⁰

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (C₅D₅N): 77,7(d), 69,9(d), 72,1(d), 33,3(t), 48,0(d), 204,8(s), 123,3(d), 167,2(s), 36,4(d), 45,0(s), 23,0(t), 33,6(t), 49,3(s), 85,7(s), 31,4(t), 21,5(t), 51,7(d), 19,5(q), 22,9(q), 78,4(s), 23,0(q), 78,4(d), 29,8(t), 38,7(t), 31,4(d), 24,0(q), 24,9(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	79,9	20,1	---	100,0	71,5
17-Metilandrostando	---	30,1	---	---	8,1
Cardanolídeo	7,0	10,1	---	---	5,1
Androstando	---	10,4	---	---	2,8
9,11-Seco-colestano	---	9,9	---	---	2,7

29

Esqueleto: Estigmastano - Polymastiidae – *Polymastia tenax*¹⁷¹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 32,1(t), 31,1(t), 68,6(d), 39,5(t), 67,7(s), 61,3(d), 65,1(d), 125,0(s), 38,6(d), 35,8(s), 18,9(t), 36,6(t), 43,1(s), 152,6(s), 24,9(t), 26,6(t), 56,6(d), 17,9(q), 16,5(q), 34,9(d), 18,9(q), 33,6(t), 26,2(t), 46,0(d), 28,9(d), 19,0(q), 19,5(q), 22,9(t), 12,3(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Estigmastano	61,9	---	100,0	50,0	50,2
Poriferastano	1,7	---	---	50,0	16,6
Colestano	4,8	61,0	---	---	15,2
Androstano	---	19,4	---	---	4,4
Ergostano	1,5	9,8	---	---	2,7

30

Esqueleto: Ergostano - Coral – *Nephthea chabroli*¹⁷²

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{CDCl}_3/\text{CD}_3\text{OD}$): 33,6(t), 31,0(t), 71,2(d), 42,2(t), 50,6(d), 39,6(t), 21,6(t), 39,0(d), 141,5(s), 42,5(s), 128,7(d), 65,1(d), 42,3(s), 42,9(d), 24,1(t), 28,2(t), 55,8(d), 12,0(q), 63,2(t), 35,7(d), 18,8(q), 34,8(t), 31,8(t), 156,9(s), 33,9(d), 22,0(q), 21,9(q), 106,0(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	62,0	39,2	100,0	50,0	59,1
Campestando	---	19,8	---	50,0	20,6
Colestano	1,6	31,3	---	---	7,5
4-Metilergostano	2,4	9,8	---	---	2,9
Colano	9,5	---	---	---	2,8

31

Esqueleto: Pregnano - Asclepiadaceae – *Solenostemma argel*¹⁷³

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 26,6(t), 33,1(t), 199,2(s), 123,6(d), 170,5(s), 33,9(t), 35,9(t), 39,9(d), 48,5(d), 38,7(s), 20,2(t), 37,5(t), 47,4(s), 81,6(s), 73,7(d), 36,3(t), 59,9(d), 15,7(q), 17,7(q), 217,5(s), 32,9(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Pregnano	30,8	29,5	75,0	100,0	59,9
Androstano	31,0	40,0	---	---	18,0
Cardanolídeo	16,5	10,0	25,0	---	11,1
17-Metilandrostando	10,3	10,1	---	---	5,3
20-Metil-pregnano	6,1	10,4	---	---	4,1

32

Esqueleto: Vitanolídeo - Solanaceae – *Solanum cilistum*¹⁷⁴

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 203,8(s), 127,9(d), 145,7(d), 33,7(t), 136,3(s), 124,9(d), 31,2(t), 33,7(d), 43,3(d), 50,8(s), 37,5(t), 32,8(t), 48,1(s), 50,6(d), 24,0(t), 24,1(t), 85,0(s), 15,1(q), 19,0(q), 44,0(d), 10,1(q), 67,0(d), 34,0(t), 63,1(s), 63,0(s), 92,8(d), 17,4(q), 19,0(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	66,0	33,6	100,0	100,0	75,1
Colestano	8,5	9,4	---	---	4,6
17-Metilandrostando	---	19,2	---	---	4,3
Pregnano	---	18,8	---	---	4,2
9,11-Secocolestano	12,9	---	---	---	3,7

33

Esqueleto: Vitanolídeo - Solanaceae – *Solanum cilistum*¹⁷⁴

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 203,9(s), 128,0(d), 145,8(d), 33,6(t), 136,4(s), 124,9(d), 31,3(t), 33,8(d), 43,4(d), 50,8(s), 37,4(t), 32,9(t), 48,3(s), 50,7(d), 24,0(t), 24,2(t), 85,1(s), 15,3(q), 19,0(q), 43,8(d), 10,5(q), 70,1(d), 41,1(t), 75,3(s), 74,5(s), 98,4(d), 16,2(q), 22,6(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	37,9	33,0	100,0	100,0	66,8
Ergostano	19,8	---	---	---	5,8
Colestano	15,6	---	---	---	4,5
Cardanolídeo	---	19,1	---	---	4,3
17-Metilandrostando	---	19,1	---	---	4,3

34

Esqueleto: Androstano - Meliaceae – *Aglaia rubiginosa*¹⁷⁵

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 155,1(d), 127,7(d), 186,1(s), 119,8(d), 169,6(s), 68,0(d), 41,1(t), 33,6(d), 52,1(d), 43,5(s), 22,0(t), 31,1(t), 47,7(s), 50,2(d), 21,9(t), 35,6(t), 219,5(s), 13,8(q), 19,1(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Androstano	24,5	68,3	25,0	100,0	58,8
Pregnano	25,5	---	62,5	---	17,5
17-Metilandrostan	39,0	21,4	---	---	16,2
3,5-Cicloandrostan	2,2	10,3	---	---	3,0
4-Metilergostano	---	---	12,5	---	2,0

35

Esqueleto: Colestano - Meliaceae – *Aglaia rubiginosa*¹⁷⁵

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (C₅D₅N): 45,1(t), 73,0(d), 72,7(d), 76,0(d), 45,9(d), 27,2(t), 119,2(d), 138,0(s), 52,8(d), 34,4(s), 21,3(t), 39,8(t), 44,2(s), 54,9(d), 23,6(t), 27,6(t), 54,3(d), 12,1(q), 18,5(q), 39,2(d), 13,0(q), 80,6(d), 25,7(t), 39,1(t), 79,6(s), 28,2(q), 29,0(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	40,9	61,0	---	100,0	57,9
Pregnano	---	---	62,5	---	10,1
Androstano	---	9,6	25,0	---	6,2
6-Metil-colestano	2,7	19,6	---	---	5,2
4-Metilergostano	5,9	---	12,5	---	3,7

36

Esqueleto: Colestano - Meliaceae – *Aglaia rubiginosa*¹⁷⁵

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 37,3(t), 25,6(t), 72,6(d), 73,2(d), 44,4(d), 26,0(t), 117,9(d), 139,0(s), 50,7(d), 34,2(s), 21,0(t), 39,4(t), 43,9(s), 54,6(d), 23,1(t), 27,3(t), 53,7(d), 11,8(q), 15,2(q), 38,5(d), 12,4(q), 80,2(d), 24,8(t), 38,9(t), 79,7(s), 28,0(q), 28,6(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	33,7	9,3	---	100,0	44,2
Androstano	---	49,6	25,0	---	15,2
Pregnano	---	10,4	62,5	---	12,4
4-Metilergostano	7,4	10,3	12,5	---	6,5
Estigmastano	11,5	---	---	---	3,3

37

Esqueleto: Estigmastano - Sapindaceae – *Schleicheria oleosa*¹⁷⁶

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 37,1(t), 31,6(t), 71,5(d), 42,0(t), 144,0(s), 121,4(d), 81,9(d), 37,2(d), 48,4(d), 36,5(s), 21,2(t), 39,6(t), 43,2(s), 56,0(d), 25,8(t), 27,7(t), 52,5(d), 11,9(q), 19,0(q), 42,5(d), 12,4(q), 71,3(d), 29,9(t), 41,4(d), 28,8(d), 17,6(q), 20,5(q), 23,5(t), 11,8(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	43,8	29,8	---	---	23,2
Estigmastano	10,7	---	---	50,0	22,9
Poriferastano	4,3	---	---	50,0	20,7
30-Nor-gorgostano	---	20,3	---	---	5,5
Androstano	---	19,6	---	---	5,3

38

Esqueleto: Ergostano - Sapindaceae – *Schleicheria oleosa*¹⁷⁶

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 36,4(t), 32,2(t), 70,6(d), 41,8(t), 165,2(s), 126,1(d), 199,8(s), 45,5(d), 49,9(d), 38,4(s), 21,2(t), 38,7(t), 43,4(s), 49,5(d), 26,4(t), 27,9(t), 51,8(d), 12,0(q), 17,3(q), 42,7(d), 12,5(q), 71,9(d), 34,6(t), 35,4(d), 29,7(d), 16,2(q), 21,1(q), 15,7(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	33,8	20,0	---	50,0	36,3
Campestando	3,1	---	---	50,0	20,3
Colestano	13,2	40,0	---	---	15,3
Estigmastano	16,4	---	---	---	5,7
Androstano	---	20,4	---	---	5,5

39

Esqueleto: 9,11-Seco-Ergostano - Coral – *Pachyclavularia violacea*¹⁷⁷

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 28,5(t), 30,5(t), 68,0(d), 38,5(t), 65,5(s), 58,1(d), 26,3(t), 38,8(d), 214,6(s), 46,5(s), 58,8(t), 41,1(t), 45,6(s), 45,2(d), 22,5(t), 25,8(t), 49,5(d), 18,1(q), 19,7(q), 34,4(d), 19,6(q), 33,0(t), 31,4(t), 39,1(d), 31,5(d), 20,4(q), 17,6(q), 15,4(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	52,8	10,4	6,7	---	18,8
Ergostano	8,0	29,2	60,0	---	18,6
9,11-Secoergostano	5,3	---	6,7	25,0	10,7
Androstano	---	40,4	---	---	9,1
5,6-Secoergostano	---	---	---	25,0	8,1
9,11-Secocampestano	---	---	---	25,0	8,1
5,6-Secocampestano	---	---	---	25,0	8,1

40

Esqueleto: Cardanolídeo - Cruciferae – *Erysimum cheiranthoides*¹⁷⁸

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (C₅D₅N): 24,2(t), 26,7(t), 75,3(d), 36,4(t), 73,8(s), 37,2(t), 24,2(t), 40,5(d), 39,4(d), 55,1(s), 21,5(t), 40,1(t), 50,3(s), 84,8(s), 32,5(t), 27,5(t), 51,2(d), 16,5(q), 177,0(s), 176,0(s), 73,8(t), 117,7(d), 174,6(s); Dig: 97,7(d), 39,8(t), 68,9(d), 82,5(d), 67,5(d), 18,5(q); Rha: 104,1(d), 72,4(d), 72,5(d), 73,8(d), 70,4(d), 18,6(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Cardanolídeo	87,1	91,1	---	100,0	93,2
Colestano	2,6	8,9	---	---	3,3
Androstano	3,0	---	---	---	1,0
Pregnano	2,2	---	---	---	0,8
27-Nor-colestano	1,6	---	---	---	0,6

Macronó Proposto: Rhamnosil: 104,1, 73,8, 72,5, 72,4, 70,4, 18,6 – Erro médio: 0,750

41

Esqueleto: Ergostano - Coral – *Clavularia viridis*¹⁷⁹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 205,3(s), 128,6(d), 140,1(d), 31,7(t), 78,2(s), 66,0(d), 38,8(t), 34,0(d), 54,3(d), 54,5(s), 66,8(d), 48,7(t), 43,1(s), 54,7(d), 28,4(t), 23,8(t), 55,7(d), 13,2(q), 9,4(q), 40,0(d), 20,9(q), 135,1(d), 132,3(d), 42,8(d), 33,1(d), 19,6(q), 19,9(q), 17,6(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	72,0	100,0	87,5	100,0	89,8
30-Norgorgostano	2,9	---	12,5	---	2,8
3,5-Cicloestigmastano	6,6	---	---	---	1,9
Estigmastano	4,8	---	---	---	1,4
Campestando	3,8	---	---	---	1,1

Esqueleto: Colestano - Esponja – *Dysidea sp.*¹⁸⁰

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 29,5(t), 30,5(t), 66,7(d), 36,4(t), 75,6(s), 69,7(d), 67,2(d), 132,2(s), 142,8(s), 44,3(s), 66,1(d), 49,4(t), 47,4(s), 48,2(d), 22,6(t), 28,5(t), 56,1(d), 13,7(q), 24,4(q), 36,2(d), 18,6(q), 36,1(t), 24,0(t), 39,7(t), 28,2(d), 23,0(q), 22,7(q); Ac: 170,4(s), 21,1(q); 170,1(s), 20,8(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	59,4	100,0	37,5	100,0	78,1
Ergostano	8,1	---	25,0	---	6,4
9,11-Secocolestano	---	---	37,5	---	6,0
Cardanolídeo	10,5	---	---	---	3,1
Estigmastano	3,8	---	---	---	1,1

Macronós Propostos: Acetoil: 170,1, 21,1 – Erro médio: 0,600

Acetoil: 170,4, 20,8 – Erro médio: 0,800

43

Esqueleto: Pregnano - Asclepiadaceae – *Caralluma russeliana*¹⁸¹

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 38,7(t), 30,8(t), 79,9(d), 39,7(t), 140,9(s), 123,3(d), 28,4(t), 38,4(d), 47,8(d), 38,0(s), 22,3(t), 41,6(t), 49,0(s), 86,1(s), 33,9(t), 20,1(t), 58,1(d), 15,4(q), 20,3(q), 79,1(d), 21,4(q); MDGal: 102,9(d), 71,3(d), 84,7(d), 68,8(d), 71,8(d), 17,1(q), 57,3(q); Glc: 104,5(d), 75,4(d), 78,1(d), 71,4(d), 78,7(d), 62,8(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Pregnano	22,2	---	75,0	100,0	50,8
Cardanolídeo	29,3	---	25,0	---	12,5
Androstano	26,4	20,4	---	---	12,3
Colestano	---	30,1	---	---	6,8
Ergostano	---	29,6	---	---	6,7

Macronós Propostos: Glicosil: 104,5, 78,7, 78,1, 75,4, 71,4, 62,8 – Erro médio: 1,133

3-O-Metil-6-deoxigalactosil: 102,9, 84,7, 71,8, 71,4, 71,3, 68,8, 57,3, 17,1 – Erro médio: 0,321

Esqueleto: Colestano - Agavaceae – *Yucca schidigera*¹⁸²

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (CD₃OD): 31,4(t), 27,3(t), 76,0(d), 30,9(t), 37,5(d), 27,6(t), 27,7(t), 36,7(d), 41,2(d), 36,1(s), 21,9(t), 41,3(t), 42,2(s), 57,6(d), 32,7(t), 83,4(d), 65,3(d), 16,9(q), 24,5(q), 41,2(d), 16,2(q), 113,9(s), 31,4(t), 28,9(t), 35,0(d), 75,8(t), 17,4(q); Glc(2→1)Glc: 101,1(d), 80,6(d), 77,8(d), 71,5(d), 78,2(d), 62,4(t), 104,8(d), 75,9(d), 77,7(d), 71,8(d), 78,0(d), 62,4(t); Glc: 104,7(d), 74,9(d), 77,3(d), 71,5(d), 77,8(d), 62,4(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	78,7	40,7	100,0	100,0	80,4
Colano	15,7	---	---	---	4,6
2-Metilandrostando	---	19,4	---	---	4,4
Cardanolídeo	2,2	10,1	---	---	2,9
Androstando	---	9,8	---	---	2,2

Macronós Propostos: Glicosil: 104,8, 77,8, 77,3, 74,9, 71,5, 62,4 – Erro médio: 0,383

Soforosil [Glc(2→1)Glc]: 104,7, 101,1, 80,6, 78,2, 78,0, 77,8, 77,7, 75,9, 71,8, 71,5, 62,4, 62,4 – Erro médio: 0,337

45

Esqueleto: Androstano - Fungo – *Botrytis cinerea*¹⁸³

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ¹³C (CDCl₃): 37,1(t), 34,4(t), 213,5(s), 34,8(t), 40,3(d), 30,4(t), 70,1(d), 38,0(d), 59,4(d), 36,5(s), 20,5(t), 42,2(t), 43,5(s), 51,6(d), 23,2(t), 29,9(t), 81,5(d), 11,1(q), 23,1(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Androstano	41,8	50,8	---	100,0	66,6
Pregnano	25,1	---	---	---	8,7
Ergostano	---	19,4	---	---	5,2
Cardanolídeo	15,0	---	---	---	5,2
Estrano	1,6	10,3	---	---	3,3

46

Esqueleto: Pregnano - Fungo – *Botrytis cinerea*¹⁸³

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CD_3OD): 37,2(t), 34,2(t), 199,9(s), 125,6(d), 166,7(s), 43,1(t), 23,9(t), 42,4(d), 56,3(d), 39,4(s), 73,9(d), 50,5(t), 44,8(s), 54,6(d), 27,1(t), 68,8(d), 62,3(d), 14,6(q), 18,4(q), 208,9(s), 31,4(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Pregnano	47,6	70,9	---	100,0	74,0
Androstano	20,3	---	---	---	7,0
17-Metilandrostando	10,4	10,3	---	---	6,4
Cardanolídeo	13,3	---	---	---	4,6
Colestano	---	9,4	---	---	2,5

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Colestano	71,1	40,3	100,0	100,0	78,1
27-Nor-colestano	1,6	30,0	---	---	7,3
Pregnano	---	19,7	---	---	4,4
9,11-Secocolestano	---	10,0	---	---	2,3
Cardanolídeo	6,3	---	---	---	1,8

Galoil: 165,5, 153,6, 153,6, 142,9, 126,7, 107,5, 107,5 – Erro médio: 2,57

48

Esqueleto: Vitanolídeo - Solanaceae – *Vassobia lorentzii*¹⁸⁵

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 202,5(s), 131,9(d), 142,1(d), 69,7(d), 63,7(s), 62,6(d), 31,3(t), 30,9(d), 43,9(d), 47,8(s), 24,6(t), 34,6(t), 58,0(s), 56,4(d), 25,4(t), 26,2(t), 57,0(d), 101,0(d), 17,6(q), 85,1(s), 21,8(q), 80,5(d), 31,5(t), 147,7(s), 122,3(s), 165,7(s), 12,5(q), 20,5(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Vitanolídeo	86,2	33,8	100,0	100,0	81,1
Colestano	5,5	27,6	---	---	7,8
Ergostano	2,1	10,1	---	---	2,9
Colano	---	9,5	---	---	2,1
Pregnano	---	9,5	---	---	2,1

49

Esqueleto: 17(13→18)-Abeoergostano - Solanaceae – *Salpichroa origanifolia*¹⁸⁶

Subestrutura e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C (CDCl_3): 202,7(s), 128,9(d), 142,5(d), 33,6(t), 64,6(s), 59,0(d), 30,3(t), 33,2(d), 36,4(d), 48,7(s), 25,4(t), 30,6(t), 137,7(s), 137,1(d), 125,4(s), 126,5(d), 141,3(d), 128,2(s), 14,8(q), 43,9(d), 18,0(q), 71,7(d), 40,2(t), 73,7(s), 76,5(s), 97,2(d), 20,9(q), 24,2(q)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
17(13→18)-Abeoergostano	---	---	40,0	100,0	54,5
Vitanolídeo	---	---	60,0	---	13,6
Cardanolídeo	---	40,2	---	---	12,8
Ergostano	---	40,1	---	---	12,8
Androstano	---	10,1	---	---	3,2

50

Esqueleto: Ergostano - Coral – *Plexaurella grisea*¹⁸⁷

Subestruturas e atribuição de sinais propostos pelo programa SISCONST

Dados de RMN ^{13}C ($\text{C}_5\text{D}_5\text{N}$): 33,4(t), 32,6(t), 67,1(d), 43,1(t), 76,8(s), 79,5(d), 73,0(d), 39,5(d), 44,8(d), 38,6(s), 22,0(t), 40,9(t), 43,9(s), 56,5(d), 27,9(t), 29,2(t), 55,8(d), 12,6(q), 17,6(q), 36,2(d), 19,1(q), 35,2(t), 31,4(t), 156,8(s), 34,1(d), 22,2(q), 22,0(q), 106,7(t)

Esqueletos	SISCONST	C13MACH	SISOCBOT	REGRAS	GLOBAL
Ergostano	26,0	39,5	100,0	50,0	48,7
Colestano	43,6	31,4	---	---	19,7
Campestando	1,5	9,7	---	50,0	18,7
6-Metil-colestano	8,6	---	---	---	2,5
Androstano	---	9,9	---	---	2,2

1

Esqueleto: Vitanolídeo - Solanaceae – *Withania coagulans*¹⁴⁹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	204,5	204,4	0,1	204,2	0,3
2	127,9	128,0	0,1	128,0	0,1
3	145,2	145,2	0,0	145,2	0,0
4	33,4	33,5	0,1	33,5	0,1
5	135,9	136,0	0,1	136,0	0,1
6	124,7	124,8	0,1	124,6	0,1
7	31,6	31,2	0,4	30,8	0,8
8	39,7	39,0	0,3	33,3	6,4
9	40,1	40,5	0,4	43,0	2,9
10	50,1	51,0	0,9	51,8	1,7
11	21,9	22,2	0,3	23,7	1,8
12	23,4	25,0	0,6	39,9	16,5
13	49,6	48,5	1,1	43,6	6,0
14	54,7	55,0	0,3	55,9	1,2
15	29,6	30,0	0,4	23,7	5,9
16	42,9	42,0	0,9	22,4	20,5
17	56,6	56,0	0,6	54,6	2,0
18	13,6	14,0	0,4	14,1	0,5
19	18,9	18,5	0,4	18,3	0,6
20	75,2	74,8	0,4	74,6	0,6
21	20,5	20,7	0,2	20,7	0,2
22	81,0	81,1	0,1	81,1	0,1
23	30,6	31,0	0,4	31,5	0,9
24	149,0	149,3	0,3	149,4	0,4
25	122,0	121,8	0,2	121,9	0,1
26	166,2	166,4	0,2	166,4	0,2
27	12,7	12,5	0,2	12,4	0,3
28	21,7	21,3	0,4	20,5	1,2
Maior erro observado			1,10		20,50
Erro médio			0,35		2,55
Desvio padrão			0,27		4,89

Esqueleto: Estigmastano - Lamiaceae - *Ajuga salicifolia*¹⁵⁰

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,3	37,8	0,5	37,3	1,0
2	32,1	31,7	0,4	31,5	0,6
3	71,5	71,0	0,5	71,0	0,5
4	38,6	38,1	0,5	38,1	0,5
5	41,5	40,8	0,7	40,4	1,1
6	30,9	29,9	1,0	29,7	1,2
7	119,0	118,4	0,6	118,4	0,6
8	140,6	140,9	0,3	140,9	0,3
9	50,8	48,7	2,1	45,4	5,4
10	35,3	34,6	0,7	34,3	1,0
11	22,3	21,5	0,8	21,3	1,0
12	34,0	34,6	0,6	35,5	1,5
13	49,6	47,9	1,7	43,7	5,9
14	48,7	48,0	0,7	48,0	0,7
15	33,5	34,1	0,6	35,8	2,3
16	81,3	80,7	0,6	76,5	4,8
17	89,5	89,5	0,0	89,5	0,0
18	13,4	14,0	0,6	16,0	2,6
19	13,5	13,0	0,5	13,0	0,5
20	40,3	41,2	0,9	41,4	1,1
21	62,6	63,3	0,7	63,3	0,7
22	79,0	80,4	1,4	82,3	3,3
23	77,9	77,2	0,7	75,8	2,1
24	55,3	53,4	1,9	47,2	8,1
25	86,2	85,1	1,1	85,4	0,8
26	18,7	19,4	0,7	23,4	4,7
27	69,7	68,2	1,5	67,9	1,8
28	23,1	22,7	0,4	18,7	4,4
29	14,0	13,0	1,0	13,0	1,0
Maior erro observado			2,10		8,10
Erro médio			0,82		2,05
Desvio padrão			0,48		2,03

