

# *Cryptococcus neoformans* Chemotyping by Quantitative Analysis of <sup>1</sup>H Nuclear Magnetic Resonance Spectra of Glucuronoxylomannans with a Computer-Simulated Artificial Neural Network

ROBERT CHERNIAK,<sup>1\*</sup> HOMAYOUN VALAFAR,<sup>2</sup> LAURA C. MORRIS,<sup>1</sup> AND FARAMARZ VALAFAR<sup>2</sup>

*Department of Chemistry, Laboratory for Biological and Chemical Sciences, Georgia State University, Atlanta,<sup>1</sup> and Complex Carbohydrate Research Center, University of Georgia, Athens,<sup>2</sup> Georgia*

Received 15 September 1997/Returned for modification 11 December 1997/Accepted 30 December 1997

The complete assignment of the proton chemical shifts obtained by nuclear magnetic resonance (NMR) spectroscopy of de-O-acetylated glucuronoxylomannans (GXMs) from *Cryptococcus neoformans* permitted the high-resolution determination of the total structure of any GXM. Six structural motifs based on an  $\alpha$ -(1→3)-mannotriose substituted with variable quantities of 2-O- $\beta$ - and 4-O- $\beta$ -xylopyranosyl and 2-O- $\beta$ -glucopyranosyluronic acid were identified. The chemical shifts of only the anomeric protons of the mannosyl residues served as structure reporter groups (SRG) for the identification and quantitation of the six triads present in any GXM. The assigned protons for the mannosyl residues resonated at clearly distinguishable positions in the spectrum and supplied all the information essential for the assignment of the complete GXM structure. This technique for assigning structure is referred to as the SRG concept. The SRG concept was used to analyze the distribution of the six mannosyl triads of GXMs obtained from 106 isolates of *C. neoformans*. The six mannosyl triads occurred singularly or in combination with one or more of the other triads. The identification and quantitation of the SRG were simplified by using a computer-simulated artificial neural network (ANN) to automatically analyze the SRG region of the one-dimensional proton NMR spectra. The occurrence and relative distribution of the six mannosyl triads were used to chemotype *C. neoformans* on the basis of subtle variations in GXM structure determined by analysis of the SRG region of the proton NMR spectrum by the ANN. The data for the distribution of the six SRGs from GXMs of 106 isolates of *C. neoformans* yielded eight chemotypes, Chem1 through Chem8.

*Cryptococcus neoformans* is an opportunistic pathogenic yeast. In most healthy individuals the organism does not progress beyond the lungs, its normal portal of entry. Hematogenous dissemination occurs in patients who are immunosuppressed due to immunotherapy or some form of immunodeficiency (38, 44, 52). AIDS patients are vulnerable to opportunistic infections due to the precipitous decline in the competency of the immune system (35). Consequently, the occurrence of cryptococcosis has increased dramatically, concomitantly with the emergence of AIDS (24, 41, 67). Despite antifungal therapy, AIDS patients with cryptococcosis are extremely susceptible to relapse. With the exception of *Pneumocystis carinii*, cryptococcosis is the most common invasive life-threatening fungal infection associated with AIDS (27); it is a leading cause of death in these patients (25). The yeast has a propensity to involve the central nervous system, where it causes meningoencephalitis (26, 48). Other frequent sites of infection are the liver and spleen. Diagnosis of cryptococcosis in an AIDS patient may portend the progression to further immunosuppression and to a more severe phase of the disease (47). The insidious association of cryptococcosis with AIDS has resulted in increased attempts to define the virulence factors associated with the yeast and the host at the molecular level.

*C. neoformans* is atypical of pathogenic fungi in that it produces a well-developed polysaccharide capsule (18, 60). The major capsular polysaccharide is glucuronoxylomannan (GXM),

which is an important contributor to the virulence of *C. neoformans* (60). GXM is antiphagocytic (20, 36) and poorly immunogenic (37, 45), and acapsular strains have diminished virulence (32). In vitro, GXM inhibits leukocyte migration (28), enhances human immunodeficiency virus (HIV) infection in human lymphocytes (46, 47), induces the release of tumor necrosis factor alpha by peripheral-blood mononuclear cells (21, 42), and promotes L-selectin shedding from neutrophils (28).