3

Esqueleto: Gorgostano - Coral - *Capnella lacertiliensis*¹⁵¹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	36,9	37,1	0,2	37,4	0,5
2	31,2	31,5	0,3	31,8	0,6
3	71,2	71,3	0,1	71,3	0,1
4	41,8	42,0	0,2	42,5	0,7
5	146,0	145,0	1,0	144,4	1,6
6	123,9	124,5	0,6	125,3	1,4
7	64,9	65,5	0,6	69,9	5,0
8	36,5	37,0	0,5	39,9	3,4
9	41,5	42,3	0,8	44,8	3,3
10	37,5	35,2	2,3	34,9	2,6
11	27,0	27,4	0,4	27,4	0,4
12	80,8	79,9	0,9	78,0	2,8
13	46,1	46,5	0,4	47,8	1,7
14	57,4	56,4	1,0	50,1	7,3
15	23,6	24,1	0,5	25,4	1,8
16	27,9	27,0	0,9	26,7	1,2
17	48,3	48,4	0,1	50,5	2,2
18	9,0	9,9	0,9	11,5	2,5
19	18,1	18,1	0,0	18,5	0,4
20	33,6	34,3	0,7	35,9	2,3
21	22,2	21,0	1,2	18,7	3,5
22	30,6	30,2	0,4	19,3	11,3
23	25,3	26,0	0,7	26,0	0,7
24	50,6	50,9	0,3	50,9	0,3
25	32,2	32,0	0,2	27,3	5,9
26	21,5	20,1	1,4	19,0	2,5
27	22,2	21,8	0,4	21,0	1,2
28	15,4	14,3	1,1	11,3	4,1
29	21,5	21,6	0,1	21,6	0,1
30	13,8	15,1	1,3	23,1	9,3
Ac	170,7	169,9	0,8	169,8	0,9
	21,8	21,5	0,3	20,9	0,9
Maior erro observado			2,30		11,30
Erro médio			0,64		2,58
Desvio padrão			0,49		2,67

4

Esqueleto: Ergostano - Coral - *Capnella lacertiliensis*¹⁵¹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	26,3	28,0	1,7	39,0	12,7
2	31,4	31,8	0,4	31,9	0,5
3	71,3	71,5	0,2	71,5	0,2
4	41,3	41,9	0,6	42,4	1,1
5	141,5	141,4	0,1	141,3	0,2
6	120,4	120,7	0,3	121,7	1,3
7	31,7	32,0	0,3	32,8	1,1
8	27,2	28,1	0,9	31,9	4,7
9	53,2	51,5	1,7	47,2	6,0
10	36,9	37,2	0,3	37,8	0,9
11	70,4	70,8	0,4	70,8	0,4
12	82,7	81,9	0,8	80,0	2,7
13	45,9	46,3	0,4	46,6	0,7
14	57,2	55,5	1,7	52,1	5,1
15	41,1	40,0	1,1	38,5	2,6
16	70,7	71,1	0,4	72,4	1,7
17	55,1	55,8	0,7	57,5	2,4
18	10,7	11,2	0,5	13,2	2,5
19	22,6	21,6	1,0	21,9	0,7
20	39,6	39,0	0,6	31,0	8,6
21	13,5	14,2	0,7	18,7	5,2
22	23,6	25,0	1,4	34,9	11,3
23	36,5	35,3	1,2	30,7	5,8
24	35,0	35,7	0,7	38,7	3,7
25	32,7	32,5	0,2	31,7	1,0
26	18,0	18,3	0,3	18,3	0,3
27	20,1	20,5	0,4	20,5	0,4
28	15,2	15,5	0,3	15,5	0,3
Ac	170,0	170,9	0,9	172,0	2,0
	21,7	21,3	0,4	21,0	0,7
Maior erro observado			1,70		12,70
Erro médio			0,69		2,89
Desvio padrão			0,47		3,28

Esqueleto: Vitanolídeo - Solanaceae - *Withania somnifera*¹⁵²

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	210,2	210,4	0,2	210,3	0,1
2	42,6	43,1	0,5	44,1	1,5
3	70,4	70,2	0,2	69,2	1,2
4	73,8	74,5	0,7	78,8	5,0
5	63,8	64,2	0,4	65,2	1,4
6	56,6	57,8	1,2	58,7	2,1
7	29,8	30,3	0,5	31,3	1,5
8	31,0	30,0	1,0	29,6	1,4
9	42,5	42,9	0,4	43,0	0,5
10	49,1	49,9	0,8	51,0	1,9
11	20,7	21,5	0,8	21,6	0,9
12	27,1	28,9	1,8	38,1	11,0
13	42,6	42,7	0,1	42,7	0,1
14	56,1	56,2	0,1	56,2	0,1
15	24,2	24,4	0,2	24,5	0,3
16	39,1	39,0	0,1	24,3	14,8
17	51,8	52,0	0,2	52,1	0,3
18	10,7	11,0	0,3	15,8	5,1
19	13,6	14,0	0,4	15,4	1,8
20	39,0	38,8	0,2	38,8	0,2
21	12,5	13,0	0,5	13,3	0,8
22	78,9	78,5	0,4	78,2	0,7
23	29,6	30,1	0,5	30,1	0,5
24	159,1	158,7	0,4	156,6	2,5
25	122,5	122,8	0,3	123,1	0,6
26	167,4	167,2	0,2	166,2	1,2
27	61,6	62,4	0,8	63,4	1,8
28	19,5	19,0	0,5	20,6	1,1
Glc-1'	102,7	103,0	0,3	104,6	1,9
2'	75,6	75,3	0,3	75,0	0,6
3'	76,9	76,8	0,1	78,4	1,5
4'	72,4	72,7	0,3	71,4	1,0
5'	76,8	76,9	0,1	78,4	1,6
6'	62,4	63,1	0,7	62,8	0,4
Maior erro observado			1,80		14,80
Erro médio			0,46		1,98
Desvio padrão			0,36		3,03

Esqueleto: Colestano - Estrela-do-mar - *Pteraster pulvillus*¹⁵³

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	33,6	34,0	0,4	36,3	2,7
2	23,4	24,1	0,7	28,7	5,3
3	78,6	78,9	0,3	78,9	0,3
4	74,1	75,0	0,9	75,7	1,6
5	45,4	45,0	0,4	36,1	9,3
6	26,5	27,0	0,5	27,6	1,1
7	33,6	31,5	2,1	28,7	4,9
8	36,9	36,5	0,4	36,2	0,7
9	56,7	55,1	1,6	43,1	13,6
10	36,6	36,0	0,6	36,2	0,4
11	21,3	21,8	0,5	22,7	1,4
12	40,4	40,8	0,4	40,8	0,4
13	43,5	43,6	0,1	43,6	0,1
14	57,9	57,8	0,1	57,8	0,1
15	25,1	25,1	0,0	25,1	0,0
16	28,6	29,3	0,7	29,4	0,8
17	51,9	51,1	0,8	41,5	10,4
18	12,7	12,7	0,0	12,7	0,0
19	14,8	15,0	0,2	24,0	9,2
20	41,4	41,5	0,1	41,5	0,1
21	69,4	70,9	1,5	73,9	4,5
22	30,8	31,8	1,0	31,2	0,4
23	24,4	24,5	0,1	24,4	0,0
24	40,7	40,5	0,2	40,6	0,1
25	29,1	29,0	0,1	28,9	0,2
26	23,0	23,0	0,0	22,7	0,3
27	23,2	23,0	0,2	22,7	0,5
Maior erro observado			2,10		13,60
Erro médio			0,51		2,53
Desvio padrão			0,53		3,82

7

Esqueleto: Bufanolídeo - Ranunculaceae - *Helleborus orientalis*¹⁵⁴

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	18,7	20,2	1,5	25,0	6,3
2	25,6	25,8	0,2	25,6	0,0
3	73,1	73,5	0,4	75,0	1,9
4	35,4	35,5	0,1	35,6	0,2
5	74,0	74,2	0,2	77,0	3,0
6	37,0	37,2	0,2	37,2	0,2
7	25,0	24,0	1,0	21,6	1,6
8	42,3	41,1	1,2	38,6	3,7
9	39,6	39,4	0,2	39,4	0,2
10	55,3	55,5	0,2	55,6	0,3
11	22,8	22,3	0,5	18,3	4,5
12	41,0	40,3	0,7	39,1	1,9
13	49,4	48,0	1,4	43,5	5,9
14	84,3	84,5	0,2	84,5	0,2
15	42,3	42,0	0,3	41,7	0,6
16	72,4	71,7	0,7	70,4	2,0
17	58,9	58,6	0,3	58,0	0,9
18	17,1	17,8	0,7	18,5	1,4
19	208,4	209,7	1,3	210,3	1,9
20	119,2	119,4	0,2	119,5	0,3
21	150,7	150,1	0,6	149,7	1,0
22	151,3	150,8	0,5	146,1	5,2
23	112,6	113,3	0,7	114,9	2,3
24	162,2	162,2	0,0	162,2	0,0
Rha-1'	100,6	101,0	0,4	101,5	0,9
2'	72,5	72,4	0,1	70,8	1,7
3'	72,9	72,5	0,4	70,8	2,1
4'	70,8	71,5	0,7	73,6	2,8
5'	73,6	72,7	0,9	68,3	5,3
6'	18,6	18,3	0,3	18,0	0,6
Maior erro observado			1,50		6,30
Erro médio			0,54		1,96
Desvio padrão			0,41		1,86

8

Esqueleto: Colestano - Liliaceae - *Tupistra chinensis*¹⁵⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	74,6	74,0	0,6	73,8	0,8
2	40,0	39,3	0,7	38,3	1,7
3	68,4	67,2	1,2	65,4	3,0
4	37,1	37,5	0,4	39,6	2,5
5	75,9	75,7	0,2	75,8	0,1
6	35,6	35,0	0,6	34,4	1,2
7	30,3	28,7	1,6	27,4	2,9
8	34,9	35,3	0,4	36,1	1,2
9	46,4	46,0	0,4	45,4	1,0
10	44,6	42,6	2,0	38,3	6,3
11	22,2	22,0	0,2	21,3	0,9
12	40,7	40,0	0,7	39,8	0,9
13	41,3	41,0	0,3	40,7	0,6
14	56,9	56,5	0,4	56,4	0,5
15	32,8	32,0	0,8	31,9	0,9
16	82,2	81,7	0,5	80,8	1,4
17	63,6	63,0	0,6	62,5	1,1
18	17,2	16,6	0,6	16,3	0,9
19	14,4	15,4	1,0	16,8	2,4
20	42,6	42,0	0,6	41,6	1,0
21	15,7	15,0	0,7	14,4	1,3
22	110,2	109,5	0,7	109,3	0,9
23	33,8	33,0	0,8	30,2	3,6
24	29,6	31,2	1,6	32,3	2,7
25	144,9	144,0	0,9	143,5	1,4
26	65,7	65,1	0,6	64,9	0,8
27	109,6	109,3	0,3	108,5	1,1
Maior erro observado			2,00		6,30
Erro médio			0,72		1,60
Desvio padrão			0,44		1,28

9

Esqueleto: Pregnano - Liliaceae - *Tupistra chinensis*¹⁵⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	75,1	76,3	1,2	78,3	3,2
2	40,8	42,4	1,6	44,0	3,2
3	66,5	67,2	0,7	68,2	1,7
4	41,1	42,0	0,9	43,6	2,5
5	140,5	140,5	0,0	140,5	0,0
6	124,3	125,0	0,7	125,4	0,1
7	31,9	31,8	0,1	31,7	0,2
8	31,5	31,0	0,5	30,1	1,4
9	51,8	51,5	0,3	51,4	0,4
10	46,0	44,9	1,1	44,6	1,4
11	24,3	23,3	1,0	20,7	3,6
12	35,8	35,0	0,8	34,7	1,1
13	44,7	45,2	0,5	46,1	1,4
14	56,8	56,3	0,5	56,4	0,4
15	32,5	32,4	0,1	32,2	0,3
16	144,5	144,5	0,0	144,4	0,1
17	155,6	155,5	0,1	155,3	0,3
18	16,2	16,0	0,2	15,7	0,5
19	13,2	13,6	0,4	14,0	0,8
20	196,3	196,5	0,2	196,6	0,3
21	27,0	27,0	0,0	27,1	0,1
Maior erro observado			1,60		3,60
Erro médio			0,52		1,10
Desvio padrão			0,45		1,14

Esqueleto: Colestano - Caryophyllaceae - *Silene italica* ssp. *nemoralis*¹⁵⁶

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	37,6	38,0	0,4	38,4	0,8
2	69,7	69,0	0,7	68,9	0,8
3	69,8	69,3	0,5	67,9	1,9
4	36,3	35,0	1,3	32,2	4,1
5	51,4	50,1	1,3	49,7	1,7
6	206,7	205,0	1,7	202,4	4,3
7	123,8	123,1	0,7	122,3	1,5
8	162,6	163,8	1,2	166,3	3,7
9	76,6	76,9	0,3	76,6	0,0
10	43,0	43,1	0,1	42,6	0,4
11	30,9	31,4	0,5	32,6	1,7
12	29,6	30,5	0,9	32,5	2,9
13	49,1	50,0	0,9	50,9	1,8
14	87,2	86,3	0,9	85,0	2,2
15	31,6	31,8	0,2	31,8	0,2
16	21,4	21,6	0,2	21,6	0,2
17	50,4	50,2	0,2	50,1	0,3
18	18,5	19,3	0,8	20,0	1,5
19	29,6	28,1	1,5	17,4	12,2
20	78,0	77,7	0,3	76,9	1,1
21	21,2	21,2	0,0	21,3	0,1
22	78,6	78,1	0,5	77,5	1,1
23	27,5	27,5	0,0	27,2	0,3
24	42,5	42,5	0,0	42,3	0,2
25	71,4	70,7	0,7	69,9	1,5
26	29,1	29,5	0,4	29,6	0,5
27	29,9	29,5	0,4	29,6	0,3
Maior erro observado			1,70		12,20
Erro médio			0,61		1,75
Desvio padrão			0,47		2,42

Esqueleto: Ergostano - Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	39,7	39,7	0,0	39,7	0,0
2	26,2	26,5	0,3	26,8	0,6
3	73,7	73,5	0,2	72,8	0,9
4	69,1	69,4	0,3	71,5	2,4
5	57,3	57,0	0,3	56,0	1,3
6	64,8	65,9	1,1	67,8	3,0
7	49,7	49,0	0,7	48,2	1,5
8	77,4	77,4	0,0	77,2	0,2
9	58,4	57,9	0,5	57,7	0,7
10	38,1	37,5	0,6	37,3	0,8
11	19,2	18,8	0,4	18,8	0,4
12	43,3	42,6	0,7	42,2	1,1
13	44,4	44,0	0,4	44,0	0,4
14	62,7	62,9	0,2	62,9	0,2
15	71,1	70,2	0,9	69,9	1,2
16	42,5	42,3	0,2	42,2	0,3
17	57,8	57,4	0,4	57,1	0,7
18	16,5	16,3	0,2	16,3	0,2
19	17,0	17,5	0,5	17,5	0,5
20	36,0	36,2	0,2	36,2	0,2
21	18,9	18,5	0,4	18,3	0,6
22	35,5	34,8	0,7	34,8	0,7
23	32,7	31,4	1,3	31,1	1,6
24	153,8	154,2	0,4	154,3	0,5
25	43,3	40,7	2,6	39,7	3,6
26	67,5	66,9	0,6	66,1	1,4
27	17,2	16,9	0,3	16,9	0,3
28	109,4	109,5	0,1	109,9	0,5
Maior erro observado			1,30		3,60
Erro médio			0,52		0,92
Desvio padrão			0,51		0,86

CC(C)C(O)C=CC[C@H]1[C@@H](O)[C@H]2[C@@H](O)[C@H]3[C@H]1CC[C@@H](O)C4=CC(=O)CC[C@]34C

Esqueleto: Colestano - Estrela-do-mar - *Certnardoia semiregularis*¹⁵⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,6	36,4	2,2	35,7	2,9
2	32,0	31,3	0,7	30,8	1,2
3	72,0	71,5	0,5	71,4	0,6
4	33,0	33,3	0,3	33,8	0,8
5	53,1	52,7	0,4	50,0	3,1
6	70,0	70,6	0,6	71,0	1,0
7	41,8	41,1	0,7	40,0	1,8
8	31,6	30,3	1,3	28,5	3,1
9	55,7	54,7	1,0	52,0	3,7
10	37,5	37,0	0,5	36,3	1,2
11	22,2	21,8	0,4	21,3	0,9
12	42,5	41,9	0,6	41,7	0,8
13	43,4	44,2	0,8	45,0	1,6
14	62,2	63,0	0,8	63,9	1,7
15	70,6	72,1	1,5	74,4	3,8
16	42,8	42,0	0,8	41,9	0,9
17	57,5	56,3	1,2	54,5	3,0
18	15,3	15,0	0,3	14,1	1,2
19	13,8	14,6	0,8	15,8	2,0
20	41,1	39,6	1,5	35,7	5,4
21	21,2	20,9	1,3	20,2	1,0
22	139,2	138,7	0,5	136,6	2,6
23	130,1	131,4	1,3	134,3	4,2
24	78,9	78,0	0,9	77,0	1,9
25	35,3	35,2	0,1	35,2	0,1
26	18,8	18,4	0,4	18,3	0,5
27	18,7	18,4	0,3	18,1	0,6
Maior erro observado			2,20		5,40
Erro médio			0,80		1,91
Desvio padrão			0,48		1,33

13

Esqueleto: Colestano - Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,6	36,4	2,2	35,7	2,9
2	32,0	31,3	0,7	30,8	1,2
3	72,0	71,5	0,5	71,4	0,6
4	33,0	33,3	0,3	33,8	0,8
5	53,1	52,7	0,4	50,0	3,1
6	70,0	70,6	0,6	71,0	1,0
7	41,8	41,1	0,7	40,0	1,8
8	31,6	30,3	1,3	28,5	3,1
9	55,7	54,7	1,0	52,0	3,7
10	37,5	37,0	0,5	36,3	1,2
11	22,2	21,8	0,4	21,3	0,9
12	42,5	41,9	0,6	41,5	1,0
13	43,4	44,0	0,6	45,1	1,7
14	62,2	63,2	1,0	64,7	2,5
15	70,6	72,3	1,7	75,4	4,8
16	42,3	41,1	1,2	39,6	2,7
17	57,8	56,0	1,8	55,3	2,5
18	15,2	15,2	0,0	13,9	1,3
19	13,8	14,6	0,8	15,8	2,0
20	37,1	38,0	0,9	35,9	1,2
21	19,4	18,9	0,5	18,9	0,5
22	33,5	33,0	0,5	33,0	0,5
23	31,7	31,7	0,0	31,7	0,0
24	78,1	77,6	0,5	77,0	1,1
25	34,5	34,4	0,1	34,2	0,3
26	19,5	19,5	0,0	19,7	0,2
27	17,5	17,7	0,2	17,5	0,0
Maior erro observado			2,20		4,80
Erro médio			0,70		1,58
Desvio padrão			0,55		1,22

14

Esqueleto: 23-Metilergostano - Estrela-do-mar - *Certonardoa semiregularis*¹⁵⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	39,4	39,2	0,2	39,2	0,2
2	31,5	31,9	0,4	31,9	0,4
3	72,2	71,6	0,6	71,6	0,6
4	32,4	33,0	0,6	33,1	0,7
5	53,9	53,7	0,2	53,7	0,2
6	67,6	67,0	0,6	66,5	1,1
7	49,5	50,1	0,6	50,8	1,3
8	77,5	77,0	0,5	76,8	0,7
9	57,4	57,9	0,5	57,9	0,5
10	38,0	37,5	0,5	37,2	0,8
11	19,7	19,0	0,7	18,8	0,9
12	43,3	41,4	1,9	39,6	3,7
13	44,2	45,0	0,8	45,1	0,9
14	62,7	63,0	0,3	63,2	0,5
15	71,2	70,9	0,3	70,9	0,3
16	42,8	42,0	0,8	41,8	1,0
17	58,5	57,2	1,3	55,5	3,0
18	16,8	15,9	0,9	15,7	1,1
19	14,1	14,4	0,3	14,4	0,3
20	35,3	34,7	0,6	27,4	7,9
21	20,7	20,5	0,2	20,5	0,2
22	132,9	131,1	1,8	130,7	2,2
23	136,2	137,0	0,8	138,5	2,3
24	45,9	47,2	1,3	49,4	3,5
25	39,2	38,7	0,5	38,4	0,8
26	67,2	66,8	0,4	66,2	1,0
27	14,3	15,3	1,0	17,3	3,0
28	14,0	15,1	1,1	15,1	1,1
29	16,4	16,0	0,4	15,4	1,0
Maior erro observado			1,90		7,90
Erro médio			0,69		1,42
Desvio padrão			0,44		1,60

15

Esqueleto: Estigmastano - Verbenaceae - *Verbena littoralis*¹⁵⁸

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	36,9	36,5	0,4	36,0	0,9
2	24,4	24,6	0,2	25,2	0,8
3	72,2	73,0	0,8	74,0	1,8
4	77,1	76,7	0,4	74,6	2,5
5	146,6	146,4	0,2	146,4	0,2
6	128,6	128,1	0,5	124,0	4,6
7	64,6	64,9	0,3	72,7	8,1
8	37,8	39,0	1,2	40,2	2,4
9	42,5	43,3	0,8	46,6	4,1
10	36,8	35,1	1,7	32,3	4,5
11	23,9	23,0	0,9	21,3	2,6
12	39,2	39,7	0,5	39,7	0,5
13	42,2	44,0	1,8	45,8	3,6
14	49,0	51,0	2,0	55,0	6,0
15	19,9	22,0	2,1	26,1	6,2
16	27,5	26,5	1,0	26,4	1,1
17	53,0	52,7	0,3	52,9	0,1
18	10,7	12,0	1,3	12,2	1,5
19	18,6	19,4	0,8	20,7	2,1
20	42,9	43,5	0,6	44,2	1,3
21	11,6	12,9	1,3	13,1	1,5
22	70,5	71,8	1,3	74,3	3,8
23	29,4	31,3	1,9	35,3	5,9
24	41,1	42,7	1,6	42,9	1,8
25	28,8	30,0	1,2	31,4	2,6
26	16,9	19,4	2,5	19,8	2,9
27	19,4	19,4	0,0	19,4	0,0
28	22,8	23,2	0,4	25,0	2,2
29	10,8	12,1	1,3	12,7	1,9
Maior erro observado			2,50		8,10
Erro médio			1,01		2,67
Desvio padrão			0,66		2,03

16

Esqueleto: 23-Metilergostano - Coral - *Isis hippuris*¹⁵⁹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,1	37,9	0,2	37,9	0,2
2	38,0	38,0	0,0	37,7	0,3
3	211,8	211,9	0,1	211,9	0,1
4	44,5	44,0	0,5	43,6	0,9
5	46,5	46,1	0,4	45,7	0,8
6	28,6	27,8	0,8	27,3	1,3
7	31,5	31,0	0,5	30,7	0,8
8	33,8	34,2	0,4	34,2	0,4
9	53,4	54,3	0,9	54,5	1,1
10	35,7	35,4	0,3	35,4	0,3
11	20,9	21,3	0,4	21,3	0,4
12	39,3	39,3	0,0	39,3	0,0
13	43,6	43,6	0,0	43,6	0,0
14	50,9	51,7	0,8	52,3	1,4
15	38,8	39,5	0,7	39,7	0,9
16	218,8	218,5	0,3	218,0	0,8
17	67,4	67,0	0,4	66,4	1,0
18	14,4	14,7	0,3	15,9	1,5
19	11,4	13,2	1,8	16,4	5,0
20	77,1	76,8	0,3	72,1	5,0
21	20,6	21,3	0,7	22,0	1,4
22	81,1	82,0	0,9	82,0	0,9
23	32,3	33,6	1,3	34,7	2,4
24	44,5	42,2	2,3	38,9	5,6
25	85,7	84,7	1,0	83,4	2,3
26	23,6	25,1	1,5	28,8	5,2
27	25,1	25,1	0,0	25,2	0,1
28	9,5	11,0	1,5	13,2	3,7
29	10,7	11,9	1,2	15,9	5,2
Ac	171,4	171,6	0,2	171,6	0,2
	21,1	21,3	0,2	21,3	0,2
Ac	170,3	170,0	0,3	170,0	0,3
	22,7	22,5	0,2	22,5	0,2
Maior erro observado			2,30		5,60
Erro médio			0,62		1,51
Desvio padrão			0,56		1,77

17

Esqueleto: Ergostano - Coral - *Sinularia gibberosa*¹⁶⁰

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	70,7	71,4	0,7	72,8	2,1
2	37,5	38,0	0,5	38,2	0,7
3	67,7	67,1	0,6	64,4	3,3
4	37,3	38,3	0,0	38,6	1,3
5	83,7	82,5	1,2	82,8	0,9
6	211,5	211,9	0,4	212,4	0,9
7	41,2	39,7	1,5	35,2	6,0
8	37,4	37,0	0,4	36,7	0,7
9	43,3	44,4	1,1	47,2	3,9
10	49,1	48,6	0,5	38,2	10,9
11	23,6	23,0	0,6	21,4	2,2
12	39,7	41,3	1,6	41,3	1,6
13	42,4	43,0	0,6	43,0	0,6
14	56,8	56,7	0,1	56,7	0,1
15	24,2	23,9	0,3	23,9	0,3
16	27,7	28,2	0,5	28,2	0,5
17	56,0	56,0	0,0	56,0	0,0
18	11,9	12,0	0,1	12,0	0,1
19	13,4	15,1	1,7	17,3	3,9
20	35,6	35,7	0,1	35,7	0,1
21	18,5	18,9	0,4	18,8	0,3
22	34,5	34,8	0,3	34,8	0,3
23	30,9	31,3	0,4	31,1	0,2
24	156,7	156,8	0,1	156,8	0,1
25	33,8	33,9	0,1	33,9	0,1
26	21,8	22,0	0,2	22,0	0,2
27	22,0	22,0	0,0	22,0	0,0
28	106,0	106,3	0,3	106,1	0,1
Maior erro observado			1,70		10,90
Erro médio			0,51		1,48
Desvio padrão			0,49		2,37

18

Esqueleto: Ergostano - Fungo - *Gymnascella dankaliensis*¹⁶¹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	32,6	32,1	0,5	31,6	1,0
2	29,6	29,6	0,0	29,6	0,0
3	65,6	66,0	0,4	66,6	1,0
4	121,9	121,5	0,4	121,0	0,9
5	146,2	145,7	0,5	142,3	3,9
6	66,5	65,8	0,7	64,9	1,6
7	124,1	122,8	1,3	115,6	8,5
8	136,0	138,3	2,3	143,9	7,9
9	48,2	45,1	3,1	42,0	6,2
10	38,2	35,9	2,3	35,4	2,8
11	21,3	21,3	0,0	21,3	0,0
12	40,1	38,9	1,2	30,6	9,5
13	45,3	47,1	1,8	48,3	3,0
14	72,7	73,2	0,5	74,9	2,2
15	67,9	66,5	1,4	62,0	5,9
16	29,6	30,3	0,7	32,9	3,3
17	53,6	52,4	1,2	51,9	1,7
18	15,9	15,0	0,9	14,5	1,4
19	22,3	21,1	1,2	18,0	4,3
20	39,0	40,1	1,1	40,3	1,3
21	23,4	21,4	2,0	20,3	3,1
22	135,5	134,6	0,9	134,1	1,4
23	132,9	133,4	0,5	133,2	0,3
24	44,0	42,7	1,3	40,1	3,9
25	33,9	33,1	0,8	33,2	0,7
26	20,1	19,7	0,4	19,8	0,3
27	20,4	19,7	0,7	20,3	0,1
28	18,1	17,8	0,3	17,8	0,3
Maior erro observado			3,10		9,50
Erro médio			1,01		2,73
Desvio padrão			0,74		2,69

19

Esqueleto: Vitanolídeo - Solanaceae – *Physalis philadelphica*¹⁶²

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	202,1	202,8	0,7	203,5	1,4
2	129,3	130,2	0,9	133,9	4,6
3	142,4	143,0	0,6	143,9	1,5
4	69,1	68,5	0,6	68,0	1,1
5	143,0	141,8	1,2	139,6	3,4
6	131,4	130,2	1,2	128,8	2,6
7	64,4	66,7	2,3	75,3	10,9
8	37,2	38,5	1,3	40,3	3,1
9	35,0	37,1	2,1	41,1	6,1
10	50,0	49,3	0,7	49,0	1,0
11	21,7	22,0	0,3	22,2	0,5
12	39,7	40,0	0,3	40,0	0,3
13	42,9	43,2	0,3	43,5	0,6
14	50,0	51,4	1,4	56,3	6,3
15	23,8	24,6	0,8	26,2	2,4
16	22,6	22,4	0,2	22,4	0,2
17	55,4	54,9	0,5	54,9	0,5
18	13,4	13,8	0,4	13,9	0,5
19	21,7	21,0	0,7	20,7	1,0
20	75,5	76,0	0,5	76,3	0,8
21	20,7	20,9	0,2	20,9	0,2
22	80,4	81,1	0,7	82,1	1,7
23	31,2	31,1	0,1	31,1	0,1
24	31,3	31,3	0,0	31,3	0,0
25	40,5	40,7	0,2	40,7	0,2
26	176,0	175,2	0,8	175,1	0,9
27	14,2	14,1	0,1	14,1	0,1
28	21,4	21,1	0,3	21,1	0,3
Maior erro observado			2,30		10,90
Erro médio			0,69		1,87
Desvio padrão			0,57		2,48

Esqueleto: Colestano - Liliaceae – *Tupistra wattii*¹⁶³

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	73,4	73,0	0,4	73,0	0,4
2	33,0	34,2	1,2	37,3	4,3
3	68,2	67,8	0,4	67,3	0,9
4	34,4	35,0	0,6	35,9	1,5
5	31,3	32,3	1,0	34,8	3,5
6	26,7	27,5	0,8	28,5	1,8
7	26,6	28,6	2,0	30,6	4,0
8	35,7	36,0	0,3	36,2	0,5
9	42,2	43,5	1,3	52,4	10,2
10	40,4	40,2	0,2	40,1	0,3
11	21,1	21,1	0,0	21,1	0,0
12	40,7	40,0	0,7	39,8	0,9
13	40,4	40,5	0,1	40,7	0,3
14	56,5	56,4	0,1	56,4	0,1
15	32,1	31,9	0,2	31,9	0,2
16	81,6	81,4	0,2	81,4	0,2
17	62,2	62,5	0,3	62,5	0,3
18	16,8	16,5	0,3	16,5	0,3
19	19,4	17,0	2,4	12,7	6,7
20	42,2	42,3	0,1	42,3	0,1
21	14,9	15,0	0,1	15,0	0,1
22	111,8	111,8	0,0	111,8	0,0
23	40,7	41,0	0,3	41,0	0,3
24	81,6	81,5	0,1	81,5	0,1
25	38,2	38,3	0,1	38,3	0,1
26	65,1	65,4	0,3	65,4	0,3
27	13,6	13,6	0,0	13,6	0,0
Glc-1'	106,1; 105,6	106,0; 105,5	0,1; 0,1	104,4; 105,5	1,7; 0,1
2'	75,1; 75,5	75,0; 75,0	0,1; 0,5	75,0; 75,0	0,1; 0,5
3'	78,5; 78,4	78,3; 78,3	0,2; 0,1	76,7; 77,7	1,8; 0,7
4'	71,3; 71,6	71,5; 71,5	0,2; 0,1	71,8; 71,7	0,5; 0,1
5'	76,7; 78,5	77,0; 78,1	0,3; 0,4	77,2; 78,5	0,5; 0,0
6'	70,0; 62,8	70,3; 62,8	0,3; 0,0	70,5; 62,7	0,5; 0,1
Maior erro observado			2,40		10,20
Erro médio			0,41		1,13
Desvio padrão			0,53		2,06

21

Esqueleto: Ergostano - Fungo – *Byssoschlamys nivea*¹⁶⁴

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	35,7	36,1	0,4	36,6	0,9
2	26,4	26,7	0,3	26,9	0,5
3	72,2	72,0	0,2	70,9	1,3
4	25,5	26,2	0,7	27,0	1,5
5	54,4	53,4	1,0	52,6	1,8
6	209,3	209,5	0,2	209,5	0,2
7	78,8	78,6	0,2	78,6	0,2
8	42,4	40,9	1,5	37,1	5,3
9	55,2	54,3	0,9	51,5	3,7
10	40,6	39,8	0,8	39,8	0,8
11	68,3	68,9	0,6	69,1	0,8
12	48,8	49,3	0,5	49,3	0,5
13	42,9	43,3	0,4	48,2	5,3
14	56,0	55,4	0,6	51,2	4,8
15	36,4	35,7	0,7	32,7	3,7
16	74,6	75,1	0,5	75,7	1,1
17	60,1	61,0	0,9	61,5	1,4
18	15,4	15,6	0,2	15,6	0,2
19	15,5	16,5	1,0	16,5	1,0
20	34,3	32,5	1,8	29,6	4,7
21	21,1	21,9	0,8	22,1	1,0
22	134,9	134,4	0,5	134,4	0,5
23	133,0	133,3	0,3	133,1	0,1
24	43,2	42,8	0,4	37,7	5,5
25	33,0	33,2	0,2	33,2	0,2
26	19,6	19,9	0,3	19,9	0,3
27	20,1	20,4	0,3	20,4	0,3
28	18,0	17,8	0,2	17,8	0,2
Ac	170,6	171,1	0,5	171,3	0,7
	21,3	21,3	0,0	21,3	0,0
	170,0	170,4	0,4	171,9	1,9
	21,5	21,4	0,1	21,6	0,1
Maior erro observado			1,80		5,50
Erro médio			0,54		1,58
Desvio padrão			0,40		1,79

22

Esqueleto: Vitanolideo - Solanaceae – *Withania somnifera*¹⁶⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	204,2	204,3	0,1	204,1	0,1
2	127,7	128,2	0,5	128,7	1,0
3	142,1	141,5	0,6	141,3	1,2
4	31,9	33,0	1,1	36,0	4,1
5	78,0	77,8	0,2	77,2	0,8
6	73,4	74,0	0,6	74,3	0,9
7	30,7	29,3	1,4	26,7	4,0
8	35,4	35,8	0,4	35,5	0,1
9	35,8	35,2	0,6	34,6	1,2
10	54,2	53,7	0,5	52,2	2,0
11	22,2	22,5	0,3	22,8	0,6
12	33,6	34,0	0,4	34,6	1,0
13	51,3	53,0	1,7	54,4	3,1
14	82,5	82,1	0,4	82,4	0,1
15	28,9	29,7	0,8	30,5	1,6
16	34,4	36,6	2,2	37,0	2,6
17	87,5	88,1	0,6	88,3	0,8
18	20,8	20,4	0,4	20,1	0,7
19	9,8	10,3	0,5	15,1	5,3
20	76,3	77,1	0,8	79,0	2,7
21	19,3	20,0	0,7	20,1	0,8
22	81,3	81,4	0,1	81,4	0,1
23	33,1	32,5	0,6	32,7	0,4
24	150,8	151,3	0,5	151,4	0,6
25	120,1	120,9	0,8	121,4	1,3
26	166,1	166,7	0,6	167,0	0,9
27	12,1	12,5	0,4	12,4	0,3
28	20,3	20,5	0,2	20,5	0,2
Maior erro observado			2,20		5,30
Erro médio			0,64		1,38
Desvio padrão			0,47		1,36

23

Esqueleto: Colestano - Alga – *Hypnea musciformis*¹⁶⁶

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,2	37,9	0,3	37,7	0,5
2	37,8	37,6	0,2	37,6	0,2
3	211,3	210,8	0,5	208,0	3,3
4	37,6	37,3	0,3	36,8	0,8
5	57,5	57,3	0,2	57,4	0,1
6	209,2	210,2	1,0	210,7	1,5
7	46,4	45,9	0,5	45,7	1,3
8	37,5	37,0	0,5	36,9	0,6
9	53,6	54,1	0,5	55,5	1,9
10	41,5	41,1	0,4	41,0	0,5
11	21,8	21,5	0,3	21,3	0,5
12	40,0	39,8	0,2	39,9	0,1
13	43,5	45,0	1,5	47,0	3,5
14	56,9	56,7	0,2	56,8	0,1
15	23,8	24,0	0,2	23,9	0,1
16	29,5	28,4	1,1	24,0	5,5
17	60,1	60,6	0,5	61,7	1,6
18	13,9	15,2	1,3	16,3	2,4
19	12,6	12,3	0,3	12,5	0,1
20	75,2	76,1	0,9	76,7	1,5
21	29,5	29,2	0,3	29,4	0,1
22	139,3	139,1	0,2	138,6	0,7
23	125,6	125,3	0,3	125,4	0,2
24	42,1	42,0	0,1	42,0	0,1
25	28,7	29,0	0,3	29,0	0,3
26	22,5	22,5	0,0	22,5	0,0
27	22,4	22,5	0,1	22,5	0,1
Maior erro observado			1,50		5,50
Erro médio			0,45		0,99
Desvio padrão			0,38		1,33

Esqueleto: Colestano - Alga – *Hypnea musciformis*¹⁶⁶

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	37,1	37,1	0,0	37,1	0,0
2	34,3	34,0	0,3	34,2	0,1
3	200,1	200,0	0,1	200,1	0,0
4	126,4	126,1	0,3	126,0	0,4
5	168,0	168,3	0,3	168,5	0,5
6	73,3	73,0	0,3	73,1	0,2
7	38,5	38,7	0,2	38,6	0,1
8	29,7	29,9	0,2	29,8	0,1
9	53,6	53,7	0,1	53,7	0,1
10	37,9	38,0	0,1	37,9	0,0
11	21,0	21,0	0,0	21,0	0,0
12	39,4	39,7	0,3	39,4	0,0
13	42,2	42,7	0,5	42,5	0,3
14	56,2	56,0	0,2	55,9	0,3
15	24,8	24,2	0,6	24,1	0,7
16	28,5	28,0	0,5	28,1	0,4
17	55,9	56,3	0,4	56,1	0,2
18	12,0	12,0	0,0	12,0	0,0
19	19,5	19,3	0,2	19,4	0,1
20	32,8	34,0	1,2	35,7	2,9
21	18,7	18,9	0,2	18,7	0,0
22	37,3	36,6	0,7	36,1	1,2
23	24,5	23,4	1,1	23,8	0,7
24	39,3	39,8	0,5	39,6	0,3
25	28,1	27,6	0,5	27,9	0,2
26	22,5	22,8	0,3	22,7	0,2
27	22,5	22,8	0,3	22,7	0,2
Maior erro observado			1,20		2,90
Erro médio			0,35		0,34
Desvio padrão			0,29		0,58

25

Esqueleto: Bufanolídeo - Apocynaceae – *Catharanthus roseus*¹⁶⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	30,6	30,5	0,1	30,2	0,4
2	27,0	27,0	0,0	27,0	0,0
3	74,2	74,0	0,2	72,2	2,0
4	30,3	30,5	0,2	30,5	0,2
5	36,5	36,3	0,2	36,3	0,2
6	26,2	26,9	0,7	26,9	0,7
7	21,2	20,7	0,5	20,8	0,4
8	35,5	34,8	0,7	33,4	2,1
9	39,8	39,0	0,8	38,0	1,8
10	33,4	34,6	1,2	35,4	2,0
11	21,0	21,5	0,5	21,3	0,3
12	39,6	38,5	1,1	38,6	1,0
13	45,3	46,4	1,1	46,8	1,5
14	72,5	70,9	1,6	69,9	2,6
15	59,8	60,2	0,4	62,3	2,5
16	75,4	75,9	0,5	78,1	2,7
17	50,5	51,7	1,2	55,2	4,7
18	17,2	17,4	0,2	18,1	0,9
19	23,6	23,9	0,3	23,7	0,1
20	116,7	116,3	0,4	115,5	1,2
21	152,6	152,2	0,4	151,5	1,1
22	148,9	147,7	1,2	143,5	5,4
23	113,9	114,6	0,7	118,3	4,4
24	161,4	161,8	0,4	162,2	0,8
Glc-1'	103,2	103,1	0,1	102,6	0,6
2'	75,4	75,3	0,1	75,3	0,1
3'	78,8	78,6	0,2	78,6	0,2
4'	71,9	72,3	0,4	72,3	0,4
5'	78,4	78,0	0,4	77,9	0,5
6'	63,0	63,2	0,2	63,3	0,3
Ac	169,9	170,1	0,2	171,3	0,4
	20,2	21,0	0,8	20,9	0,7
Maior erro observado			1,60		5,40
Erro médio			0,53		1,32
Desvio padrão			0,41		1,40

26

Esqueleto: Colano – Lampréia – *Petromyzon marinus*¹⁶⁸

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	22,7	25,2	2,5	37,6	14,9
2	44,0	38,4	1,6	37,5	6,5
3	210,6	211,7	1,1	213,2	2,6
4	28,8	30,8	2,0	45,8	17,0
5	41,4	42,0	0,6	42,3	0,9
6	36,8	35,4	1,4	33,9	2,9
7	65,4	66,6	1,2	70,3	4,9
8	46,0	44,7	1,3	39,8	6,2
9	38,9	37,5	1,4	27,8	11,1
10	34,9	34,2	0,7	33,8	1,1
11	27,1	27,6	0,5	28,4	1,3
12	70,8	71,4	0,6	73,6	2,8
13	45,8	46,3	0,5	46,8	1,0
14	38,2	39,0	0,8	41,5	3,3
15	30,8	29,9	0,9	23,3	7,5
16	30,7	29,3	1,4	28,0	2,7
17	39,5	41,0	1,5	47,4	7,9
18	10,0	11,2	1,2	12,8	2,8
19	12,3	13,0	0,7	17,5	5,2
20	35,0	35,6	0,6	35,7	0,7
21	16,9	17,2	0,3	17,6	0,7
22	37,9	36,8	1,1	31,3	6,6
23	37,7	35,9	1,8	31,2	6,5
24	174,8	174,5	0,3	174,5	0,3
Maior erro observado			2,50		17,00
Erro médio			1,08		4,89
Desvio padrão			0,56		4,43

Esqueleto: Estigmastano - Lamiaceae - *Ajuga salicifolia*¹⁶⁹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,5	38,0	0,5	37,7	0,8
2	30,4	30,1	0,3	30,2	0,2
3	79,0	78,4	0,6	77,9	1,1
4	35,2	35,1	0,1	35,1	0,1
5	41,4	40,7	0,7	40,4	1,0
6	30,9	30,0	0,9	30,0	0,9
7	118,9	118,4	0,5	118,4	0,5
8	140,7	140,9	0,2	140,9	0,2
9	50,8	50,0	0,8	45,4	5,4
10	35,5	34,5	1,0	34,3	1,2
11	22,2	21,5	0,7	21,3	0,9
12	34,0	34,9	0,9	35,5	1,5
13	49,5	48,6	0,9	45,1	4,4
14	49,0	49,7	0,7	49,4	0,4
15	33,6	34,6	1,0	34,6	1,0
16	82,2	80,3	1,9	75,6	6,6
17	87,8	88,4	0,6	88,3	0,5
18	13,6	14,7	1,1	16,3	2,7
19	13,5	12,6	0,9	12,9	0,6
20	35,3	36,4	1,1	36,6	1,3
21	12,8	13,1	0,3	14,9	2,1
22	81,8	82,3	0,5	84,2	2,4
23	78,5	76,6	1,9	75,5	3,0
24	56,2	54,9	1,3	48,0	8,2
25	84,1	84,8	0,7	86,0	1,9
26	18,0	19,3	1,3	23,6	5,6
27	72,0	71,4	0,6	66,7	5,3
28	22,9	22,2	0,7	18,7	4,2
29	13,8	13,0	0,8	13,0	0,8
Glc-1'	102,3	102,5	0,2	102,6	0,3
2'	75,1	75,3	0,2	75,3	0,2
3'	78,1	78,3	0,2	78,6	0,5
4'	71,7	72,1	0,4	72,3	0,6
5'	77,9	77,9	0,0	77,9	0,0
6'	62,8	63,3	0,5	63,3	0,5
Ac	172,7	171,1	1,6	170,7	2,0
	20,7	21,0	0,3	20,7	0,3
Maior erro observado			1,90		8,20
Erro médio			0,73		1,87
Desvio padrão			0,46		2,07

28

Esqueleto: Colestano - Zoanthidae – *Zoanthus* sp.¹⁷⁰

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	77,7	77,3	0,4	75,6	2,1
2	69,9	70,0	0,1	75,5	5,6
3	72,1	72,8	0,7	75,9	3,8
4	33,3	33,0	0,3	32,5	0,8
5	48,0	48,5	0,5	49,7	1,7
6	204,8	204,7	0,1	202,4	2,4
7	123,3	123,3	0,0	121,4	1,9
8	167,2	167,0	0,2	166,3	0,9
9	36,4	36,9	0,5	43,0	6,6
10	45,0	44,3	0,7	40,4	4,6
11	---	21,5		21,3	
12	33,6	32,1	1,5	30,7	2,9
13	49,3	49,7	0,4	48,2	1,1
14	85,7	85,7	0,0	85,0	0,7
15	31,4	31,8	0,4	31,8	0,4
16	---	23,7		21,4	
17	51,7	51,1	0,6	50,0	1,7
18	19,5	19,1	0,4	19,2	0,3
19	22,9	22,7	0,2	22,0	0,9
20	78,4	77,9	0,5	76,7	1,7
21	23,0	22,1	0,9	21,1	1,9
22	78,4	77,6	0,8	76,7	1,7
23	29,8	30,3	0,5	30,2	0,4
24	38,7	38,0	0,7	37,1	1,6
25	31,4	30,3	1,1	28,1	3,3
26	24,0	24,5	0,5	22,3	1,7
27	24,9	24,5	0,4	22,3	2,6
Maior erro observado			1,50		6,60
Erro médio			0,50		2,13
Desvio padrão			0,34		1,60

29

Esqueleto: Estigmastano - Polymastiidae – *Polymastia tenax*¹⁷¹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	32,1	32,1	0,0	32,0	0,1
2	31,1	31,4	0,3	33,7	2,6
3	68,6	67,9	0,7	67,5	1,1
4	39,5	39,9	0,4	40,6	1,1
5	67,7	66,1	1,6	65,4	2,3
6	61,3	61,9	0,6	61,5	0,2
7	65,1	65,9	0,8	71,3	6,2
8	125,0	126,0	1,0	133,2	8,2
9	38,6	38,0	0,6	37,1	1,5
10	35,8	36,6	0,8	45,1	9,3
11	18,9	18,7	0,2	20,5	1,6
12	36,6	36,9	0,3	37,3	0,7
13	43,1	42,5	0,6	42,4	0,7
14	152,6	151,2	1,4	138,0	14,6
15	24,9	25,3	0,4	26,1	1,2
16	26,6	27,0	0,4	27,2	0,6
17	56,6	56,3	0,3	56,3	0,3
18	17,9	18,1	0,2	18,4	0,5
19	16,5	16,2	0,3	15,5	1,0
20	34,9	36,3	1,4	36,8	1,9
21	18,9	19,5	0,6	19,1	0,2
22	33,6	33,5	0,1	33,6	0,0
23	26,2	26,0	0,2	26,4	0,2
24	46,0	45,7	0,3	46,0	0,0
25	28,9	29,3	0,4	29,2	0,3
26	19,0	19,5	0,5	19,1	0,1
27	19,5	19,5	0,0	19,7	0,2
28	22,9	23,4	0,5	23,2	0,3
29	12,3	12,5	0,2	12,4	0,1
Maior erro observado			1,60		14,60
Erro médio			0,52		1,97
Desvio padrão			0,41		3,38