Two distinct varieties of *C. neoformans* have been described: *C. neoformans* var. *neoformans* and *C. neoformans* var. *gattii* (39). The two varieties were subdivided serologically into four serotypes based on the reactivities of whole yeast cells to polyclonal sera that were prepared by selective absorption with whole yeast cells (31, 33, 68). The variety *neoformans* consists of serotypes A, D, and A/D, and the variety *gattii* consists of serotypes B and C. Most cryptococcal infections in AIDS patients are due to *C. neoformans* var. *neoformans*, with serotype A comprising the majority of isolates (19, 40, 51, 57). Monoclonal antibodies and factor sera react specifically with GXMs. These observations substantiate the role GXMs play in conferring serotypes on *C. neoformans*.

The presence of at least eight antigenic factors, distributed among the serotypes of *C. neoformans*, has been proposed based on the reactivities of factor sera in yeast cell agglutination reactions (33). No structure corresponding to any one of the eight antigenic factors has been delineated.

The typical GXM consists of a linear (1→3)- $\alpha$ -D-mannopyranan bearing  $\beta$ -D-xylopyranosyl (Xylp),  $\beta$ -D-glucopyranosyluronic acid (Glc<sub>p</sub>A), and 6-O-acetyl substituents (18, 60, 63). The disposition of the O-acetyl substituents is the major determinant of the antigenic activity observed among GXMs

\* Corresponding author. Mailing address: Department of Chemistry, Georgia State University, University Plaza, Atlanta, GA 30303. Phone: (404) 651-3868. Fax: (404) 651-1416. E-mail: cherniak@gsu.edu.

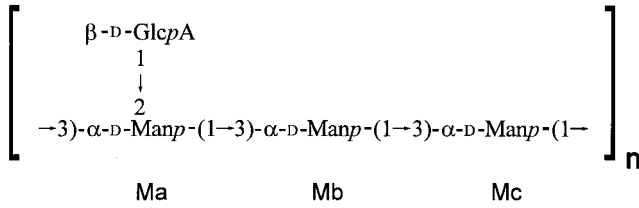


FIG. 1. The core repeating structure of GXM.

obtained from all serotypes (A, B, C, D, and A/D) (15). Generally, the serological activity observed with type-specific antibody is lost after de-O-acetylation (15, 36). Despite considerable structural and antigenic diversity, a simple structural relationship exists between GXMs of reference isolates for the four serotypes. They are all comprised of a core repeating unit (Fig. 1) to which (1→2)-linked and (1→4)-linked β-D-Xylp units are added in increments of one to four residues. In this way, explicit molar ratios of Xyl/Man/GlcA in serotypes D, A, B, and C have been assigned as 1:3:1, 2:3:1, 3:3:1, and 4:3:1, respectively (Fig. 2, M1 through M4) (18). GXM from serotypes A and D are mainly substituted with Xylp at O-2, whereas GXM from serotypes B and C are substituted at O-2 and at O-4. Additional analytical data show that the precise molar ratios and typical substitution patterns, as proposed in the original models of GXM structure, are an oversimplifica-

tion. In addition, substituent dispositions previously thought to be characteristic of one serotype have been identified in heterologous isolates (23, 24, 62, 63).

The primary sequences of de-O-acetylated GXMs from serotypes A, B, C, and D were determined previously by two-dimensional (2D) nuclear magnetic resonance (NMR) spectroscopy (11, 56, 58, 59), and the assignments of the <sup>1</sup>H and <sup>13</sup>C chemical shifts were reported. These data, in conjunction with earlier chemical and NMR analyses of GXMs, document the existence of antigenic multiplicity among the serotypes, particularly in serotypes A and C (11, 24, 62). GXMs within a particular serotype cannot be subdivided serologically because factor sera and monoclonal antibodies have low discriminatory power for determining structural heterogeneity (15, 17). Consequently, the variation in GXM structure, as determined serologically, cannot be correlated to the range of virulence observed among *C. neoformans* isolates.