30

Esqueleto: Ergostano - Coral – *Nephthea chabroli*¹⁷²

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	33,6	34,2	0,6	34,6	1,0
2	31,0	31,3	0,3	31,5	0,5
3	71,2	70,1	1,1	69,6	1,6
4	42,2	41,9	0,3	38,2	4,0
5	50,6	49,8	0,8	38,7	11,9
6	39,6	38,4	1,2	27,8	11,8
7	21,6	23,2	1,6	33,5	11,9
8	39,0	38,4	0,6	38,4	0,6
9	141,5	142,3	0,8	147,1	5,6
10	42,5	42,0	0,5	42,1	0,4
11	128,7	127,5	1,2	116,4	12,3
12	65,1	65,3	0,2	73,3	8,2
13	42,3	43,2	0,9	43,1	0,8
14	42,9	43,5	0,6	50,1	7,2
15	24,1	27,0	2,9	27,0	2,9
16	28,2	27,6	0,6	25,7	2,5
17	55,8	54,8	1,0	54,0	1,8
18	12,0	11,7	0,3	10,9	1,1
19	63,2	63,5	0,3	63,7	0,5
20	35,7	36,1	0,4	37,1	1,4
21	18,8	18,5	0,3	18,5	0,3
22	34,8	34,9	0,1	34,9	0,1
23	31,8	31,5	0,3	31,2	0,6
24	156,9	156,7	0,2	156,7	0,2
25	33,9	34,0	0,1	34,0	0,1
26	22,0	22,0	0,0	22,0	0,0
27	21,9	22,0	0,1	22,0	0,1
28	106,0	106,3	0,3	106,2	0,2
Maior erro observado			2,90		12,30
Erro médio			0,63		3,20
Desvio padrão			0,60		4,22

31

Esqueleto: Pregnano - Asclepiadaceae – *Solenostemma argel*¹⁷³

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	26,6	27,9	1,3	35,8	9,2
2	33,1	34,0	0,9	33,9	0,8
3	199,2	199,0	0,2	198,9	0,3
4	123,6	124,2	0,6	123,8	0,2
5	170,5	170,7	0,2	170,6	0,1
6	33,9	33,1	0,8	32,1	1,8
7	35,9	34,6	1,3	30,0	5,9
8	39,9	38,4	1,5	37,3	2,6
9	48,5	49,3	0,8	50,2	1,7
10	38,7	39,2	0,5	41,9	3,2
11	20,2	20,6	0,4	20,9	0,7
12	37,5	36,6	0,9	35,6	1,9
13	47,4	47,0	0,4	46,1	1,3
14	81,6	83,0	1,4	85,6	4,0
15	73,7	74,5	0,8	76,1	2,4
16	36,3	35,8	0,5	31,1	5,2
17	59,9	58,7	1,2	58,7	1,2
18	15,7	16,3	0,6	17,0	1,3
19	17,7	18,0	0,3	18,0	0,3
20	217,5	217,3	0,2	209,3	8,2
21	32,9	32,1	0,8	30,2	2,7
Maior erro observado			1,50		9,20
Erro médio			0,74		2,62
Desvio padrão			0,41		2,57

32

Esqueleto: Vitanolídeo - Solanaceae – *Solanum cilistum*¹⁷⁴

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	203,8	204,2	0,4	204,2	0,4
2	127,9	128,0	0,1	128,0	0,1
3	145,7	145,2	0,5	145,2	0,5
4	33,7	33,5	0,2	33,5	0,2
5	136,3	136,0	0,3	136,0	0,3
6	124,9	124,6	0,3	124,6	0,3
7	31,2	30,8	0,4	30,8	0,4
8	33,7	32,7	1,0	31,0	2,7
9	43,3	43,0	0,3	43,0	0,3
10	50,8	50,6	0,2	50,6	0,2
11	37,5	36,8	0,7	21,8	15,7
12	32,8	33,5	0,7	33,6	0,8
13	48,1	47,4	0,7	46,4	1,7
14	50,6	51,2	0,6	51,2	0,6
15	24,0	24,0	0,0	24,0	0,0
16	24,1	27,8	3,7	35,9	11,8
17	85,0	83,9	1,1	82,9	2,1
18	15,1	15,5	0,4	15,7	0,6
19	19,0	18,5	0,5	18,3	0,7
20	44,0	42,9	1,1	41,2	2,8
21	10,1	11,4	1,3	15,7	5,6
22	67,0	67,3	0,3	67,5	0,5
23	34,0	33,5	0,5	33,1	0,9
24	63,1	63,7	0,6	64,0	0,9
25	63,0	63,3	0,3	63,3	0,3
26	92,8	92,1	0,7	91,8	1,0
27	17,4	16,8	0,6	16,6	0,8
28	19,0	18,6	0,4	18,6	0,4
Maior erro observado			3,70		15,70
Erro médio			0,64		1,88
Desvio padrão			0,68		3,59

33

Esqueleto: Vitanolídeo - Solanaceae – *Solanum cilistum*¹⁷⁴

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	203,9	204,2	0,3	204,2	0,3
2	128,0	128,0	0,0	128,0	0,0
3	145,8	145,2	0,6	145,2	0,6
4	33,6	33,5	0,1	33,5	0,1
5	136,4	136,0	0,6	136,0	0,6
6	124,9	124,6	0,3	124,6	0,3
7	31,3	30,8	0,5	30,8	0,5
8	33,8	32,7	1,1	31,0	2,8
9	43,4	43,0	0,4	43,0	0,4
10	50,8	50,6	0,2	50,6	0,2
11	37,4	36,8	0,6	21,8	15,6
12	32,9	33,5	0,6	33,6	0,7
13	48,3	47,4	0,9	46,4	1,9
14	50,7	51,2	0,5	51,2	0,5
15	24,0	24,0	0,0	24,0	0,0
16	24,2	27,8	3,6	35,9	11,7
17	85,1	83,9	1,2	82,9	2,2
18	15,3	15,5	0,2	15,7	0,4
19	19,0	18,5	0,5	18,3	0,7
20	43,8	42,9	0,9	41,2	2,6
21	10,5	11,4	0,9	15,7	5,2
22	70,1	69,6	0,5	69,2	0,9
23	41,1	40,2	0,9	39,0	2,1
24	75,3	76,7	1,4	77,3	2,0
25	74,5	75,9	1,4	76,9	2,4
26	98,4	96,3	2,1	92,3	6,1
27	16,2	16,8	0,6	14,2	2,0
28	22,6	23,1	0,5	28,3	5,7
Maior erro observado			3,60		15,60
Erro médio			0,76		2,45
Desvio padrão			0,73		3,62

34

Esqueleto: Androstano - Meliaceae – *Aglaia rubiginosa*¹⁷⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	155,1	156,0	0,9	156,3	1,2
2	127,7	127,5	0,2	127,3	0,4
3	186,1	186,0	0,1	185,9	0,2
4	119,8	120,3	0,5	124,2	4,4
5	169,6	170,0	0,4	165,0	4,6
6	68,0	68,9	0,9	69,9	1,9
7	41,1	40,3	0,8	39,5	1,6
8	33,6	32,9	0,7	32,6	1,0
9	52,1	53,5	1,4	54,7	2,6
10	43,5	43,0	0,5	42,4	1,1
11	22,0	22,0	0,0	22,0	0,0
12	31,1	32,0	0,9	32,5	1,4
13	47,7	48,0	0,3	48,2	0,5
14	50,2	51,0	0,8	52,3	2,1
15	21,9	21,9	0,0	21,8	0,1
16	35,6	35,7	0,1	35,8	0,2
17	219,5	219,9	0,4	220,2	0,7
18	13,8	13,9	0,1	13,9	0,1
19	19,1	19,7	0,6	19,8	0,7
Maior erro observado			1,40		4,60
Erro médio			0,51		1,31
Desvio padrão			0,38		1,35

35

Esqueleto: Colestano - Meliaceae – *Aglaia rubiginosa*¹⁷⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	45,1	42,5	2,6	38,7	6,4
2	73,0	70,5	2,5	67,3	5,7
3	72,7	74,1	1,4	79,1	6,4
4	76,0	75,0	1,0	71,0	5,0
5	45,9	46,0	0,1	46,1	0,2
6	27,2	27,0	0,2	25,9	1,3
7	119,2	119,0	0,2	118,1	1,1
8	138,0	138,3	0,3	139,0	1,0
9	52,8	50,3	2,5	45,3	7,5
10	34,4	35,5	1,1	37,0	2,6
11	21,3	21,5	0,2	21,6	0,3
12	39,8	39,5	0,3	39,1	0,7
13	44,2	41,5	2,7	39,3	4,9
14	54,9	54,9	0,0	54,9	0,0
15	23,6	23,3	0,3	22,9	0,7
16	27,6	28,0	0,4	28,6	1,0
17	54,3	54,5	0,2	54,7	0,4
18	12,1	12,2	0,1	12,4	0,3
19	18,5	17,5	1,0	14,6	3,9
20	39,2	39,0	0,2	38,9	0,3
21	13,0	13,1	0,1	13,3	0,3
22	80,6	80,5	0,1	80,1	0,5
23	25,7	27,1	1,4	28,4	2,7
24	39,1	38,9	0,2	38,7	0,4
25	79,6	80,0	0,4	80,4	0,8
26	28,2	28,0	0,2	27,6	0,6
27	29,0	28,7	0,3	28,5	0,5
Maior erro observado			2,70		7,50
Erro médio			0,74		2,06
Desvio padrão			0,87		2,34

36

Esqueleto: Colestano - Meliaceae – *Aglaia rubiginosa*¹⁷⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	37,3	36,0	1,3	34,3	3,0
2	25,6	26,0	0,4	26,9	1,3
3	72,6	72,7	0,1	72,7	0,1
4	73,2	73,3	0,1	73,5	0,3
5	44,4	45,0	0,6	45,8	1,4
6	26,0	26,0	0,0	25,9	0,1
7	117,9	118,0	0,1	118,1	0,2
8	139,0	139,2	0,2	139,2	0,2
9	50,7	48,1	2,6	45,0	5,7
10	34,2	35,3	1,1	37,1	2,9
11	21,0	21,3	0,3	21,6	0,6
12	39,4	39,5	0,1	39,1	0,3
13	43,9	41,5	2,4	39,3	4,6
14	54,6	54,9	0,3	54,9	0,3
15	23,1	23,3	0,2	22,9	0,2
16	27,3	28,0	0,7	28,6	1,3
17	53,7	54,5	0,8	54,7	1,0
18	11,8	12,2	0,4	12,4	0,6
19	15,2	15,0	0,2	14,2	1,0
20	38,5	39,0	0,5	38,9	0,4
21	12,4	13,1	0,7	13,3	0,9
22	80,2	80,5	0,3	80,1	0,1
23	24,8	27,1	2,3	28,4	3,6
24	38,9	38,9	0,0	38,7	0,2
25	79,7	80,0	0,3	80,4	0,7
26	28,0	28,0	0,0	27,6	0,4
27	28,6	28,7	0,1	28,5	0,1
Maior erro observado			2,60		5,70
Erro médio			0,60		1,17
Desvio padrão			0,74		1,49

37

Esqueleto: Estigmastano - Sapindaceae – *Schleicheria oleosa*¹⁷⁶

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	37,1	37,3	0,2	37,5	0,4
2	31,6	32,0	0,4	32,1	0,5
3	71,5	71,9	0,4	72,2	0,7
4	42,0	42,9	0,9	43,2	1,2
5	144,0	143,4	0,6	141,5	2,5
6	121,4	122,5	1,1	124,7	3,3
7	81,9	81,0	0,9	79,9	2,0
8	37,2	37,9	0,7	38,7	1,5
9	48,4	48,5	0,1	48,8	0,4
10	36,5	37,0	0,5	37,7	1,2
11	21,2	21,3	0,1	21,3	0,1
12	39,6	39,7	0,1	39,7	0,1
13	43,2	43,3	0,1	43,0	0,2
14	56,0	55,2	0,8	49,8	6,4
15	25,8	26,3	0,5	27,1	1,3
16	27,7	27,0	0,7	26,5	1,2
17	52,5	52,9	0,4	53,2	0,7
18	11,9	12,0	0,1	12,2	0,3
19	19,0	19,0	0,0	18,9	0,1
20	42,5	43,5	1,0	44,4	1,9
21	12,4	13,0	0,6	13,2	0,8
22	71,3	72,5	1,2	74,2	2,9
23	29,9	30,3	0,4	35,3	5,4
24	41,4	42,0	0,6	42,9	1,5
25	28,8	29,5	0,7	31,4	2,6
26	17,6	18,1	0,5	19,4	1,8
27	20,5	20,1	0,4	19,8	0,7
28	23,5	24,1	0,6	25,0	1,5
29	11,8	12,3	0,5	12,7	0,9
Maior erro observado			1,20		6,40
Erro médio			0,52		1,52
Desvio padrão			0,32		1,50

38

Esqueleto: Ergostano - Sapindaceae – *Schleicheria oleosa*¹⁷⁶

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	36,4	36,7	0,3	36,9	0,5
2	32,2	32,1	0,1	32,1	0,1
3	70,6	70,2	0,4	70,2	0,4
4	41,8	42,1	0,3	42,9	1,1
5	165,2	166,0	0,8	166,9	1,7
6	126,1	126,0	0,1	125,7	0,4
7	199,8	200,0	0,2	200,2	0,4
8	45,5	45,6	0,1	45,4	0,1
9	49,9	50,0	0,1	50,1	0,2
10	38,4	38,2	0,2	38,1	0,3
11	21,2	21,4	0,2	21,4	0,2
12	38,7	39,2	0,5	39,6	0,9
13	43,4	44,2	0,8	45,5	2,1
14	49,5	50,1	0,6	50,4	0,9
15	26,4	26,5	0,1	26,6	0,2
16	27,9	27,0	0,9	26,5	1,4
17	51,8	52,2	0,4	52,3	0,5
18	12,0	12,0	0,0	12,1	0,1
19	17,3	17,4	0,1	17,4	0,1
20	42,7	43,5	0,8	44,3	1,6
21	12,5	13,0	0,5	13,2	0,7
22	71,9	72,3	0,4	73,3	1,4
23	34,6	35,7	1,1	37,7	3,1
24	35,4	36,6	1,2	39,4	4,0
25	29,7	30,3	0,6	31,4	1,7
26	16,2	17,0	0,8	18,3	2,1
27	21,1	20,8	0,3	20,5	0,6
28	15,7	16,5	0,8	18,8	3,1
Maior erro observado			1,20		4,00
Erro médio			0,45		1,07
Desvio padrão			0,34		1,04

39

Esqueleto: 9,11-Seco-Ergostano - Coral – *Pachyclavularia violacea*¹⁷⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	28,5	29,1	0,6	30,5	2,0
2	30,5	31,0	0,5	30,7	0,2
3	68,0	68,3	0,3	68,7	0,7
4	38,5	39,4	0,9	40,5	2,0
5	65,5	66,1	0,6	69,5	4,0
6	58,1	58,3	0,2	60,8	2,7
7	26,3	26,5	0,2	33,0	6,7
8	38,8	39,5	0,7	39,5	0,7
9	214,6	213,7	0,9	212,6	2,0
10	46,5	46,4	0,1	46,4	0,1
11	58,8	58,9	0,1	58,9	0,1
12	41,1	41,5	0,4	41,2	0,1
13	45,6	45,1	0,5	46,4	0,8
14	45,2	46,0	0,8	46,5	1,3
15	22,5	22,9	0,4	23,1	0,6
16	25,8	26,7	0,9	26,5	0,7
17	49,5	50,1	0,6	50,0	0,5
18	18,1	17,3	0,8	17,8	0,3
19	19,7	20,5	0,8	22,7	3,0
20	34,4	34,9	0,5	34,9	0,5
21	19,6	19,1	0,5	18,8	0,8
22	33,0	33,0	0,0	34,1	1,1
23	31,4	31,2	0,2	30,7	0,7
24	39,1	38,3	0,8	38,6	0,5
25	31,5	31,7	0,2	31,7	0,2
26	20,4	20,0	0,4	20,5	0,1
27	17,6	18,5	0,9	18,3	0,7
28	15,4	15,7	0,3	15,5	0,1
Maior erro observado			0,90		6,70
Erro médio			0,50		1,19
Desvio padrão			0,28		1,46

Esqueleto: Cardanolídeo - Cruciferae - *Erysimum cheiranthoides*¹⁷⁸

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	24,2	25,0	0,8	25,5	1,3
2	26,7	27,2	0,5	28,4	1,7
3	75,3	76,0	0,7	77,3	2,0
4	36,4	36,9	0,5	36,9	0,5
5	73,8	75,2	1,4	76,1	2,3
6	37,2	36,4	0,8	35,7	1,5
7	24,2	23,5	0,7	22,1	2,1
8	40,5	41,2	0,7	41,7	1,2
9	39,4	40,3	0,9	40,6	1,2
10	55,1	54,8	0,3	54,8	0,3
11	21,5	21,7	0,2	21,7	0,2
12	40,1	40,0	0,1	40,0	0,1
13	50,3	49,7	0,6	49,9	0,4
14	84,8	85,7	0,9	86,1	1,3
15	32,5	32,7	0,2	32,7	0,2
16	27,5	27,0	0,5	27,1	0,4
17	51,2	50,6	0,6	50,8	0,4
18	16,5	16,1	0,4	16,1	0,4
19	177,0	177,3	0,3	183,3	6,3
20	176,0	176,0	0,0	176,0	0,0
21	73,8	73,9	0,1	73,9	0,1
22	117,7	117,5	0,2	117,4	0,3
23	174,6	175,2	0,6	175,0	0,4
Dig-1'	97,7	97,9	0,2	97,7	0,0
2'	39,8	---	---	37,7	2,1
3'	68,9	---	---	67,7	1,2
4'	82,5	---	---	83,4	0,9
5'	67,5	---	---	69,3	1,8
6'	18,5	---	---	18,4	0,1
Rha-1'	104,1	---	---	103,8	0,3
2'	72,4	---	---	72,1	0,3
3'	72,5	---	---	72,3	0,2
4'	73,8	---	---	73,6	0,2
5'	70,4	---	---	70,2	0,2
6'	18,6	---	---	18,1	0,5
Maior erro observado			1,40		6,30
Erro médio			0,51		0,93
Desvio padrão			0,33		1,17

41

Esqueleto: Ergostano - Coral – *Clavularia viridis*¹⁷⁹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	205,3	204,6	0,7	201,4	3,9
2	128,6	128,8	0,2	128,8	0,2
3	140,1	141,1	1,0	144,3	4,2
4	31,7	32,0	0,3	32,6	0,9
5	78,2	79,4	1,2	82,8	4,6
6	66,0	65,3	0,7	62,4	3,6
7	38,8	37,4	1,4	35,8	3,0
8	34,0	32,6	1,4	31,5	2,5
9	54,3	52,9	1,4	52,4	1,9
10	54,5	52,1	2,4	47,0	7,5
11	66,8	66,8	0,0	66,7	0,1
12	48,7	49,4	0,7	49,5	0,8
13	43,1	44,0	0,9	43,9	0,8
14	54,7	55,4	0,7	56,4	1,7
15	28,4	27,5	0,9	24,1	4,3
16	23,8	24,9	1,1	28,9	5,1
17	55,7	56,3	0,6	56,4	0,7
18	13,2	14,9	1,7	15,4	2,2
19	9,4	10,1	0,7	16,4	7,0
20	40,0	38,6	1,4	35,6	4,4
21	20,9	20,9	0,0	20,9	0,0
22	135,1	134,9	0,2	134,9	0,2
23	132,3	132,8	0,5	133,9	1,6
24	42,8	41,1	1,7	37,7	5,1
25	33,1	33,5	0,4	33,3	0,2
26	19,6	20,0	0,4	20,0	0,4
27	19,9	20,0	0,1	20,5	0,6
28	17,6	18,0	0,4	17,9	0,3
Maior erro observado			2,40		7,50
Erro médio			0,83		2,42
Desvio padrão			0,59		2,20

Esqueleto: Colestano - Esponja – *Dysidea sp.*¹⁸⁰

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	29,5	30,1	0,6	30,8	1,3
2	30,5	30,9	0,4	30,8	0,3
3	66,7	67,2	0,5	67,7	1,0
4	36,4	38,3	1,9	40,3	3,9
5	75,6	76,1	0,5	80,6	5,0
6	69,7	71,1	1,4	77,2	7,5
7	67,2	66,6	0,6	66,7	0,5
8	132,2	131,0	1,2	130,1	2,1
9	142,8	143,3	0,5	143,2	0,4
10	44,3	43,8	0,5	41,3	3,0
11	66,1	64,8	1,3	63,8	2,3
12	49,4	47,2	2,2	39,6	9,8
13	47,4	45,1	2,3	42,0	5,4
14	48,2	49,9	1,7	51,3	3,1
15	22,6	23,7	1,1	24,3	1,7
16	28,5	28,0	0,5	28,2	0,3
17	56,1	56,7	0,6	56,5	0,4
18	13,7	14,0	0,3	14,0	0,3
19	24,4	22,9	1,5	21,4	3,0
20	36,2	35,1	1,1	35,3	0,9
21	18,6	19,2	0,6	19,0	0,4
22	36,1	36,3	0,2	36,2	0,1
23	24,0	24,2	0,2	24,1	0,1
24	39,7	39,5	0,2	39,6	0,1
25	28,2	28,3	0,1	28,1	0,1
26	23,0	22,8	0,2	22,7	0,3
27	22,7	22,8	0,1	22,7	0,0
Ac	170,4	170,1	0,3	169,8	0,6
	21,1	21,1	0,0	21,1	0,0
Ac	170,1	170,1	0,0	171,1	1,0
	20,8	21,1	0,3	21,2	0,4
Maior erro observado			2,30		9,80
Erro médio			0,74		1,78
Desvio padrão			0,65		2,38

43

Esqueleto: Pregnano - Asclepiadaceae – *Caralluma russeliana*¹⁸¹

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	38,7	38,0	0,7	37,4	1,3
2	30,8	30,1	0,7	29,9	0,9
3	79,9	79,3	0,6	78,4	1,5
4	39,7	39,2	0,5	39,2	0,5
5	140,9	141,5	0,6	141,5	0,6
6	123,3	122,4	0,9	121,9	1,4
7	28,4	27,8	0,6	27,8	0,6
8	38,4	39,0	0,6	39,0	0,6
9	47,8	46,7	1,1	46,4	1,4
10	38,0	38,0	0,0	38,0	0,0
11	22,3	21,5	0,8	21,5	0,8
12	41,6	40,4	1,2	35,9	5,7
13	49,0	47,3	1,7	46,2	2,8
14	86,1	84,4	1,7	81,4	4,7
15	33,9	34,2	0,3	33,8	0,1
16	20,1	22,5	2,4	27,2	7,1
17	58,1	56,6	1,5	54,9	3,2
18	15,4	15,7	0,3	15,4	0,0
19	20,3	20,0	0,3	19,7	0,6
20	79,1	78,8	0,3	76,5	2,6
21	21,4	20,3	1,1	19,9	1,5
Gal-1'	102,9	---	---	101,2	1,7
2'	71,3	---	---	70,6	0,7
3'	84,7	---	---	83,0	1,7
4'	68,8	---	---	68,1	0,7
5'	71,8	---	---	70,4	1,4
6'	17,1	---	---	16,5	0,6
MeO	57,3	---	---	57,4	0,1
Glc-1'	104,5	104,5	0,0	103,7	0,8
2'	75,4	75,3	0,1	73,8	1,6
3'	78,1	77,2	0,9	76,9	1,2
4'	71,4	71,6	0,2	71,6	0,2
5'	78,7	77,9	0,8	77,9	0,8
6'	62,8	61,8	1,0	62,8	0,0
Maior erro observado			2,40		7,10
Erro médio			0,77		1,45
Desvio padrão			0,57		1,61

Esqueleto: Colestano - Agavaceae – *Yucca schidigera*¹⁸²

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	31,4	32,7	1,3	37,3	5,9
2	27,3	28,1	0,8	30,0	2,7
3	76,0	76,8	0,8	77,8	1,8
4	30,9	31,5	0,6	35,3	4,4
5	37,5	38,9	1,4	44,8	7,3
6	27,6	28,0	0,4	29,1	1,5
7	27,7	28,5	0,8	32,3	4,6
8	36,7	35,8	0,9	35,3	1,4
9	41,2	43,7	2,5	53,9	12,7
10	36,1	36,0	0,1	36,0	0,1
11	21,9	21,4	0,5	21,4	0,5
12	41,3	40,4	0,9	39,1	2,2
13	42,2	42,0	0,2	40,8	1,4
14	57,6	56,9	0,7	56,8	0,8
15	32,7	32,3	0,4	32,3	0,4
16	83,4	82,1	1,3	79,0	4,4
17	65,3	64,6	0,7	63,5	1,8
18	16,9	17,0	0,1	17,0	0,1
19	24,5	22,3	2,2	12,6	11,9
20	41,2	42,5	1,3	44,8	3,6
21	16,2	15,7	0,5	13,9	2,3
22	113,9	113,2	0,7	112,9	1,0
23	31,4	32,4	1,0	34,6	3,2
24	28,9	28,4	0,5	28,1	0,8
25	35,0	36,6	1,6	37,2	2,2
26	75,1	75,5	0,4	75,4	0,3
27	17,4	17,8	0,4	17,4	0,0
Glc-1'	101,1; 104,8	100,8; 104,5	0,3; 0,3	100,2; 105,3	0,9; 0,5
2'	80,6; 75,9	81,3; 75,3	0,7; 0,6	82,8; 75,3	2,2; 0,6
3'	77,8; 77,7	77,2; 77,2	0,6; 0,5	76,5; 76,5	1,3; 1,2
4'	71,5; 71,8	71,8; 71,6	0,3; 0,2	71,8; 71,6	0,3; 0,2
5'	78,2; 78,0	77,9; 77,9	0,3; 0,1	76,7; 77,8	1,5; 0,2
6'	62,4; 62,4	61,8; 61,8	0,6; 0,6	61,2; 62,9	1,2; 0,5
Glc-1'	104,7	104,8	0,1	105,2	0,5
2'	74,9	75,3	0,4	75,4	0,5
3'	77,3	77,2	0,1	78,5	1,2
4', 5'	71,5; 77,8	71,6; 77,9	0,1; 0,1	71,8; 78,5	0,3; 0,7
6'	62,4	61,8	0,6	62,9	0,5
Maior erro observado			2,50		12,70
Erro médio			0,66		2,08
Desvio padrão			0,53		2,75

45

Esqueleto: Androstano - Fungo – *Botrytis cinerea*¹⁸³

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	37,1	36,9	0,2	36,9	0,2
2	34,4	35,3	0,9	36,8	2,4
3	213,5	213,3	0,2	213,2	0,3
4	34,8	36,7	1,9	45,8	11,0
5	40,3	41,4	1,1	43,7	3,4
6	30,4	31,3	0,9	33,9	3,5
7	70,1	70,9	0,8	71,5	1,4
8	38,0	39,1	1,1	39,1	1,1
9	59,4	57,6	1,8	40,2	19,2
10	36,5	35,5	1,0	35,3	1,2
11	20,5	21,0	0,5	21,2	0,7
12	42,2	41,1	1,1	37,1	5,1
13	43,5	44,6	1,1	45,2	1,7
14	51,6	50,8	0,8	49,7	1,9
15	23,2	23,3	0,1	23,3	0,1
16	29,9	30,5	0,6	30,5	0,6
17	81,5	81,5	0,0	81,5	0,0
18	11,1	11,3	0,2	11,3	0,2
19	23,1	21,9	1,2	21,9	1,2
Maior erro observado			1,90		19,20
Erro médio			0,82		2,91
Desvio padrão			0,53		4,70

46

Esqueleto: Pregnano - Fungo – *Botrytis cinerea*¹⁸³

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	37,2	36,7	0,5	36,0	1,2
2	34,2	33,8	0,4	33,8	0,4
3	199,9	199,8	0,1	199,8	0,1
4	125,6	124,6	1,0	123,0	2,6
5	166,7	167,8	1,1	172,1	5,4
6	43,1	41,4	1,7	32,9	10,2
7	23,9	25,6	1,7	32,2	8,3
8	42,4	40,1	2,3	33,0	9,4
9	56,3	56,0	0,3	56,0	0,3
10	39,4	38,5	0,9	38,5	0,9
11	73,9	71,3	2,6	65,5	8,4
12	50,5	49,5	1,0	49,5	1,0
13	44,8	44,4	0,4	44,4	0,4
14	54,6	55,7	1,1	55,7	1,1
15	27,1	29,7	2,6	38,4	11,3
16	68,8	69,6	0,8	70,6	1,8
17	62,3	64,1	1,8	69,0	6,7
18	14,6	15,8	1,2	16,1	1,5
19	18,4	19,6	1,2	19,6	1,2
20	208,9	207,7	1,2	206,8	2,1
21	31,4	32,0	0,6	32,0	0,6
Maior erro observado			2,60		11,30
Erro médio			1,17		3,57
Desvio padrão			0,72		3,81

47

Esqueleto: Colestano - Liliaceae – *Galtonia candicans*¹⁸⁴

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	68,1	68,7	0,6	68,8	0,7
2	35,0	35,3	0,3	35,7	0,7
3	72,4	72,1	0,3	71,6	0,8
4	31,4	32,3	0,9	33,7	2,3
5	32,0	33,8	1,8	37,2	5,2
6	26,5	27,2	0,7	28,4	1,9
7	26,9	27,8	0,9	32,1	5,2
8	34,4	35,1	0,7	35,9	1,5
9	47,8	48,4	0,6	49,1	1,3
10	47,0	46,9	0,1	46,9	0,1
11	65,6	67,7	2,1	67,7	2,1
12	51,1	49,7	1,4	49,9	1,2
13	42,5	40,4	2,1	39,9	2,6
14	52,8	53,5	0,7	54,9	2,1
15	34,4	33,6	0,8	32,3	2,1
16	72,5	74,8	2,3	77,3	4,8
17	58,8	59,0	0,2	59,0	0,2
18	16,0	15,1	0,9	14,8	1,2
19	59,7	60,8	1,1	61,8	2,1
20	33,5	33,4	0,1	33,4	0,1
21	17,1	16,9	0,2	14,8	2,3
22	86,0	85,5	0,5	83,4	2,6
23	76,0	74,3	1,7	65,2	10,8
24	125,6	124,8	0,8	123,8	1,8
25	136,9	135,1	1,8	134,1	2,8
26	26,1	25,7	0,4	25,7	0,4
27	18,9	18,5	0,4	18,5	0,4
Gal ₂ -1'	126,7	125,4	1,3	120,3	6,4
2'	107,5	108,4	0,9	111,0	3,5
3'	153,6	153,1	0,5	146,0	7,6
4'	142,9	142,3	0,6	139,1	3,8
5'	153,6	153,1	0,5	146,0	7,6
6'	107,5	108,4	0,9	111,0	3,5
7'	165,5	165,8	0,3	166,1	0,6

47

Esqueleto: Colestano - Liliaceae – *Galtonia candicans*¹⁸⁴

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
Glc-1'	105,1	105,3	0,2	105,2	0,1
2'	75,4	75,3	0,1	73,6	1,8
3'	76,5	77,2	0,7	75,8	0,7
4'	81,5	81,3	0,2	80,9	0,6
5'	76,0	76,5	0,5	76,3	0,3
6'	62,5	61,8	0,7	61,5	1,0
Ara-1'	105,8	105,8	0,0	105,8	0,0
2'	72,3	---	---	75,1	2,8
3'	74,6	---	---	76,3	1,7
4'	69,6	---	---	70,9	1,3
5'	67,9	---	---	67,4	0,5
Maior erro observado			2,30		10,80
Erro médio			0,78		2,29
Desvio padrão			0,60		2,32

48

Esqueleto: Vitanolídeo - Solanaceae – *Vassobia lorentzii*¹⁸⁵

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	202,5	202,1	0,4	202,2	0,3
2	131,9	132,2	0,3	132,2	0,3
3	142,1	142,9	0,8	142,9	0,8
4	69,7	70,0	0,3	70,0	0,3
5	63,7	64,0	0,3	64,0	0,3
6	62,6	62,0	0,6	62,0	0,6
7	31,3	31,3	0,0	31,3	0,0
8	30,9	30,4	0,5	30,4	0,5
9	43,9	42,7	1,2	42,9	1,0
10	47,8	48,0	0,2	47,9	0,1
11	24,6	24,2	0,4	23,6	1,0
12	34,6	35,4	0,8	35,7	1,1
13	58,0	58,5	0,5	58,8	0,8
14	56,4	55,9	0,5	55,7	0,7
15	25,4	25,7	0,3	25,5	0,1
16	26,2	26,8	0,6	26,9	0,7
17	57,0	57,0	0,0	57,0	0,0
18	101,0	101,5	0,5	102,2	1,2
19	17,6	17,3	0,3	17,4	0,2
20	85,1	85,3	0,2	85,0	0,1
21	21,8	22,6	0,8	23,3	1,5
22	80,5	80,4	0,1	80,5	0,0
23	31,5	31,7	0,2	31,7	0,2
24	147,7	148,2	0,5	148,1	0,4
25	122,3	122,4	0,1	122,2	0,1
26	165,7	165,3	0,4	165,0	0,7
27	12,5	12,5	0,0	12,5	0,0
28	20,5	20,4	0,1	20,4	0,1
Maior erro observado			1,20		1,50
Erro médio			0,39		0,47
Desvio padrão			0,28		0,42

49

Esqueleto: 17(13→18)-Abeoergostano - Solanaceae – *Salpichroa organifolia*¹⁸⁶

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	202,7	202,4	0,3	202,2	0,5
2	128,9	128,5	0,4	128,5	0,4
3	142,5	142,4	0,1	142,4	0,1
4	33,6	33,1	0,5	32,8	0,8
5	64,6	64,1	0,5	63,4	1,2
6	59,0	59,3	0,3	60,4	1,4
7	30,3	31,4	1,1	35,4	5,1
8	33,2	---	---	35,7	2,5
9	36,4	---	---	42,2	5,8
10	48,7	49,2	0,5	49,1	0,4
11	25,4	---	---	22,9	2,5
12	30,6	---	---	29,4	1,2
13	137,7	--	--	135,5	2,2
14	137,1	---	---	138,8	1,7
15	125,4	---	---	124,6	0,8
16	126,5	---	---	124,5	2,0
17	141,3	---	---	141,6	0,3
18	128,2	---	---	127,9	0,3
19	14,8	14,5	0,3	14,6	0,2
20	43,9	---	---	42,7	1,2
21	18,0	---	---	17,4	0,6
22	71,7	---	---	70,7	1,0
23	40,2	---	---	39,6	0,6
24	73,7	74,5	0,8	77,1	3,4
25	76,5	77,1	0,6	76,9	0,4
26	97,2	96,8	0,4	92,7	4,5
27	20,9	19,8	1,1	14,2	6,7
28	24,2	25,6	1,4	28,3	4,1
Maior erro observado			1,40		6,70
Erro médio			0,59		1,85
Desvio padrão			0,37		1,84

50

Esqueleto: Ergostano - Coral – *Plexaurella grisea*¹⁸⁷

Carbono	Literatura	FOCOS	Erro	ACD Labs	Erro
1	33,4	32,8	0,6	32,3	1,1
2	32,6	33,1	0,5	33,2	0,6
3	67,1	66,9	0,2	67,0	0,1
4	43,1	43,0	0,1	42,8	0,3
5	76,8	76,5	0,3	76,6	0,2
6	79,5	79,0	0,5	79,2	0,3
7	73,0	72,6	0,4	72,8	0,2
8	39,5	39,3	0,2	39,3	0,2
9	44,8	44,5	0,3	44,5	0,3
10	38,6	38,4	0,2	38,4	0,2
11	22,0	21,8	0,2	21,8	0,2
12	40,9	40,7	0,2	40,7	0,2
13	43,9	43,5	0,4	43,2	0,7
14	56,5	56,1	0,4	56,3	0,2
15	27,9	27,7	0,2	27,7	0,2
16	29,2	28,4	0,8	28,5	0,7
17	55,8	56,5	0,7	56,7	0,9
18	12,6	12,1	0,5	12,2	0,4
19	17,6	17,3	0,3	17,4	0,2
20	36,2	36,0	0,2	36,1	0,1
21	19,1	19,2	0,1	19,2	0,1
22	35,2	35,6	0,4	35,6	0,4
23	31,4	31,9	0,5	32,1	0,7
24	156,8	157,1	0,3	157,4	0,6
25	34,1	34,5	0,4	34,4	0,3
26	22,2	22,2	0,0	22,2	0,0
27	22,0	22,2	0,2	22,2	0,2
28	106,7	106,5	0,2	106,5	0,2
Maior erro observado			0,80		1,10
Erro médio			0,33		0,35
Desvio padrão			0,19		0,27

Tabela XVIII. Fragmentos estruturais propostos pelo programa ESTGERA

Teste	Anel A	Anel B	Anel C	Anel D	Cadeia Lateral
1	1-Oxo; 2,5-EN	5EN	SF	SF	20 β -OH; 24EN; 26Oxo;22,26Oxi
2	3 β -OH 3 α -OH	7EN	SF 11 β -OH	NP	NP
3	3 β -OH; 5EN	5EN; 7 α -OH	12 β -OR 12 α -OR	SF	22,23,28-Me; 23,29Ciclo
4	3 β -OH; 5EN	5EN	NP	16 β -OH	28 α -Me
5	1Oxo; 3 β -OR; 5 β ,6 β -EP	5 β ,6 β -EP	SF	SF	20 β -OH; 24EN; 26Oxo;22,26Oxi
6	3 β ,4 β -OH	SF	SF	SF	21-OR
7	3 β ,5 β -OR	5 β -OR; 19Oxo	14 β -OH	14 β ,16 β -OH	20,22EN;24Oxo 21,24Oxi
8	1 β ,3 β -OH	5 β -OH	SF	16 β ,22-Oxi	22 α ,26Oxi; 25(27)EN
9	1 β ,3 α OH; 5EN	5EN	SF	16EN; 20Oxo	---
10	2 β ,3 β -OH;5 β H	6Oxo; 7EN	9 α ,14 α -OH	14 α ,20 β -OH	20 β ,22 β ,25-OH 20 β ,22 α ,25-OH
11	2 β ,3 β -OH 3 β ,4 β -OH	6 α ,8 β -OH	8 β -OH	15 α -OH 15 β -OH	24(28)EN;26OH
12	3 β -OH	6 α -OH	SF	15 α -OH 15 β -OH	22EN; 24 α -OH
13	3 β -OH	6 α -OH	SF	15 α -OH 15 β -OH	24 α -OH
14	3 β -OH	6 α ,8 β -OH	8 β -OH	15 α -OH 15 β -OH	NP
15	3 β ,4 β -OH;5EN	5EN; 7 α -OH	SF	SF	22 α -OH; 24 β -Et
16	3-Oxo	SF	SF	17-Oxo	NP
17	NP	5 β -OH; 6Oxo	SF	SF	24(28)EN
18	3 β -OH; 4EN	NP	14 β ,15 β -EP	14 β ,15 β -EP	22EN; 24 α -Me
19	1Oxo; 2,5EN; 4 β -OH	5EN; 7 β -OH	SF	SF	20 β -OH; 24EN; 26Oxo;22,26Oxi

Tabela XVIII. Continuação

Teste	Anel A	Anel B	Anel C	Anel D	Cadeia Lateral
20	1 β ,3 β -OH	SF	SF	16 β ,22-Oxi	22 α ,26Oxi; 24 β -OR
21	3 β -OR	NP	11 α -OH 11 β -OH	NP	22EN; 24 α -Me
22	1-Oxo; 2EN; 5 β -OH	5 β ,6 α -OH	14 α -OH	14 α ,17 β ,20 β - OH	20 β -OH; 24EN; 26Oxo;22,26Oxi
23	3-Oxo	6-Oxo 7-Oxo	SF	SF	20 β -OH; 22EN
24	3-Oxo; 4EN	4EN; 6 α -OH	SF	SF	SF
25	3 β -OR; 5 β -H	5 β -H	14 β ,15 β -EP	NP	20,22EN;24Oxo 21,24Oxi
26	3-Oxo	5 α ,6 α -EP 7 β -OH	11 α -OH 12 α -OH	SF	24Oxo; 24-OH
27	3 β -OR	7EN	7EN	NP	NP
28	NP	6-Oxo; 7EN	7EN; 14 α -OH	14 α ,20 β -OH	20 β ,22 β -OH 20 β ,22 α -OH
29	3 β -OH; 5 α ,6 α - EP	5 α ,6 α -EP; 7 α - OH	8(14)EN	SF	24 β -Et
30	3 β ,19-OH	9(11)EN	NP	SF	24(28)EN
31	3-Oxo; 4EN	4EN	14 β -OH	14 β ,15 α -OH 14 β ,16 α -OH	20-Oxo
32	1-Oxo; 2,5EN	5EN	SF	17 α -OH	22,26-Oxi; 24 α ,25 α -EP; 26 α -OH
33	1-Oxo; 2,5EN	5EN	SF	17 α -OH	NP
34	1,4EN; 3-Oxo	4EN; 6 α -OH	SF	17-Oxo	---
35	3 β ,4 β -OH	7EN	7EN	SF	22,25-Oxi
36	3 β ,4 β -OH	7EN	7EN	SF	22,25-Oxi
37	3 β -OH; 5EN	NP	SF	SF	22 α -OH; 24 β -Et
38	3 β -OH; 5EN	5EN; 7-Oxo	SF	SF	22 α -OH;24 α -Et

Tabela XVIII. Continuação

Teste	Anel A	Anel B	Anel C	Anel D	Cadeia Lateral
39	3 β -OH; 5 β ,6 β -EP	5 β ,6 β -EP; 9-Oxo	9,11-Seco; 11-OH	SF	24 α -Me
40	3 β -OR; 5 α -OH 3 β -OR; 5 β -OH	5 β -OH	14 β -OH	14 β -OH	20(22)EN; 21,23Oxi;23Oxo
41	1-Oxo; 2EN; 5 β -OH	NP	SF 11 α -OH 11 β -OH	SF	22EN; 24 α -Me
42	3 β ,5 α -OH	NP	8EN 8EN; 11 α -OH	SF	SF
43	3 β -OR; 5EN	5EN	14 α -OH 14 β -OH	14 β -OH; 20OR	---
44	3 β -OR; 5 β -H	5 β -H	SF	16 β ,22-Oxi	NP
45	3-Oxo	7 α -OH 7 β -OH	SF	17 β -OH	---
46	3-Oxo; 4EN	4EN	SF 11 α -OH 11 β -OH	16 β -OH;20Oxo	---
47	NP	5 β -H	SF 11 α -OH 11 β -OH	16 β ,22-Oxi	NP
48	1-Oxo; 2EN; 4 β -OH	5 β ,6 β -EP	SF	SF	20 β -OR; 24EN; 26Oxo;22,26Oxi
49	1-Oxo; 2EN	5 α ,6 α -EP	NP	NP	NP
50	3 β ,5 α -OH	5 α ,6 β ,7 β -OH 6 β ,7 α ,8 β -OH 6 α ,7 α ,8 β -OH	SF	SF	24(28)EN

VII - DISCUSSÃO DE RESULTADOS

7.1. Programas aplicativos do SISTEMAT

Analisando os resultados dos testes com esteróides, apresentados anteriormente, verificamos que o sistema PESQUISA apresentou o esqueleto correto da substância, como primeira opção, em 90,0% dos casos. Em cinco casos o sistema não forneceu o esqueleto correto da substância, testes de números 11, 14, 16, 37 e 39, respectivamente, apresentados nas págs. 173, 176, 178, 199 e 201. Dentre esses casos, pode-se verificar que nos testes 14 e 16 o sistema não forneceu o esqueleto correto da substância em nenhuma das opções pois o esqueleto 23-Metilergostano é considerado um esqueleto novo, pois não apresenta dados de RMN ^{13}C na base de dados. Para os outros três casos, o sistema forneceu o esqueleto correto como segunda opção (testes 11 e 37) ou terceira opção (teste 39). Se considerarmos apenas os testes das substâncias cujos esqueletos apresentam dados de RMN ^{13}C no banco de dados, o acerto percentual do sistema passa a ser de 93,8% para previsão de esqueletos.

O programa SISCONST apresentou um índice de acerto de 78,0%, sendo os 22,0% restantes distribuídos entre 20,0% de erros (testes 14, 16, 17, 21, 31, 34, 37, 39, 43 e 50) e 2,0% de testes onde o programa não apresentou previsão de esqueleto (teste 49). Comparado com o programa C13MACH, que apresentou um acerto de 60,0% nos testes, o programa SISCONST apresentou-se mais eficiente. Isto deve-se ao fato de que o programa SISCONST seleciona para a determinação dos esqueletos mais prováveis apenas os esteróides que contenham no mínimo a metade do número de átomos de carbono da estrutura interligados, com os deslocamentos químicos e as multiplicidades requeridas pelo usuário. Já o programa C13MACH realiza apenas uma confrontação entre os deslocamentos químicos do espectro questionado e os dados espectrais presentes na base de dados, sem a restrição de interligação entre os átomos de carbono que apresentam os respectivos deslocamentos químicos. Nas tabelas XIX e XX são apresentados os resumos das análises realizadas através dos programas SISCONST e C13MACH, respectivamente.

Em geral, observamos que o programa SISCONST erra ou não apresenta previsão de esqueletos quando:

1. o número de substâncias referentes àquele esqueleto no banco de dados é muito pequeno ou inexistente (testes 14, 16 e 39);

2. a substância apresenta uma funcionalização incomum ao esqueleto (teste 17);
3. o espectro de RMN ^{13}C da substância questionada é muito semelhante ao de uma substância de um outro esqueleto (testes 31 e 34).

Tabela XIX. Resumo das análises realizadas através do programa SISCONST

Esqueletos	Nº Espectros	Nº Análises	Nº Acertos	Nº Erros	Sem Previsão
Vitanolídeo	60	7	7	---	---
Estigmastano	34	5	4	1	---
Gorgostano	4	1	1	---	---
Ergostano	100	9	6	3	---
Colestano	371	14	14	---	---
Colano	38	3	3	---	---
Pregnano	115	4	2	2	---
23-Metilergostano	0	2	---	2	---
Androstano	167	2	1	1	---
9,11-Secoergostano	6	1	---	1	---
Cardanolídeo	99	1	1	---	---
17(13→18)- Abeoergostano	5	1	---	---	1
TOTAL		50	39	10	1

É importante salientar aqui que, entre as dez propostas incorretas de esqueleto apresentadas pelo programa SISCONST, em oito casos o programa apresentou o esqueleto correto da substância como segunda ou terceira proposta de maior probabilidade de esqueleto. Somente nos dois testes do esqueleto 23-Metilergostano o esqueleto correto não foi proposto.