The complete assignment of the proton chemical shifts for the structural elements present in various GXMs permits the high-resolution determination of the total structure of any GXM. In addition, the primary structural assignment of GXM can be made by using only the anomeric protons of the mannose residues. This is based on the fact that the assigned proton chemical shifts for the mannose residues resonate at clearly distinguishable positions in the spectrum and supply all the information essential for the assignment of the complete

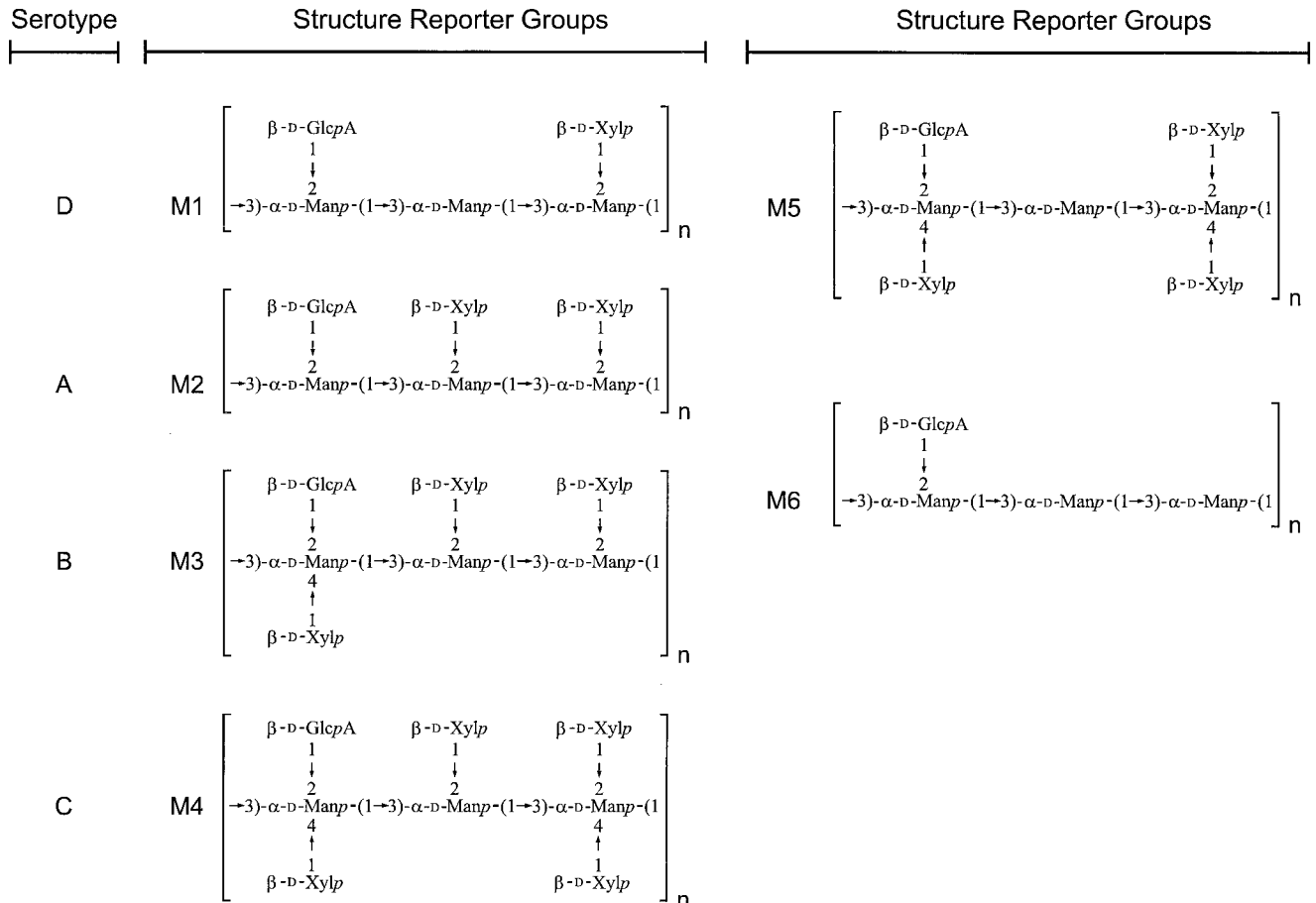


FIG. 2. The six mannosyl triads found in GXMs of *C. neoformans*.