Analisando as subestruturas propostas pelo programa SISCONST, observamos que em vários testes, a sobreposição ou a união de algumas subestruturas pode conduzir a uma subestrutura maior ou até mesmo à estrutura completa da substância como, por exemplo, nos testes 1, 7, 23, 24 e 34. Dentre as subestruturas apresentadas nos testes, em 96,0% dos casos essas são corretas e compatíveis com as estruturas propostas na literatura. Muitas vezes, as subestruturas propostas são muito semelhantes, uma vez que um mesmo sinal do espectro de RMN ^{13}C analisado pode ser atribuído a diferentes átomos de carbono em diferentes subestruturas. Com relação à atribuição dos sinais de

RMN ^{13}C , dentre os sinais atribuídos pelo programa SISCONST nos testes realizados, 93,0% foram atribuídos corretamente, de acordo com os dados da literatura.

Tabela XX. Resumo das análises realizadas através do programa C13MACH

Esqueletos	Nº Análises	Nº Acertos	Nº Erros	2ª Opção	3ª Opção
Vitanolídeo	7	7	---	---	---
Estigmastano	5	---	5	1	---
Gorgostano	1	1	---	---	---
Ergostano	9	5	4	2	2
Colestano	14	10	4	2	1
Colano	3	2	1	---	---
Pregnano	4	2	2	1	---
23-Metilergostano	2	---	2	---	---
Androstano	2	2	---	---	---
9,11-Secoergostano	1	---	1	---	---
Cardanolídeo	1	1	---	---	---
17(13→18)- Abeoergostano	1	---	1	---	---
TOTAL		30	20	6	3

O programa REGRAS, que desfuncionaliza o espectro de RMN ^{13}C e realiza a pesquisa nos tipos de átomos de carbono presentes no esqueleto (Tabela III) e nas faixas de deslocamento químico características de esqueletos e subesqueletos (Tabelas XIII à XVII), apresentou um percentual de acerto de 90,0%. O programa cometeu apenas cinco erros, testes 11-14 e 16. No teste 11 (pg. 173) o programa forneceu uma previsão errônea de esqueleto pois, durante a desfuncionalização do espectro de RMN ^{13}C , os deslocamentos químicos 58,4 e 62,7 de multiplicidade dubleto, referentes aos carbonos 9 e 14 da substância, respectivamente, que não apresentam funções orgânicas, foram desfuncionalizados de acordo com as faixas de deslocamento químico da Tabela IV como sendo um CH-OH, sendo atribuídos como carbonos metilênicos no esqueleto totalmente desfuncionalizado. Assim, o programa propôs incorretamente o esqueleto 8,9-secoergostano ao invés do esqueleto ergostano.

Nos testes 12-14 e 16, o programa REGRAS não apresentou propostas de esqueleto, portanto não interferiu negativamente na probabilidade global de esqueletos da substância analisada. Nos testes 12, 13 e 14 o erro cometido pelo programa foi o

mesmo identificado no teste 11, isto é, o programa desfuncionalizou o carbono 14 das substâncias cujos deslocamentos químicos encontram-se à 62,2δ, 62,2δ e 62,7δ. respectivamente. Ainda no teste 14, o programa também desfuncionalizou o carbono 17 considerando o deslocamento químico que encontra-se à 58,5δ. Esse mesmo erro, voltou a ocorrer no teste 16, com a desfuncionalização do sinal à 67,4δ referente ao carbono 17 da substância.

A partir dos erros identificados acima, podemos verificar que, em todos os casos, a segunda faixa de desfuncionalização de hidroxilas (Tabela IV) foi a responsável pelo problema. Uma das soluções para esse problema é a alteração das faixas de desfuncionalização dos dados de RMN ^{13}C . Na Tabela XXI são apresentadas as novas faixas de desfuncionalização de dados a serem empregadas a partir desses resultados no sistema PESQUISA. As principais alterações ocorreram na primeira e segunda faixa de hidroxilas, que somente serão desfuncionalizadas a partir de 63,0δ. Com essa alteração, os erros verificados nos testes 11-14 seriam eliminados e o programa REGRAS apresentaria um acerto de 98,0%.

Tabela XXI. Novas faixas de deslocamentos químicos usadas para desfuncionalização dos dados de RMN ^{13}C

Função Química	Multiplicidade Inicial	Multiplicidade Final	Faixas de deslocamento químico
C=O	1	3	190,0 - 250,0
CHO	2	4	190,0 - 250,0
R-COO-R'	1	4	165,5 - 190,0
	1	2	113,0 - 165,5
C=C	2	3	104,0 - 167,0
	3	4	100,0 - 167,0
	1	2	63,0 - 100,0
C-OH	2	3	63,0 - 90,0
	3	4	54,0 - 90,0
C(OR) ₂	1	3	100,0 - 113,0
CH(OR) ₂	2	4	90,0 - 104,0

Em quinze dos cinquenta testes apresentados, foi utilizado o programa MACRONO, uma vez que as substâncias apresentavam macronós. Em 91,3% dos testes o programa identificou corretamente os grupos substituintes presentes. Na Tabela XXII é apresentado um resumo das análises realizadas através do programa MACRONO.

Tabela XXII. Resumo das análises realizadas através do programa MACRONO

Macronós	Nº Análises	Nº Acertos	Nº Erros
Acetoila	10	10	---
Glicosila	5	5	---
Rhamnosila	1	---	1
Gentiobiosila	1	1	---
Metoxila	1	1	---
Digitalosila-(4→1)-Rhamnosila	1	---	1
3-O-Metil-6-deoxigalactosila	1	1	---
Soforosila	1	1	---
Galoila	1	1	---
Glicosila-(4→1)-Arabinosila	1	1	---
TOTAL	23	21	2

Nos dois erros cometidos pelo programa, podemos verificar que:

A- No teste 7, foi proposto o açúcar rhamnosila, entretanto o macronó fucosila foi apresentado com uma menor faixa de erro médio.

B- No teste 40, o macronó Digitalosila-(4→1)-Rhamnosila não foi apresentado pois esse não consta da base de dados, assim como o açúcar digitalose isoladamente, o que torna impossível sua apresentação pelo programa. No entanto, foi proposta pelo programa uma das unidades do dissacarídeo, o açúcar rhamnosila.

É necessário comentar também que, no teste 47, o macronó Glicosila-(4→1)-Arabinosila não está presente na base de dados, porém o programa foi capaz de fornecer corretamente os dois glicosídeos pertencentes à estrutura, restando ao usuário a determinação da posição de ligação entre os mesmos.

Em certos casos, verificamos que o programa atribuiu ao substituinte um ou dois sinais do espectro de RMN ^{13}C que pertenciam, segundo os dados da literatura, a carbonos do esqueleto da substância. Isto se deve ao fato de que aqueles deslocamentos

químicos estavam mais próximos dos deslocamentos químicos do macronó presente no banco de dados. Nos testes aqui apresentados isto não prejudicou a previsão de esqueletos ou a obtenção de subestruturas, mas sim a atribuição dos sinais do espectro de RMN ^{13}C questionado. Para evitar esse tipo de erro e também a proposta de outros macronós que não correspondam aqueles presentes na substância analisada, mas que podem aparecer nos resultados devido à necessidade de se aumentar o gradiente na pesquisa do macronó, pretende-se armazenar os dados de deslocamentos químicos dos diversos macronós em vários solventes, de modo que a pesquisa possa ser realizada de acordo com o solvente no qual foi obtido o espectro de RMN ^{13}C da amostra. Muitas vezes, o efeito do solvente pode conduzir a uma grande diferença entre os deslocamentos químicos do macronó presente no banco de dados e aqueles do macronó da substância analisada em outro solvente, atingindo valores de δ , como é o caso dos açúcares.

O programa SISOCBOT que analisa os dados botânicos foi utilizado em 64,0% dos casos pois quando a substância analisada pertencia a um gênero ausente no banco de dados, embora a pesquisa pudesse ser realizada no nível de família, conduziria a resultados menos confiáveis. Assim, a pesquisa somente foi realizada para os gêneros presentes na base de dados. Se considerarmos todos os testes, o programa apresentou um percentual de acerto de 52,0%. No entanto, se usarmos para o cálculo somente os trinta e dois testes onde o programa SISOCBOT pôde ser utilizado, verifica-se que em 81,25% dos casos o programa apresentou o esqueleto correto da substância.

Comparando o resultado do SISOCBOT com o C13MACH, verificamos que o primeiro apresentou um acerto percentual maior, demonstrando assim a validade do dado botânico para fins de previsão de esqueletos, quando a revisão bibliográfica das ocorrências botânicas de uma classe de substâncias é significativa.

7.2. Programas de simulação de dados de RMN ^{13}C

Os resultados obtidos com os simuladores FOCOS e da ACD Labs apresentam-se sumariados na Tabela XXIII. A partir dessa, pode-se verificar que, dentre os cinquenta testes, o maior erro obtido com o programa FOCOS ocorreu no teste 32, 3,708. Em contrapartida, o maior erro observado através do programa da ACD Labs, 20,508, ocorreu no teste 1. Nesse último, em dezoito testes o maior erro observado foi superior a 10,08. O valor médio dos maiores erros observados foram de 1,978 e 8,788 para os simuladores FOCOS e da ACD Labs, respectivamente.

Tabela XXIII. Resumo das análises realizadas com esteróides através dos simuladores de dados de RMN ^{13}C

Teste	SIMULADOR FOCOS			SIMULADOR ACD Labs		
	Maior erro	Erro médio	Desvio padrão	Maior erro	Erro médio	Desvio padrão
1	1,10	0,35	0,27	20,50	2,55	4,89
2	2,10	0,82	0,48	8,10	2,05	2,03
3	2,30	0,64	0,49	11,30	2,58	2,67
4	1,70	0,69	0,47	12,70	2,89	3,28
5	1,80	0,46	0,36	14,80	1,98	3,03
6	2,10	0,51	0,53	13,60	2,53	3,82
7	1,50	0,54	0,41	6,30	1,96	1,86
8	2,00	0,72	0,44	6,30	1,60	1,28
9	1,60	0,52	0,45	3,60	1,10	1,14
10	1,70	0,61	0,47	12,20	1,75	2,42
11	1,30	0,52	0,51	3,60	0,92	0,86
12	2,20	0,80	0,48	5,40	1,91	1,33
13	2,20	0,70	0,55	4,80	1,58	1,22
14	1,90	0,69	0,44	7,90	1,42	1,60
15	2,50	1,01	0,66	8,10	2,67	2,03
16	2,30	0,62	0,56	5,60	1,51	1,77
17	1,70	0,51	0,49	10,90	1,48	2,37
18	3,10	1,01	0,74	9,50	2,73	2,69
19	2,30	0,69	0,57	10,90	1,87	2,48
20	2,40	0,41	0,53	10,20	1,13	2,06
21	1,80	0,54	0,40	5,50	1,58	1,79
22	2,20	0,64	0,47	5,30	1,38	1,36
23	1,50	0,45	0,38	5,50	0,99	1,33
24	1,20	0,35	0,29	2,90	0,34	0,58
25	1,60	0,53	0,41	5,40	1,32	1,40
26	2,50	1,08	0,56	17,00	4,89	4,43
27	1,90	0,73	0,46	8,20	1,87	2,07
28	1,50	0,50	0,34	6,60	2,13	1,60
29	1,60	0,52	0,41	14,60	1,97	3,38
30	2,90	0,63	0,60	12,30	3,20	4,22
31	1,50	0,74	0,41	9,20	2,62	2,57

Tabela XXIII. Continuação

Teste	SIMULADOR FOCOS			SIMULADOR ACD Labs		
	Maior erro	Erro médio	Desvio padrão	Maior erro	Erro médio	Desvio padrão
32	3,70	0,64	0,68	15,70	1,88	3,59
33	3,60	0,76	0,73	15,60	2,45	3,62
34	1,40	0,51	0,38	4,60	1,31	1,35
35	2,70	0,74	0,87	7,50	2,06	2,34
36	2,60	0,60	0,74	5,70	1,17	1,49
37	1,20	0,52	0,32	6,40	1,52	1,50
38	1,20	0,45	0,34	4,00	1,07	1,04
39	0,90	0,50	0,28	6,70	1,19	1,46
40	1,40	0,51	0,33	6,30	0,93	1,17
41	2,40	0,83	0,59	7,50	2,42	2,20
42	2,30	0,74	0,65	9,80	1,78	2,38
43	2,40	0,77	0,57	7,10	1,45	1,61
44	2,50	0,66	0,53	12,70	2,08	2,75
45	1,90	0,82	0,53	19,20	2,91	4,70
46	2,60	1,17	0,72	11,30	3,57	3,81
47	2,30	0,78	0,60	10,80	2,29	2,32
48	1,20	0,39	0,28	1,50	0,47	0,42
49	1,40	0,59	0,37	6,70	1,85	1,84
50	0,80	0,33	0,19	1,10	0,35	0,27
Média	1,97	0,64	0,49	8,78	1,87	2,19

O maior erro médio observado no simulador FOCOS foi de 1,178 no teste 46, enquanto para o simulador da ACD Labs o maior erro médio ocorreu no teste 26, sendo 4,898. Comparativamente, a média dos erros médios dos programas foram de 0,648 e 1,878 para os simuladores FOCOS e da ACD Labs, respectivamente.

Os simuladores de dados de RMN ^{13}C também foram analisados através de regressão linear múltipla (RLM). Para tanto, foi realizada uma RLM entre os dados previstos pelo simulador FOCOS e os dados experimentais publicados na literatura, onde se obteve um $R = 0,9998$ com um $F = 3.590.214,0$. O mesmo procedimento foi realizado para os dados previstos pelo simulador da ACD Labs e os dados experimentais, obtendo-se um $R = 0,9973$ com um $F = 266.655,6$. Os gráficos dos valores experimentais vs. previstos são apresentados nas Figuras 49 e 50 para os simuladores FOCOS e da ACD Labs, respectivamente.

Analisando-se os dois gráficos de correlação obtidos, pode-se verificar que os dados previstos pelo simulador FOCOS encontram-se mais alinhados à equação da reta, enquanto os dados previstos pelo programa da ACD Labs encontram-se mais dispersos.

Figura 49. Dados experimentais versus dados previstos pelo simulador FOCOS

Figura 50. Dados experimentais versus dados previstos pelo simulador da ACD Labs

Esses dados demonstram que o simulador tridimensional de dados de RMN ^{13}C apresenta um desempenho superior ao de programas comerciais de previsão de dados de RMN ^{13}C . Esses resultados são promissores e demonstram que o modelo de codificação do FOCOS, o qual foi desenvolvido nessa tese e que contempla a estereoquímica dos centros quirais e a influência espacial exercida pelos demais átomos da estrutura sobre um determinado átomo de carbono, mostrou-se bom e adequado para descrever o ambiente químico de um átomo e, portanto, seu deslocamento químico. Além disso, através da metodologia dos focos que descrevem o ambiente químico existente ao redor de um determinado átomo de carbono, os dados de RMN ^{13}C de outras classes de substâncias poderão ser simulados com esta base de dados, uma vez que, se o ambiente químico de um determinado carbono é o mesmo, o seu deslocamento químico deverá se enquadrar em uma determinada faixa de absorção já representada no programa, independente da classe química à qual pertence a substância.

Em alguns casos, o programa FOCOS não apresentou previsão de deslocamentos químicos de certos átomos de carbono das estruturas como, por exemplo, nos testes 40, 43, 47 e 49. Nesses casos, os deslocamentos químicos não previstos referiam-se a carbonos de glicosídeos que se encontram pouco representados na base de dados. Durante a elaboração dessa, privilegiou-se a inserção de estruturas de esteróides não glicosilados, ficando para uma etapa posterior o arquivamento das saponinas esteroidais. Especificamente no teste 49, os dados de RMN ^{13}C não simulados referem-se a átomos de carbono do núcleo aromático ou a átomos influenciados por esses, uma vez que o número de esteróides que apresentam sistemas aromáticos na estrutura é extremamente raro na base de dados, encontrando-se apenas poucos exemplos desses em anéis A do esqueleto estrano.

Vale ressaltar, também, que a base de dados do simulador da ACD Labs apresenta cerca de 2.000 estruturas de produtos naturais dos quais mais de um terço são esteróides, portanto, o bom resultado desse simulador para essa classe de substâncias já era esperado. Entretanto, para testes realizados com outras classes de produtos naturais como, por exemplo, as lactonas sesquiterpênicas¹⁸⁸, observou-se que os erros médios apresentados pelo simulador da ACD Labs foram muito superiores aos observados para esteróides. Nesse último, a boa representação da classe aliada à maior rigidez estrutural em termos conformacionais garantiram o bom desempenho do programa.

7.3. Programa de geração de estruturas - ESTGERA

Analisando os resultados obtidos com o gerador ESTGERA pode-se verificar que, em 58,0% dos casos, o programa forneceu a estrutura correta da substância e, em 36,0% dos casos, o programa não apresentou nenhuma proposta estrutural devido à ausência de um dos anéis da estrutura ou da cadeia lateral. Apenas em 6,0% dos casos, o programa forneceu propostas estruturais incorretas. Verificando detalhadamente os testes onde o programa não apresentou resultados, pode-se verificar que, em todos os casos, o programa selecionou fragmentos estruturais corretos, porém a falta de um dos fragmentos impossibilitou a construção da estrutura completa. Se analisarmos somente os casos onde foram apresentadas propostas estruturais, pode-se verificar que, em 90,6% dos testes, o gerador forneceu a estrutura correta da substância, sendo que desses, em 46,8% dos casos, essa foi a única proposta estrutural exibida. Na Tabela XXIV é apresentado um resumo das análises realizadas através do programa ESTGERA.

Tabela XXIV. Resumo das análises realizadas com o gerador ESTGERA

Teste	Estruturas propostas	Estrutura correta	Justificativa
1	1	Presente	
2	0	Ausente	O anel D e a cadeia lateral não foram identificados
3	2	Presente	
4	0	Ausente	O anel C não foi identificado
5	1	Ausente	O anel A foi identificado erroneamente
6	1	Ausente	O anel A foi identificado erroneamente
7	1	Presente	
8	1	Presente	
9	1	Presente	
10	2	Presente	
11	4	Presente	
12	2	Presente	
13	2	Presente	
14	0	Ausente	A cadeia lateral não foi identificada
15	1	Presente	
16	0	Ausente	O anel D não foi identificado
17	0	Ausente	O anel A não foi identificado
18	0	Ausente	O anel B não foi identificado
19	1	Presente	
20	1	Presente	
21	0	Ausente	Os anéis B e D não foram identificados
22	1	Presente	
23	2	Presente	
24	1	Presente	
25	0	Ausente	O anel D não foi identificado

Tabela XXIV. Continuação

Teste	Estruturas propostas	Estrutura correta	Justificativa
26	4	Presente	
27	0	Ausente	O anel D e a cadeia lateral não foram identificados
28	0	Ausente	O anel A não foi identificado
29	1	Presente	
30	0	Ausente	O anel C não foi identificado
31	2	Presente	
32	1	Presente	
33	0	Ausente	A cadeia lateral não foi identificada
34	1	Presente	
35	1	Ausente	O anel A foi identificado erroneamente
36	1	Presente	
37	0	Ausente	O anel B não foi identificado
38	1	Presente	
39	1	Presente	
40	2	Presente	O ácido carboxílico em C ₁₉ não foi identificado
41	0	Ausente	O anel B não foi identificado
42	0	Ausente	O anel B não foi identificado
43	2	Presente	
44	0	Ausente	A cadeia lateral não foi identificada
45	2	Presente	
46	3	Presente	
47	0	Ausente	O anel A e a cadeia lateral não foram identificados
48	2	Presente	O C ₁₈ não foi identificado
49	0	Ausente	A cadeia lateral e os anéis C e D não foram identificados
50	3	Presente	

Nos casos onde o programa não forneceu propostas estruturais, os fragmentos apresentados servirão como pontos de partida no processo de geração estrutural exaustiva através de geradores comerciais, tais como o MOLGEN¹⁴³. O uso desses fragmentos restringe o número de prováveis estruturas apresentadas pelo gerador além de diminuir o tempo computacional utilizado.

Para os testes onde foram apresentadas mais de uma proposta estrutural cabe ao usuário discernir entre as prováveis substâncias. Nestes casos, quando os dados de RMN ¹³C das propostas estruturais foram simulados através do programa FOCOS, foi possível discernir, em todos os casos, a estrutura correta da substância.

VIII - CONCLUSÕES

O SISTEMAT é um sistema especialista criado com o objetivo principal de auxiliar os espectroscopistas no processo de determinação estrutural de produtos naturais. Esse sistema encontra-se atualmente em fase de aperfeiçoamento, entretanto, os programas aplicativos que compõem o SISTEMAT têm sido consideravelmente testados para a previsão do tipo de esqueleto de produtos naturais tais como monoterpenos, iridóides, sesquiterpenos lactonizados e não lactonizados, diterpenos, triterpenos e, nessa tese, esteróides, tendo apresentado bons resultados em todos os casos.

As análises realizadas durante esse trabalho com esteróides mostram um bom desempenho dos programas aplicativos do SISTEMAT para identificação e previsão dos tipos de esqueletos, com um acerto médio de 90,0%. A partir dos resultados observados, podemos afirmar que o SISTEMAT é uma valiosa ferramenta que pode ser utilizada no processo de determinação de estruturas, atribuição de sinais de RMN ^{13}C e identificação de esqueletos de esteróides, aumentando assim a capacidade de interpretação espectral dos pesquisadores, tendo em vista que a caracterização do tipo de esqueleto de um produto natural desconhecido é o ponto-chave na determinação estrutural dessas substâncias.

O programa SISCONST, além dos prováveis esqueletos, fornece também, para uma dada substância, grandes subestruturas com sinais de RMN ^{13}C atribuídos, cuja sobreposição ou união pode, muitas vezes, conduzir à estrutura completa de uma substância ou pelo menos a estruturas maiores relevantes para o estabelecimento de uma proposta estrutural. Além disso, o SISTEMAT é o único sistema especialista que utiliza dados sobre a origem natural da substância como mais uma ferramenta para previsão do tipo de esqueleto do produto natural isolado.

O processo de desfuncionalização dos dados de RMN ^{13}C é um procedimento realizado mentalmente por espectroscopistas experientes e, a partir dessa metodologia, foi desenvolvido, nesse trabalho, um programa que realiza este procedimento automaticamente. Os resultados obtidos com o programa REGRAS demonstram a importância desse tipo de abordagem para a identificação de esqueletos de terpenóides, uma vez que essa análise restringe o número de prováveis esqueletos.

O desenvolvimento do programa PESQUISA apresentado e utilizado no decorrer desse trabalho, possibilitou o cálculo automatizado da média global dos resultados

fornecidos pelos diversos programas, além da facilidade para o usuário que precisará digitar apenas uma vez cada tipo de dado durante a consulta. Este aplicativo é de fundamental importância para o projeto SISTEMAT uma vez que apresenta uma interface de fácil acesso e utilização para o usuário e, também, possibilitará a distribuição do mesmo para diversos grupos de pesquisa no país.

As informações obtidas a partir dos programas aplicativos, tais como os prováveis tipos de esqueletos para uma substância analisada, as grandes subestruturas, o percentual de reconhecimento de esqueletos e os prováveis grupos substituintes presentes na estrutura, serão utilizadas como grandes restrições por programas comerciais geradores de estruturas, de modo a obter-se um número menor de propostas estruturais e uma redução do tempo de análise computacional. É importante ressaltar aqui que nenhum outro sistema inicia o processo de geração de estruturas a partir do provável esqueleto da substância. Se o SISTEMAT for comparado com outros sistemas especialistas em determinação estrutural, podemos verificar que a grande diferença entre esse e os demais sistemas é que, no primeiro, o usuário só interfere na criação da base de dados e das regras de classificação dos esqueletos, enquanto nos outros sistemas, para que o usuário obtenha respostas compatíveis com estruturas de produtos naturais, é necessária a interferência direta do mesmo e o fornecimento de inúmeras restrições estruturais.

Os testes realizados através do simulador tridimensional de dados de RMN ^{13}C demonstram sua superioridade em relação a um programa comercial de mesma finalidade. A metodologia usada permite prever com elevado grau de segurança e exatidão o deslocamento químico de RMN ^{13}C de esteróides, considerando-se a influência espacial dos grupos ligados a uma determinada distância do carbono analisado, a estereoquímica dos carbonos assimétricos e a conformação da substância, sendo assim, outras classes de produtos naturais serão inseridas em uma próxima etapa e avaliadas através do simulador tridimensional.

Através dessa fase de desenvolvimento do simulador, já é possível auxiliar na atribuição de sinais de novos esteróides estimando seus dados de RMN ^{13}C , prever os deslocamentos químicos de RMN ^{13}C para uma substância publicada sem esses dados, ou atribuir teoricamente deslocamentos químicos publicados na literatura sem atribuição experimental dos sinais. Nos dois últimos casos a utilização do simulador é de extrema validade e importância, uma vez que inúmeros esteróides foram e ainda são publicados na literatura sem dados de RMN ^{13}C ou sem atribuição desses dados. Em ambos os

casos, o número de substâncias geradas com dados de RMN ^{13}C será bastante elevado e, conseqüentemente, todas essas substâncias poderão ser inseridas no sistema especialista SISTEMAT com seus respectivos deslocamentos químicos previstos, aumentando assim a capacidade de identificação de subestruturas e de previsão de esqueletos do sistema quando um novo conjunto de dados for submetido à análise.

As informações obtidas a partir do simulador FOCOS também servirão para validar e/ou descartar propostas estruturais apresentadas por um gerador de estruturas. Sendo assim, os programas aplicativos do SISTEMAT fornecerão possíveis subestruturas e prováveis esqueletos para um determinado conjunto de dados. Essas informações serão usadas como ponto inicial para geração de estruturas. Ao final desse processo, as substâncias apresentadas serão validadas pelo simulador desenvolvido. Esse procedimento, como descrito anteriormente, reduzirá o número de propostas estruturais e o tempo computacional gasto durante uma análise, além de extinguir ou minimizar o problema de explosão combinatória verificado em outros geradores já desenvolvidos para outros sistemas especialistas^{2,35,189,190}.

O gerador de estruturas ESTGERA, desenvolvido para o sistema, utiliza a sobreposição de anéis proveniente de uma análise de faixas de deslocamento químico características de RMN ^{13}C . Essa abordagem mostrou-se muito eficiente na elucidação estrutural dos testes realizados nessa tese, fornecendo um baixo número de substâncias propostas e, em alguns casos, de fragmentos estruturais presentes na estrutura da substância. A partir de uma revisão bibliográfica e do melhor de nosso conhecimento da literatura, é imprescindível ressaltar aqui que esse tipo de abordagem não é utilizada por nenhum outro sistema especialista já desenvolvido, o que demonstra a originalidade do método empregado. Se faz necessário citar também que essa metodologia apesar de aplicada a esteróides, não é restrita a tal classe, podendo ser aplicada a outras classes químicas de produtos naturais. As propostas estruturais exibidas pelo ESTGERA tiveram seus dados de RMN ^{13}C simulados através do programa FOCOS, o qual permitiu o discernimento entre a estrutura correta e as demais propostas estruturais para a substância. Sendo assim, foi possível atingir com sucesso os escopos pretendidos nesse trabalho.

IX - REFERÊNCIAS BIBLIOGRÁFICAS

1. E. Rich, “Inteligência Artificial”; São Paulo, McGraw-Hill, (1988).
2. R.K. Lindsay, B.G. Buchanan, E.A. Feigenbaum, J. Lederberg, “Applications of Artificial Intelligence for Organic Chemistry: the DENDRAL project”; New York, McGraw-Hill, (1980).
3. S.M. Weiss, C.A. Kulikowski, “Guia prático para projetar sistemas especialistas”; Rio de Janeiro, Livros Técnicos e Científicos, (1988).
4. H.M. Cartwright, “Applications of Artificial Intelligence in Chemistry”; New York, Oxford University Press Inc., (1993).
5. D.S.K. Chan, “Expert system for product manufacturability and cost evaluation”; *Materials and Manufacturing Processes*, **17**, 855 (2002).
6. M.G. Abou-Ali, M. Khamis, “TIREDDX: an integrated intelligent defects diagnostic system for tire production and service”; *Expert Systems with Applications*, **24**, 247 (2003).
7. A. Kazaz, “Application of an expert system on the fracture mechanics of concrete”; *Artificial Intelligence Review*, **19**, 177 (2003).
8. I.V. Filis, M. Sabrakos, C.P. Yialouris, A.B. Sideridis, B. Mahaman, “GEDAS: an integrated geographical expert database system”; *Expert Systems with Applications*, **24**, 25 (2003).
9. F. Witlox, “MATISSE: a relational expert system for industrial site selection”; *Expert Systems with Applications*, **24**, 133 (2003).
10. R.O. Duda, P.E. Hart, K. Konolige, R. Reboh, “A computer based consultant for mineral exploration”; Technical Report, SRI International, (1979).
11. I. Roy, B.C. Sarkar, “Prototype development of MODELER: A knowledge based system for identification of iron ore deposit types”; *Journal of Geological Society of India*, **61**, 177 (2003).
12. R.F. Kokaly, D.G. Despain, R.N. Clark, K.E. Livo, “Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data”; *Remote Sensing of Environment*, **84**, 437 (2003).
13. B.D. Mahaman, P. Harizanis, I. Filis, E. Antonopoulou, C.P. Yialouris, A.B. Sideridis, “A diagnostic expert system for honeybee pests”; *Computers and Electronics in Agriculture*, **36**, 17 (2002).

14. D.L. Lia, Z.T. Fu, Y.Q. Duan, "Fish-Expert: a web-based expert system for fish disease diagnosis"; *Expert Systems with Applications*, **23**, 311 (2002).
15. B. Mahaman, H.C. Passam, A.B. Sideridis, C.P. Yialouris, "DIARES-IPM: a diagnostic advisory rule-based expert system for integrated pest management in Solanaceous crop systems"; *Agricultural Systems*, **76**, 1119 (2003).
16. E.H. Shortliffe, "Computer-based medical consultations: MYCIN"; New York, Elsevier, (1976).
17. I. Turkoglu, A. Arslan, E. Ilkay, "An expert system for diagnosis of the heart valve diseases"; *Expert Systems with Applications*, **23**, 229 (2002).
18. I. Turkoglu, A. Arslan, E. Ilkay, "An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks"; *Computers in Biology and Medicine*, **33**, 319 (2003).
19. M. de Santo, M. Molinara, F. Tortorella, M. Vento, "Automatic classification of clustered microcalcifications by a multiple expert system"; *Pattern Recognition*, **36**, 1467 (2003).
20. A. Rabelo Jr., A.R. Rocha, A. Souza, A. Araujo, A. Ximenes, C. Andrade, D. Onnis, I. Olivaes, K. Oliveira, N. Lobo, N. Ferreira, S. Lopes, V. Werneck, "The validation of an expert system for diagnosis of acute myocardial infarction"; In: *Artificial Intelligence in Medicine*, Berlin, Springer, (1997).
21. A. Geissbühler, R.A. Miller, W.W. Stead, "The clinical spectrum of decision-support in oncology with a case report of a real world system"; In: *Artificial Intelligence in Medicine*, Berlin, Springer, (1997).
22. G. Klopman, S.K. Chakravarti, "Structure-activity relationship study of a diverse set of estrogen receptor ligands (I) using MultiCASE expert system"; *Chemosphere*, **51**, 445 (2003).
23. G. Klopman, S.K. Chakravarti, "Screening of high production volume chemicals for estrogen receptor binding activity (II) by the MultiCASE expert system"; *Chemosphere*, **51**, 461 (2003).
24. M. Praisler, J. Van Bocxlaer, A. de Leenheer, D.L. Massart, "Chemometric detection of thermally degraded samples in the analysis of drugs of abuse with gas chromatography-Fourier-transform infrared spectroscopy"; *Journal of Chromatography A*, **962**, 161 (2002).
25. I. Tp, R. Smith, S. Guhan, K. Taksen, M. Vavra, D. Myers, M.T.W. Hearn, "Intelligent automation of high-performance liquid chromatography method

- development by means of a real-time knowledge-based approach”; *Journal of Chromatography A*, **972**, 27 (2002).
26. J.A. Van Leeuwen, B.G.M. Vandeginste, G. Kateman, M. Mulholland, A. Cleland, “An expert system for the choice of factors for a ruggedness test in liquid chromatography”; *Analytica Chimica Acta*, **228**, 145 (1990).
27. A.T. Bozdana, O. Eyercioglu, “Development of an expert system for the determination of injection moulding parameters of thermoplastic materials: EX-PIMM”; *Journal of Materials Processing Technology*, **128**, 113 (2002).
28. L. Mannina, M. Patumi, N. Proietti, D. Bassi, A.L. Segre, “Geographical characterization of Italian extra virgin olive oils using high-field H-1 NMR spectroscopy”; *Journal of Agricultural and Food Chemistry*, **49**, 2687 (2001).
29. Y. Qian, X.X. Li, Y.R. Jiang, Y.Q. Wen, “An expert system for real-time fault diagnosis of complex chemical processes”; *Expert Systems with Applications*, **24**, 425 (2003).
30. K. Nabeshima, T. Suzudo, T. Ohno, K. Kudo, “Nuclear reactor monitoring with the combination of neural network and expert system”; *Mathematics and Computers in Simulation*, **60**, 233 (2002).
31. W.L. Jorgensen, E.R. Laird, A.J. Gushurst, J.M. Fleischer, S.A. Gothe, H.E. Helson, G.D. Paderes, S. Sinclair, “CAMEO: a program for the logical prediction of the products of organic reactions”; *Pure & Applied Chemistry*, **62**, 1921 (1990).
32. E.J. Corey, A.K. Long, S.D. Rubenstein, “Computer-assisted analysis in organic synthesis”; *Science*, **228**, 408 (1985).
33. A. Németh, T. Vidóczy, K. Héberger, Z. Kúti, J. Wágner, “MECHGEN: Computer aided generation and reduction of reaction mechanisms”; *Journal of Chemical Information and Computer Sciences*, **42**, 208 (2002).
34. W. Bremser, “Multidimensional spectroscopy”; *Magnetic Resonance in Chemistry*, **23**, 1056 (1985).
35. R. Attias, “DARC substructure search system: a new approach to chemical information”; *Journal of Chemical Information and Computer Sciences*, **23**, 102 (1983).
36. G. Carrier, A. Panaye, J.E. Dubois, “Topological structural information in the CAS file: statistical occurrences of DARC concentric fragments. Basic carbon substructures”; *Journal of Chemical Information and Computer Sciences*, **30**, 110 (1990).

37. W. Bremser, M. Grzonka, "SpecInfo - A multidimensional spectroscopic interpretation system"; *Mikrochimica Acta*, **II**, 483 (1991).
38. J.-M. Nuzillard, G. Massiot, "Logic for structure determination"; *Tetrahedron*, **47**, 3655 (1991).
39. M.E. Elyashberg, E.R. Martirosian, Y.Z. Karasev, H. Thiele, H. Somberg, "X-PERT: a user-friendly expert system for molecular structure elucidation by spectral methods"; *Analytica Chimica Acta*, **337**, 265 (1997).
40. M.E. Elyashberg, Y.Z. Karasev, E.R. Martirosian, H. Thiele, H. Somberg, "Expert systems as a tool for the molecular structure elucidation by spectral methods. Strategies of solution to the problems"; *Analytica Chimica Acta*, **348**, 443 (1997).
41. M. Badertscher, A. Korytko, K.-P. Schulz, M. Madison, M.E. Munk, P. Portmann, M. Junghans, P. Fontana, E. Pretsch, "Assemble 2.0: a structure generator"; *Chemometrics and Intelligent Laboratory Systems*, **51**, 73 (2000).
42. C. Steinbeck, "SENECA: a platform-independent, distributed, and parallel system for computer-assisted structure elucidation in organic chemistry"; *Journal of Chemical Information and Computer Sciences*, **41**, 1500 (2001).
43. N.A.B. Gray, "Computer-assisted structure elucidation"; New York, John Wiley & Sons, (1986).
44. R.E. Carhart, D.H. Smith, N.A.B. Gray, J.G. Nourse, C. Djerassi, "GENOA: a computer program for structure elucidation utilizing overlapping and alternative substructures"; *Journal of Organic Chemistry*, **46**, 1708 (1981).
45. C. Djerassi, D.H. Smith, C.W. Crandell, N.A.B. Gray, J.G. Nourse, M.R. Lindley, "The DENDRAL project: computational aid to natural products structure elucidation"; *Pure & Applied Chemistry*, **54**, 2425 (1982).
46. I. Kubo, Y.M. Lee, M. Pettei, F. Pilkiewicz, K. Nakanishi, "Potent army worm antifeedants from the east african warburgia plants"; *Journal of the Chemical Society - Chemical Communications*, 1013 (1976).
47. M.A. de Alvarenga, H.E. Gottlieb, O.R. Gottlieb, M.T. Magalhães, V.O. da Silva, "Diasin, a diterpene from *Croton diasii*"; *Phytochemistry*, **17**, 1773 (1978).
48. D.H. Smith, N.A.B. Gray, J.G. Nourse, C.W. Crandell, "The DENDRAL project: recent advances in computer-assisted structure elucidation"; *Analytica Chimica Acta*, **133**, 471 (1981).
49. R. Neudert, M. Penk, "Enhanced structure elucidation"; *Journal of Chemical Information and Computer Sciences*, **36**, 244 (1996).

50. W. Bremser, "HOSE - A novel substructure code"; *Analytica Chimica Acta*, **103**, 355 (1978).
51. M. Will, W. Fachinger, J.R. Richert, "Fully automated structure elucidation - A spectroscopist's dream comes true"; *Journal of Chemical Information and Computer Sciences*, **36**, 221 (1996).
52. M.J. Pestchanker, M.S. Ascheri, O.S. Giordano, "Uspallatine, a pyrrolizidine alkaloid from *Senecio uspallatensis*"; *Phytochemistry*, **24**, 1622 (1985).
53. T. Konoshima, M. Takasaki, M. Kozuka, M. Haruna, K. Ito, J.R. Estes, K. Lee, "Constituents of Rosaceous plants. I. Structures of new triterpenoids from *Cowania mexicana*"; *Chemical & Pharmaceutical Bulletin*, **41**, 1612 (1993).
54. B.D. Christie, M.E. Munk, "The role of two-dimensional Nuclear Magnetic Resonance spectroscopy in computer-enhanced structure elucidation"; *Journal of the American Chemical Society*, **113**, 3750 (1991).
55. C. Peng, S.-G. Yuan, C.-Z. Zheng, Y.-Z. Hui, "Efficient application of 2D NMR correlation information in computer-assisted structure elucidation of complex natural products"; *Journal of Chemical Information and Computer Sciences*, **34**, 805 (1994).
56. C. Steinbeck, "LUCY - A program for structure elucidation from NMR correlation experiments"; *Angewandte Chemie International Edition*, **35**, 1984 (1996).
57. C. Peng, S.-G. Yuan, C.-Z. Zheng, Y.-Z. Hui, H.-M. Wu, K. Ma, X.-W. Han, "Application of expert system CISOC-SES to the structure elucidation of complex natural products"; *Journal of Chemical Information and Computer Sciences*, **34**, 814 (1994).
58. C. Peng, S.-G. Yuan, C.-Z. Zheng, Z. Shi, H. Wu, "Practical computer-assisted structure elucidation for complex natural products: Efficient use of ambiguous 2D NMR correlation information"; *Journal of Chemical Information and Computer Sciences*, **35**, 539 (1995).
59. P.S. Steyn, R. Vleggaar, "A reinvestigation of the structure of monochaetin, a metabolite of *Monochaetia compta*"; *Journal of the Chemical Society Perkin Transactions I*, 1975 (1986).
60. A.M. El-Sayed, M.A. Al-Yahaja, A.H. Shah, R. Hartmann, E. Breitmaier, "Eurabidiol and other sesquiterpenes from *Euryops arabicus*"; *Chemiker-Zeitung*, **114**, 159 (1990).

61. M. Hamonnière, M. Leboeuf, A. Cavé, "Alcaloides aporfiniques et composés terpéniques du *Polyalthia oliveri*"; *Phytochemistry*, **16**, 1029 (1977).
62. F.L. Stahura, J.W. Golden, L. Xue, J. Bajorath, "Distinguishing between natural products and synthetic molecules by descriptor Shannon entropy analysis and binary QSAR calculations"; *Journal of Chemical Information and Computer Sciences*, **40**, 1245 (2000).
63. J.P. Gastmans, M. Furlan, M.N. Lopes, J.H.G. Borges, V.P. Emerenciano, "A inteligência artificial aplicada à química de produtos naturais. O programa SISTEMAT. Parte I - Bases teóricas"; *Química Nova*, **13**, 10 (1990).
64. J.P. Gastmans, M. Furlan, M.N. Lopes, J.H.G. Borges, V.P. Emerenciano, "A inteligência artificial aplicada à química de produtos naturais. O programa SISTEMAT. Parte II - Organização do programa e aplicativos"; *Química Nova*, **13**, 75 (1990).
65. V.P. Emerenciano, G.V. Rodrigues, J.P. Gastmans, "Um método de codificação semi-automático assistido por computador para o SISTEMAT"; *Química Nova*, **16**, 431 (1993).
66. M.J.P. Ferreira, G.V. Rodrigues, V.P. Emerenciano, "MONOREG – an expert system for structural elucidation of monoterpenes"; *Canadian Journal of Chemistry*, **79**, 1915 (2001).
67. G.V. Rodrigues, "Um sistema especialista em determinação estrutural de sesquiterpenos lactonizados. Um enfoque multiespectral aliado a dados botânicos"; Tese de Doutorado, IQ-USP (1996).
68. F.C. Oliveira, M.J.P. Ferreira, C.V. Núñez, G.V. Rodrigues, V.P. Emerenciano, "¹³C NMR spectroscopy of eudesmane sesquiterpenes"; *Progress in Nuclear Magnetic Resonance Spectroscopy*, **37**, 1 (2000).
69. S.A.V. Alvarenga, J.P. Gastmans, G.V. Rodrigues, V.P. Emerenciano, "Ditregra - an auxiliary program for structural determination of diterpens"; *Spectroscopy – An International Journal*, **13**, 227 (1997).
70. P.A.T. Macari, J.P. Gastmans, G.V. Rodrigues, V.P. Emerenciano, "An expert system for structure elucidation of triterpenes"; *Spectroscopy – An International Journal*, **12**, 139 (1994).
71. V.P. Emerenciano, L.D. Melo, G.V. Rodrigues, J.P. Gastmans, "Application of Artificial Intelligence in Organic Chemistry. Part XIX. Pattern recognition and

- structural determination of flavonoids using ^{13}C -NMR spectra"; *Spectroscopy – An International Journal*, **13**, 181 (1997).
72. J.H.G. Borges, "Design e Implementação de um gerador estrutural automático para o projeto SISTEMAT"; Tese de Doutorado, IQ-USP (1998).
73. M. Carabedian, I. Dagane, J.-E. Dubois, "Elucidation by progressive intersection of ordered substructures from carbon-13 nuclear magnetic resonance"; *Analytical Chemistry*, **60**, 2186 (1988).
74. J.W. Blant, J.B. Stothers, " ^{13}C NMR spectra of steroids. A survey and commentary"; *Organic Magnetic Resonance*, **9**, 439 (1977).
75. G.V. Rodrigues, I.P.A. Campos, V.P. Emerenciano, "Applications of artificial intelligence to structure determination of organic compounds. XX. Determination of groups attached to the skeleton of natural products using ^{13}C NMR spectroscopy"; *Spectroscopy – An International Journal*, **13**, 191 (1997).
76. M.J.P. Ferreira, F.C. Oliveira, S.A.V. Alvarenga, P.A.T. Macari, G.V. Rodrigues, V.P. Emerenciano, "Automatic identification by ^{13}C NMR of substituent groups bonded in natural product skeletons"; *Computers & Chemistry*, **26**, 601 (2002).
77. M. Iwashima, K. Nara, K. Iguchi, "New marine steroids, yonasterols, isolated from the Okinawan soft coral, *Clavularia viridis*"; *Steroids*, **65**, 130 (2000).
78. D.L.G. Fromanteau, J.P. Gastmans, S.A. Vestri, V.P. Emerenciano, J.H.G. Borges, "A constraints generator in structural determination by microcomputer"; *Computers & Chemistry*, **17**, 369 (1993).
79. W. Bremser, M. Klier, E. Meyer, "Mutual assignment of subspectra and substructures - A way to structure elucidation by ^{13}C NMR spectroscopy"; *Organic Magnetic Resonance*, **7**, 97 (1975).
80. M.J.P. Ferreira, S.A.V. Alvarenga, P.A.T. Macari, G.V. Rodrigues, V.P. Emerenciano, "A program for terpenoid skeleton prediction based on botanical information"; *Biochemical Systematics and Ecology*, **31**, 25 (2003).
81. V.P. Emerenciano, G.V. Rodrigues, P.A.T. Macari, S.A. Vestri, J.H.G. Borges, J.P. Gastmans, D.L.G. Fromanteau, "Applications d'intelligence artificielle dans la chimie organique. XVII. Nouveaux programmes du projet SISTEMAT"; *Spectroscopy – An International Journal*, **12**, 91 (1994).
82. M.J.P. Ferreira, A.J.C. Brant, G.V. Rodrigues, V.P. Emerenciano, "Automatic identification of terpenoid skeletons through ^{13}C NMR data disfunctionalization"; *Analytica Chimica Acta*, **429**, 151 (2001).

83. E. Pretsch, T. Clerc, J. Seibl, W. Simon, "Tables of spectral data for structural determination of organic compounds"; Berlin, Springer-Verlag, 2nd Edition (1989).
84. A. Fürst, E. Pretsch, "A computer program for the prediction of ^{13}C NMR chemical shifts of organic compounds"; *Analytica Chimica Acta*, **229**, 17 (1990).
85. A. Fürst, E. Pretsch, "Comprehensive parameter set for the prediction of the ^{13}C NMR chemical shifts of sp^3 -hybridized carbon atoms in organic compounds"; *Analytica Chimica Acta*, **233**, 213 (1990).
86. A. Fürst, E. Pretsch, "Parameter set for the prediction of the ^{13}C NMR chemical shifts of sp^2 -hybridized carbon atoms in organic compounds"; *Analytica Chimica Acta*, **248**, 415 (1991).
87. J.T. Clerc, H. Sommerauer, "A minicomputer program based on additivity rules for the estimation of ^{13}C NMR chemical shifts"; *Analytica Chimica Acta*, **95**, 33 (1977).
88. H.N. Cheng, S.J. Ellingsen, "Carbon-13 nuclear magnetic resonance spectral interpretation by a computerized substituent chemical shift method"; *Journal of Chemical Information and Computer Sciences*, **23**, 197 (1983).
89. J. Zupan, M. Novič, S. Bohanec, M. Razinger, L. Lah, M. Tusar, I. Košir, "Expert system for solving problems in carbon-13 nuclear magnetic resonance spectroscopy"; *Analytica Chimica Acta*, **200**, 333 (1987).
90. E. Pretsch, A. Fürst, M. Badertscher, R. Bürgin, M.E. Munk, "C13Shift: A computer program for the prediction of ^{13}C NMR spectra based on an open set of additivity rules"; *Journal of Chemical Information and Computer Sciences*, **32**, 291 (1992).
91. M. Tusar, L. Tusar, S. Bohanec, J. Zupan, " ^1H and ^{13}C NMR spectra simulation"; *Journal of Chemical Information and Computer Sciences*, **32**, 299 (1992).
92. W. Bremser, "Expectation ranges of ^{13}C NMR chemical shifts"; *Magnetic Resonance in Chemistry*, **23**, 271 (1985).
93. G.W. Small, "Database retrieval techniques for carbon-13 nuclear magnetic resonance spectrum simulation"; *Journal of Chemical Information and Computer Sciences*, **32**, 279 (1992).
94. R.C. Schweitzer, G.W. Small, "Performance enhancement of vector-based search systems: Application to carbon-13 nuclear magnetic resonance chemical shift prediction"; *Journal of Chemical Information and Computer Sciences*, **36**, 46 (1996).