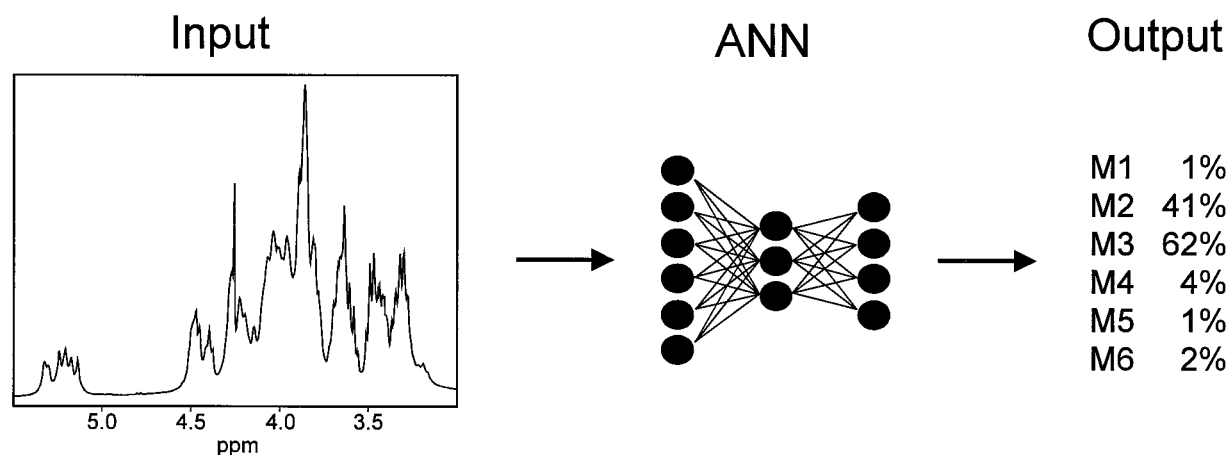


FIG. 3. ANN configuration for the analysis of the GXM NMR data for determining the chemotype of *C. neoformans*.

structure. This technique is referred to as the structural reporter group concept (SRG) (65, 66).

The identification and quantitation of the SRG can be simplified by using a computer-simulated artificial neural network (ANN) (53) to automatically analyze 1D proton NMR spectra as illustrated in Fig. 3. The 1D proton NMR spectrum serves as the input data for the ANN, and the ANN produces an output that represents the relative ratio of each SRG. The selection of an ANN to predict the chemotype of *C. neoformans* is not arbitrary. ANNs are extremely flexible in design, size, and method of training. The ANN is trained with data (1D proton NMR spectra) that define the structure of the polysaccharides obtained by conventional analytical methods.

ANN is a network that consists of simple processing elements (PEs) and connecting elements that provide communication between the PEs (53). The PEs are also referred to as "artificial neurons" or simply "neurons," and the connecting elements are referred to as "synaptic pathways" or "synaptic connections" (53).

ANNs are ideal for extracting distinguishing features from complex data patterns. ANNs have been used for a variety of classification and identification tasks, including the identification of NMR spectra of polysaccharides and other spectral data (43, 49, 50, 54, 61), the diagnosis of diseases (1–8, 14, 15), and other classification problems (8–10, 13, 29, 55, 64). This article describes the development of a computer-simulated ANN for the quantitative analysis of GXM fine structure by using the proton NMR SRG concept. The data were used to develop a chemotyping system based on the quantitation of the subtle variations in GXM structure that occur in *C. neoformans* isolates. The ANN is potentially a powerful tool for investigating the impact that the variation in the fine structure of GXM from any particular isolate has on the effects observed with *in vitro* and *in vivo* biological systems.

#### MATERIALS AND METHODS

**GXM.** *C. neoformans* isolates used in this study are listed in Table 1. Many GXMs were available from earlier studies (11, 15, 23, 24, 58, 59, 62). GXMs from *C. neoformans* isolates not studied previously were prepared as described elsewhere (56). A streamlined method for the isolation of GXM from selected isolates was used also. *C. neoformans* isolates were expanded in 50 ml of the chemically defined liquid medium (22). After 5 days, the culture was autoclaved for 25 min at 121°C and the cells were removed by centrifugation (at 18,000 × *g*) for 1 h. The culture supernatant (~50 ml) was adjusted to 0.2 M NaCl, and 0.15 g of cetyltrimethyl ammonium bromide (CTAB) was added to the stirred solution at 23°C. A 0.05% solution of CTAB (100 ml) was added slowly, with stirring. The precipitate was recovered by centrifugation (at 5,000 × *g*) for 15 min at 23°C. The