95. R.C. Schweitzer, G.W. Small, "Enhanced structural encoding algorithm for database retrievals of carbon-13 nuclear magnetic resonance chemical shifts"; *Journal of Chemical Information and Computer Sciences*, **36**, 310 (1996).
96. R.C. Schweitzer, G.W. Small, "Automated spectrum simulation methods for carbon-13 nuclear magnetic resonance spectroscopy based on database retrieval and model-building strategies"; *Journal of Chemical Information and Computer Sciences*, **37**, 249 (1997).
97. P.C. Jurs, J.W. Ball, L.S. Anker, T.L. Friedman, "Carbon-13 nuclear magnetic resonance spectrum simulation"; *Journal of Chemical Information and Computer Sciences*, **32**, 272 (1992).
98. L.S. Anker, P.C. Jurs, "Prediction of carbon-13 nuclear magnetic resonance chemical shifts by artificial neural networks"; *Analytical Chemistry*, **64**, 1157 (1992).
99. J.W. Ball, P.C. Jurs, "Automated selection of regression models using neural networks for ^{13}C NMR spectral predictions"; *Analytical Chemistry*, **65**, 505 (1993).
100. P. Geladi, B.R. Kowalski, "Partial least-squares regression: a tutorial"; *Analytica Chimica Acta*, **185**, 1 (1986).
101. S. Wold, K. Esbensen, P. Geladi, "Principal component analysis"; *Chemometrics and Intelligent Laboratory Systems*, **2**, 37 (1987).
102. E. Zupan, J. Gasteiger, "Neural Networks for Chemists - An Introduction", Weinheim, VHC-Verlag, (1993).
103. J.P. Doucet, A. Panaye, E. Feuillebois, P. Ladd, "Neural networks and ^{13}C NMR shift prediction"; *Journal of Chemical Information and Computer Sciences*, **33**, 320 (1993).
104. J. Meiler, R. Meusinger, M. Will, "Fast determination of ^{13}C NMR chemical shifts using artificial neural networks"; *Journal of Chemical Information and Computer Sciences*, **40**, 1169 (2000).
105. R.B. Schaller, M.E. Munk, E. Pretsch, "Spectra estimation for computer-aided structure determination"; *Journal of Chemical Information and Computer Sciences*, **36**, 239 (1996).
106. J. Aires-de-Souza, M.C. Hemmer, J. Gasteiger, "Prediction of ^1H NMR chemical shifts using neural networks"; *Analytical Chemistry*, **74**, 80 (2002).

- 107.L.P. Lindeman, J.Q. Adams, "Carbon-13 nuclear magnetic resonance spectrometry- Chemical shifts for the paraffins through C₉"; *Analytical Chemistry*, **43**, 1245 (1971).
- 108.D.H. Smith, P.C. Jurs, "Prediction of ¹³C nuclear magnetic resonance chemical shifts"; *Journal of the American Chemical Society*, **100**, 3316 (1978).
- 109.G.W. Small, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of cycloalkanols with computer-based structural descriptors"; *Analytical Chemistry*, **55**, 1128 (1983).
- 110.G.W. Small, P.C. Jurs, "Data reduction in the simulation of carbon-13 nuclear magnetic resonance spectra of steroids"; *Analytical Chemistry*, **56**, 2307 (1984).
- 111.D.S. Egolf, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of substituted cyclopentanes and cyclopentanols"; *Analytical Chemistry*, **59**, 1586 (1987).
- 112.D.S. Egolf, E.B. Brockett, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of methyl-substituted norbornan-2-ols"; *Analytical Chemistry*, **60**, 2700 (1988).
- 113.G.P. Sutton, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of alkyl-substituted cyclohexanones and decalones"; *Analytical Chemistry*, **61**, 863 (1989).
- 114.M.L. Ranc, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of piperidines"; *Analytical Chemistry*, **61**, 2489 (1989).
- 115.D.S. Egolf, P.C. Jurs, "Structural analysis of polychlorinated biphenyls from ¹³C NMR spectra"; *Analytical Chemistry*, **62**, 1746 (1990).
- 116.G.P. Sutton, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of alkyl-substituted aromatic compounds"; *Analytical Chemistry*, **62**, 1884 (1990).
- 117.G.P. Sutton, L.S. Anker, P.C. Jurs, "Evaluation of automated methods for the selection of models for simulation of carbon-13 nuclear magnetic resonance spectra of keto-steroids"; *Analytical Chemistry*, **63**, 443 (1991).
- 118.M.L. Ranc, P.C. Jurs, "Simulation of carbon-13 nuclear magnetic resonance spectra of quinolines and isoquinolines"; *Analytica Chimica Acta*, **248**, 183 (1991).
- 119.J.W. Ball, L.S. Anker, P.C. Jurs, "Automated model selection for the simulation of carbon-13 nuclear magnetic resonance spectra of cyclopentanones and cycloheptanones"; *Analytical Chemistry*, **63**, 2435 (1991).

- 120.M.L. Ranc, P.C. Jurs, "Simulation of ^{13}C nuclear magnetic resonance spectra of indoles"; *Analytica Chimica Acta*, **280**, 145 (1993).
- 121.J.W. Ball, P.C. Jurs, "Simulation of polysaccharide ^{13}C nuclear magnetic resonance spectra using regression analysis and neural networks"; *Analytical Chemistry*, **65**, 3615 (1993).
- 122.D.L. Clouser, P.C. Jurs, "Simulation of ^{13}C nuclear magnetic resonance spectra of tetrahydropyrans using regression analysis and neural networks"; *Analytica Chimica Acta*, **295**, 221 (1994).
- 123.D.L. Clouser, P.C. Jurs, "Simulation of the ^{13}C nuclear magnetic resonance spectra of trisaccharides using multiple linear regression analysis and neural networks"; *Carbohydrate Research*, **271**, 65 (1995).
- 124.B.E. Mitchell, P.C. Jurs, "Computer assisted simulation of ^{13}C nuclear magnetic spectra of monosaccharides"; *Journal of Chemical Information and Computer Sciences*, **36**, 58 (1996).
- 125.D.L. Clouser, P.C. Jurs, "Simulation of the ^{13}C nuclear magnetic resonance spectra of ribonucleosides using multiple linear regression analysis and neural networks"; *Journal of Chemical Information and Computer Sciences*, **36**, 168 (1996).
- 126.D.L. Clouser, P.C. Jurs, "The simulation of ^{13}C nuclear magnetic resonance spectra of dibenzofurans using multiple linear regression analysis and neural networks"; *Analytica Chimica Acta*, **321**, 127 (1996).
- 127.HyperChem, V 6.03, Hyperclube, Gainesville (Fl.) USA.
- 128.N.L. Allinger, "Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V_1 and V_2 torsional terms"; *Journal of the American Chemical Society*, **99**, 8127 (1977).
- 129.Z. Galdecki, P. Grochulski, Z. Wawrzak, W.L. Duax, M. Numazawa, Y. Osawa, "Structure of 2,4-dibromo-10 β ,17 β -dihydroxy-1,4-estradien-3-one"; *Acta Crystallographica*, **C43**, 967 (1987).
- 130.M.R. Silva, J.A. Paixão, M.-J.M. de Almeida, E.J.T. da Silva, M.L. Sá e Melo, A.S.C. Neves, "3 β ,4 α -dihydroxy-5 β -androstan-17-one"; *Acta Crystallographica*, **C52**, 2892 (1996).
- 131.P. Dastidar, V.N. Shankar, K.M. Madyastha, T.N.G. Row, "Structure of 9 α ,19-dihydroxy-4-androstene-3,17-dione"; *Acta Crystallographica*, **C48**, 486 (1992).

- 132.V.K. Gupta, Rajnikant, K.N. Goswami, "11-ketoprogesterone"; *Acta Crystallographica*, **C50**, 798 (1994).
- 133.D.S. Brown, C.A. Burns, J.C. Gill, B.A. Marples, "Structure of (20R)-20,21-epoxy-19-norpregna-1,3,5(10)-triene-3,17 β -diol"; *Acta Crystallographica*, **C45**, 1324 (1989).
- 134.W.L. Duax, J.F. Griffin, P.D. Strong, Y. Guan, "Structure of 3 β -hydroxy-16 α -methyl-5-pregnen-20-one"; *Acta Crystallographica*, **C45**, 1433 (1989).
- 135.I. Krstanović, N. Cvetković, "Structure of 17 α -hydroxyprogesterone caproate"; *Acta Crystallographica*, **C45**, 478 (1989).
- 136.G. Argay, A. Kálmán, B. Ribár, S. Vladimirov, D. Živanov-Stakić, "Structures of two related bufadienolides: gamabufotalin (3 β ,11 α ,14-trihydroxy-5 β ,14 β -bufa-20,22-dienolide) and arenobufagin (3 β ,11 α ,14-trihydroxy-12-oxo-5 β ,14 β -bufa-20,22-dienolide)"; *Acta Crystallographica*, **C43**, 922 (1987).
- 137.C.A. Mattia, L. Mazzarella, R. Puliti, R. Riccio, L. Minale, "Structure and stereochemistry of (24R)-27-nor-5 α -cholestane-3 β ,4 β ,5,6 α ,7 β ,8,14,15 α ,24-nonaol: a highly hydroxylated marine steroid from the starfish *Archaster typicus*"; *Acta Crystallographica*, **C44**, 2170 (1988).
- 138.A.D. Morales, R.P. Hernández, F.C. Manchado, M.T.H. Fonseca, G. Punte, G. Echeverría, "(25R)-5 β -hydroxyspirost-2-en-6-one"; *Acta Crystallographica*, **C53**, 1705 (1997).
- 139.M. Parvez, V. Fajardo, M. Shamma, "(+)-Jaborosalactone M, a hemiketal withanolide from *Jaborosa magellanica*"; *Acta Crystallographica*, **C44**, 553 (1988).
- 140.J.E. Dubois, M. Carabedian, I. Dagane, "Computer-aided elucidation of structures by ^{13}C NMR – The DARC-EPIOS method: characterization of ordered substructures by correlating the chemical shifts of their bonded carbon atoms"; *Analytica Chimica Acta*, **158**, 217 (1984).
- 141.M. Carabedian, I. Dagane, J.E. Dubois, "Elucidation by progressive intersection of ordered substructures from carbon-13 nuclear magnetic resonance"; *Analytical Chemistry*, **60**, 2186 (1988).
- 142.Advanced Chemistry Development Labs., CNMR Predictor v. 4.5

- 143.C. Benecke, R. Grund, R. Hohberger, A. Kerber, R. Laue, Th. Wieland, "MOLGEN⁺, a generator of connectivity isomers and stereoisomers for molecular structure elucidation"; *Analytica Chimica Acta*, **314**, 141 (1995).
- 144.J. Meiler, E. Sanli, J. Junker, R. Meusinger, T. Lindel, M. Will, W. Maier, M. Köck, "Validation of structural proposals by substructure analysis and ¹³C NMR chemical shift prediction"; *Journal of Chemical Information and Computer Sciences*, **42**, 241 (2002).
- 145.J.P. Gastmans, M.N. Lopes, M. Furlan, V.P. Emerenciano, "L'Etude des signaux de ¹³C par microordinateur un programme simple et rapide"; *Computers & Chemistry*, **14**, 75 (1990).
- 146.V.P. Emerenciano, A.C. Bussolini, G.V. Rodrigues, M. Furlan, D.L.G. Fromanteau, "A pattern recognition method on ¹³C NMR spectroscopy of sesquiterpene skeletal types"; *Spectroscopy – An International Journal*, **11**, 95 (1993).
- 147.M.J.P. Ferreira, G.V. Rodrigues, A.J.C. Brant, V.P. Emerenciano, "REGRAS: an auxiliary program for pattern recognition and substructure elucidation of monoterpenes; *Spectroscopy – An International Journal*, **15**, 65 (2001).
- 148.J.S.L.T. Militão, V.P. Emerenciano, M.J.P. Ferreira, D. Cabrol-Bass, M. Rouillard, "Structure validation in computer-supported structure elucidation: ¹³C NMR shift predictions for steroids"; *Chemometrics and Intelligent Laboratory Systems*, **67**, 5 (2003).
- 149.Atta-ur-Rahman, Dur-e-Shahwar, A. Naz, M.I. Choudhary, "Withanolides from *Withania coagulans*"; *Phytochemistry*, **63**, 387 (2003).
- 150.P. Akbay, I. Çalis, J. Heilmann, O. Sticher, "New stigmastane sterols from *Ajuga salicifolia*"; *Journal of Natural Products*, **66**, 461 (2003).
- 151.A.D. Wright, E. Goclik, G.M. König, "Oxygenated analogues of gorgosterol and ergosterol from the soft coral *Capnella lacertiliensis*"; *Journal of Natural Products*, **66**, 157 (2003).
- 152.B. Jayaprakasam, M.G. Nair, "Cyclooxygenase-2 enzyme inhibitory withanolides from *Withania somnifera*"; *Tetrahedron*, **59**, 841 (2003).
- 153.N.V. Ivanchina, A.A. Kicha, A.I. Kalinovskiy, P.S. Dmitrenok, V.A. Stonik, "Hemolytic steroid disulfates from the far eastern starfish *Pteraster pulvillus*"; *Journal of Natural Products*, **66**, 298 (2003).

- 154.K. Watanabe, Y. Mimaki, H. Sakagami, Y. Sashida, "Bufadienolide and spirostanol glycosides from the rhizomes of *Helleborus orientalis*"; *Journal of Natural Products*, **66**, 236 (2003).
- 155.W.-B. Pan, F.-R. Chang, L.-M. Wei, Y.-C. Wu, "New flavans, spirostanol sapogenins, and a pregnane genin from *Tupistra chinensis* and their cytotoxicity"; *Journal of Natural Products*, **66**, 161 (2003).
- 156.Z. Pongrácz, M. Báthori, G. Tóth, A. Simon, M. Mák, I. Máthé, "9 α ,20-dihydroxyecdysone, a new natural ecdysteroid from *Silene italica* ssp. *nemoralis*"; *Journal of Natural Products*, **66**, 450 (2003).
- 157.W. Wang, F. Li, Y. Park, J. Hong, C.-O. Lee, J.Y. Kong, S. Shin, K.S. Im, J.H. Jung, "Bioactive sterols from the starfish *Certanardoa semiregularis*"; *Journal of Natural Products*, **66**, 384 (2003).
- 158.Y. Li, M. Ishibashi, M. Satake, X. Chen, Y. Oshima, Y. Ohizumi, "Sterol and triterpenoid constituents of *Verbena littoralis* with NGF-potentiating activity"; *Journal of Natural Products*, **66**, 696 (2003).
- 159.J.-H. Sheu, L.-F. Huang, S.-P. Chen, Y.-L. Yang, P.-J. Sung, G.-H. Wang, J.-H. Su, C.-H. Chao, W.-P. Hu, J.-J. Wang, "Hippuristerones E-I, new polyoxygenated steroids from the gorgonian coral *Isis hippuris*"; *Journal of Natural Products*, **66**, 917 (2003).
- 160.A.F. Ahmed, C.-F. Dai, Y.-H. Kuo, J.-H. Sheu, "1 α ,3 β ,5 β -trihydroxy-24-methylenecholestan-6-one: a novel steroid from a soft coral *Sinularia gibberosa*"; *Steroids*, **68**, 377 (2003).
- 161.Y. Hayakawa, K. Furihata, K. Shin-ya, T. Mori, "Gymnasterol, a new antitumor steroid against IGF-dependent cells from *Gymnascella dankaliensis*"; *Tetrahedron Letters*, **44**, 1165 (2003).
- 162.J.-Q. Gu, W. Li, Y.-H. Kang, B.-N. Su, H.H.S. Fong, R.B. van Breemen, J.M. Pezzuto, A.D. Kinghorn, "Minor withanolides from *Physalis philadelphica*: Structures, quinine reductase induction activities, and liquid chromatography (LC)-MS-MS investigation as artifacts"; *Chemical & Pharmaceutical Bulletin*, **51**, 530 (2003).
- 163.P. Shen, S.-L. Wang, X.-K. Liu, C.-R. Yang, B. Cai, X.-S. Yao, "Steroidal saponins from rhizomes of *Tupistra wattii* Hook. f."; *Chemical & Pharmaceutical Bulletin*, **51**, 305 (2003).

- 164.T. Mori, K. Shin-Ya, K. Takatori, M. Aihara, Y. Hayakawa, "Byssochlamysol, a new antitumor steroid against IGF-1-dependent cells from *Byssochlamys nivea*"; *The Journal of Antibiotics*, **56**, 6 (2003).
- 165.A.M. Abou-Douh, "New withanolides and other constituents from the fruit of *Withania somnifera*"; *Archives der Pharmazie - Pharmaceutical and Medicinal Chemistry*, **6**, 267 (2002).
- 166.V. Bultel-Poncé, S. Etahiri, M. Guyot, "New ketosteroids from the red alga *Hypnea musciformis*"; *Bioorganic & Medicinal Chemistry Letters*, **12**, 1715 (2002).
- 167.M. Ye, J. Han, H. Guo, D. Guo, "Structural determination and complete NMR spectral assignments of a new bufadienolide glycoside"; *Magnetic Resonance in Chemistry*, **40**, 786 (2002).
- 168.S.-S. Yin, A.P. Scott, W. Li, "Pheromones of the male sea lamprey, *Petromyzon marinus* L.: structural studies on a new compound, 3-keto allocholic acid, and 3-keto petromyzonol sulfate"; *Steroids*, **68**, 297 (2003).
- 169.P. Akbay, J. Gertsch, I. Çalis, J. Heilmann, O. Zerbe, O. Sticher, "Novel antileukemic sterol glycosides from *Ajuga salicifolia*"; *Helvetica Chimica Acta*, **85**, 1930 (2002).
- 170.A. Suksamrarn, A. Jankam, B. Tarnchompoo, S. Putchakarn, "Ecdysteroids from a *Zoanthus* sp."; *Journal of Natural Products*, **65**, 1194 (2002).
- 171.G. Santafé, V. Paz, J. Rodríguez, C. Jiménez, "Novel cytotoxic oxygenated C29 sterols from the Colombian marine sponge *Polymastia tenax*"; *Journal of Natural Products*, **65**, 1161 (2002).
- 172.W. Zhang, W.K. Liu, C.-T. Che, "Polyhydroxylated steroids and other constituents of the soft coral *Nephthea chabroli*"; *Chemical & Pharmaceutical Bulletin*, **51**, 1009 (2003).
- 173.H.A. Hassan, A.I. Hamed, N.A. El-Emary, I.V. Springuel, H. Mitome, H. Miyaoka, "Pregnene derivatives from *Solenostemma argel* leaves"; *Phytochemistry*, **57**, 507 (2001).
- 174.X.-H. Zhu, M. Takagi, T. Ikeda, K. Midzuki, T. Nohara, "Withanolide-type steroids from *Solanum cilistum*"; *Phytochemistry*, **56**, 741 (2001).
- 175.S. Weber, J. Puripattavong, V. Brecht, A.W. Frahm, "Phytochemical investigation of *Aglaia rubiginosa*"; *Journal of Natural Products*, **63**, 636 (2000).
- 176.G.R. Pettit, A. Numata, G.M. Cragg, D.L. Herald, T. Takada, C. Iwamoto, R. Riesen, J.M. Schmidt, D.L. Doubek, A. Goswami, "Isolation and structures of

- schleicherastatins 1-7 and schleicheols 1 and 2 from the teak forest medicinal tree *Schleichera oleosa*"; *Journal of Natural Products*, **63**, 72 (2000).
- 177.C. Anta, N. González, J. Rodríguez, C. Jiménez, "A new secosterol from the Indonesian octocoral *Pachyclavularia violacea*"; *Journal of Natural Products*, **65**, 1357 (2002).
- 178.Z.-H. Lei, H. Nakayama, A. Kuniyasu, B.-S. Tai, T. Nohara, "Cardiac glycosides from *Erysimum cheiranthoides*"; *Chemical & Pharmaceutical Bulletin*, **50**, 861 (2002).
- 179.M. Iwashima, K. Nara, Y. Nakamichi, K. Iguchi, "Three new chlorinated marine steroids, yonarasterols G, H and I, isolated from the Okinawan soft coral, *Clavularia viridis*"; *Steroids*, **66**, 25 (2001).
- 180.P.A. Leone, J. Redburn, J.N.A. Hooper, R.J. Quinn, "Polyoxygenated *Dysidea* sterols that inhibit the binding of [I125] IL-8 to the human recombinant IL-8 receptor type A"; *Journal of Natural Products*, **63**, 694 (2000).
- 181.M.A.-A. Al-Yahya, E. Abdel-Sattar, E. Guittet, "Pregnane glycosides from *Caralluma russeliana*"; *Journal of Natural Products*, **63**, 1451 (2000).
- 182.W. Oleszek, M. Sitek, A. Stochmal, S. Piacente, C. Pizza, P. Cheeke, "Steroidal saponins of *Yucca schidigera* Roezl."; *Journal of Agricultural and Food Chemistry*, **49**, 4392 (2001).
- 183.A. Farooq, S. Tahara, "Biotransformation of testosterone and pregnenolone catalyzed by the fungus *Botrytis cinerea*"; *Journal of Natural Products*, **63**, 489 (2000).
- 184.M. Kuroda, Y. Mimaki, Y. Sashida, T. Yamori, T. Tsuruo, "Galtonioside A, a novel cytotoxic cholestane glycoside from *Galtonia candicans*"; *Tetrahedron Letters*, **41**, 251 (2000).
- 185.R.I. Misico, R.R. Gil, J.C. Oberti, A.S. Veleiro, G. Burton, "Withanolides from *Vassobia lorentzii*"; *Journal of Natural Products*, **63**, 1329 (2000).
- 186.M.C. Tettamanzi, A.S. Veleiro, J.R. de la Fuente, G. Burton, "Withanolides from *Salpichroa organifolia*"; *Journal of Natural Products*, **64**, 783 (2001).
- 187.A. Rueda, E. Zubía, M.J. Ortega, J. Salvá, "Structure and cytotoxicity of new polyhydroxylated sterols from the Caribbean gorgonian *Plexaurella grisea*"; *Steroids*, **66**, 897 (2001).

-
- 188.F.M.M. Magri, J.S.L.T. Militão, M.J.P. Ferreira, V.P. Emerenciano, "A new program to ^{13}C NMR spectrum prediction based on tridimensional models"; *Spectroscopy – An International Journal*, **15**, 99 (2001).
- 189.G. Massiot, J.M. Nuzillard, "Computer-assisted elucidation of structures of natural products"; *Phytochemical Analysis*, **3**, 153 (1992).
- 190.K.A. Blinov, D. Carlson, M.E. Elyashberg, G.E. Martin, E.R. Martirosian, S. Molodtsov, A.J. Williams, "Computer-assisted structure elucidation of natural products with limited 2D NMR data: application of the StrucEluc system"; *Magnetic Resonance in Chemistry*, **41**, 359 (2003).