pellet was triturated with 10% ethanol, and the suspension was centrifuged as described above. The pellet was dissolved in 1 M NaCl (25 ml) by stirring overnight. GXM was precipitated by the slow addition of 3 volumes of 95% ethanol, and the flask was placed at 4°C. The GXM was recovered by centrifugation (at 5,000 × *g*) for 15 min at 4°C. All preparations of GXM were dissolved in 1 M NaCl (~25 ml), treated by ultrasonic irradiation (with a Branson Sonifier, model 450) for 2 h at 80% power and 40% pulse at a temperature below 20°C, dialyzed, and recovered by lyophilization. A portion of each GXM was de-O-acetylated at pH 11 (NH<sub>4</sub>OH) for 24 h at 23°C (24), dialyzed, and lyophilized. De-O-acetylated GXMs previously treated by ultrasonic irradiation were used in all subsequent NMR experiments. The apparent average molecular mass of the GXMs used for NMR analysis was 1.2 × 10<sup>5</sup> Da.

GXMs isolated in prior studies and those specifically selected for use in this investigation are referred in Table 1. The serotype of each GXM, whether determined previously or by dot enzyme assay (16) as part of this study, is given in Table 1.

**NMR spectroscopy.** De-O-acetylated GXMs (~10 mg) were exchanged in 99.9% D<sub>2</sub>O and lyophilized. Each sample was dissolved in 0.80 ml of 99.96% D<sub>2</sub>O, then filtered through a Millipore MILLEX-GS 0.22-μm-pore-size filter, and the filtrate was transferred into a 5-mm-diameter NMR tube (Wilmad 528-PP). All <sup>1</sup>H NMR experiments were performed at 80°C on a Varian VXR-400 or on a Varian Unity+500 spectrometer equipped with a 5-mm <sup>1</sup>H/<sup>19</sup>F probe. <sup>1</sup>H chemical shifts were measured relative to internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate taken as 0.00 ppm. The data were processed off-line by using a FELIX 2.30 software package (Biosym/Molecular Simulations, San Diego, Calif.) on a Silicon Graphics Indy workstation. Each spectrum was resolution enhanced by applying a sine bell window function over all real data points. The portion of each proton spectrum between 5.00 and 5.40 ppm, where only the mannosyl residues resonate, was analyzed separately. This is the SRG that was used to identify the specific saccharide sequences present in a particular GXM.

Quantitation of each proton NMR signal in the SRG region required the analysis of the NMR spectrum prior to resolution enhancement. The proton NMR data were processed without a window function, and the full spectrum was saved as a FELIX ASCII file. The FELIX ASCII file was transferred to a desktop computer, where it was converted to ordered-pair data (chemical shift, intensity). The ordered-pair data (*x*, *y*) were imported into PeakFit 4.0 (Jandel Scientific, Inc., San Rafael, Calif.). Data in the SRG region (5.0 to 5.4 ppm) were selected, and the remaining data were discarded. The data in the SRG region were analyzed with the Autofit Peaks I selection (by using Lorentzian functions). The peaks designated by PeakFit were processed manually based on the assignment of the chemical shifts of the SRG regions, determined previously from the resolution-enhanced spectra (see above). Rough fitting was done visually after all the peaks were selected. Final data analysis was done with the Marquardt-Levenberg fitting algorithm. The chemical shifts of the resolved peaks and their corresponding areas for GXM were saved as an ASCII file. The data were used to quantitate the presence of particular SRGs in GXM.

ANN. In this study we used a universal ANN design that consisted of a two-stage feed-forward neural-network topology and the standard, backpropagation training algorithm (53), as well as a modified version. Some parameters, such as hidden layer size, step size, etc. (53) were optimized experimentally. The final step was to train the ANN by using the optimum parameters.

FELIX ASCII files representing the <sup>1</sup>H NMR spectra for GXMs obtained from 106 individual *C. neoformans* isolates were divided into a training set (69 spectra) and a testing set (37 spectra). The SRG region of the NMR data was used for the training set. The results of the PeakFit analysis (percent each SRG)

TABLE 1. *C. neoformans* isolates used in this study

Isolate	Serotype	Source
6	A	T. G. Mitchell (Duke University)
18	B/C	K. J. Kwon-Chung (National Institutes of Health)
34	C	K. J. Kwon-Chung
98	A	T. G. Mitchell
110	A/D	T. G. Mitchell
118	A	T. G. Mitchell
127	D	E. Jacobson (Medical College of Virginia)
132	A/D	T. Shinoda (Meiji College of Pharmacy, Tokyo, Japan)
150	A	T. G. Mitchell
184	B	K. J. Kwon-Chung
194	A	T. G. Mitchell
196	A	K. J. Kwon-Chung
201	A	K. J. Kwon-Chung
298	C	K. J. Kwon-Chung
298 Warty	B>C <sup>a</sup>	R. Cherniak (Georgia State University); derived from 298
371	A	K. J. Kwon-Chung
401	C	K. J. Kwon-Chung
409	B	E. Reiss (Centers for Disease Control and Prevention)
430	D	K. J. Kwon-Chung
1254	D	E. Reiss
3168 <sup>b</sup>	D	E. Reiss
3183	C	E. Reiss
3939	B	E. Reiss
9375	D	H. J. Shadomy (Medical College of Virginia)
9759	A	E. Reiss
9759a	A	R. Cherniak; subculture of 9759
24067 a	D	A. Casadevall (Albert Einstein College of Medicine)
24067 b	D	A. Casadevall
24067 c	D	A. Casadevall
24067 d1	D	A. Casadevall
24067 d2	D	A. Casadevall
24067 e	D	A. Casadevall
24067 f1	D	A. Casadevall
24067 f2	D	A. Casadevall
24067 g	D	A. Casadevall
A104	A	T. R. Kozel (University of Nevada, Reno)
A271	A	T. R. Kozel
A288	A	T. R. Kozel
A289	A	T. R. Kozel
A24064	A	T. R. Kozel
ATCC 6352	A	F. I. Molina (American Type Culture Collection)
ATCC 24065	B	F. I. Molina
ATCC 24066	B/C	F. I. Molina
ATCC 28958	D	F. I. Molina
ATCC 32269	B	F. I. Molina
ATCC 32608	C	F. I. Molina
ATCC 34875	D	F. I. Molina
ATCC 34878	B	F. I. Molina
ATCC 34880	C	F. I. Molina
ATCC 48184	A/D	F. I. Molina
ATCC 62066	A	F. I. Molina
ATCC 90126	A/D	F. I. Molina
AW	A	S. Levitz (Boston University Medical Center)
B3502	D	K. J. Kwon-Chung
B4538	A	K. J. Kwon-Chung
C917	C	T. R. Kozel
C24066	C	T. R. Kozel
C24066T	C	R. Cherniak; derived from C24066
C24066W	B/C	R. Cherniak; derived from C24066

Continued

TABLE 1—Continued

Isolate	Serotype	Source
CDC92-64	A/B/D	M. E. Brandt (Centers for Disease Control and Prevention)
CDC92-159	D	M. E. Brandt
CDC92-162	A/B/D	M. E. Brandt
CDC92-187	A/B/D	M. E. Brandt
CDC92-232	D	M. E. Brandt
D52	D	T. R. Kozel
D529	D	T. R. Kozel
D24067	D	T. R. Kozel
F12	D	E. Jacobson
F27	A	J. Murphy (University of Oklahoma)
F113	A	J. Murphy
F117	A	J. Murphy
Houpt 24066	B/C	T. R. Kozel
J9a	D	A. Casadevall
J9b	D	A. Casadevall
J9c	D	A. Casadevall
J9d	D	A. Casadevall
J11a	A	A. Casadevall
J11b	A	A. Casadevall
J22a	A	A. Casadevall
J22b	A	A. Casadevall
Kuhl 94-8069	B	J. Murphy
M0010	A	T. G. Mitchell
M0013	A	T. G. Mitchell
M0024	A/D	T. G. Mitchell
M0025	A	T. G. Mitchell
M0027	A	T. G. Mitchell
M0029	A	T. G. Mitchell
M0046	A	T. G. Mitchell
M0048	A	T. G. Mitchell
M0050	D	T. G. Mitchell
M0051	A	T. G. Mitchell
M0053	A	T. G. Mitchell
M0062	A	T. G. Mitchell
M0095	A	T. G. Mitchell
MU-1	A	R. Cherniak; derived from 9759
NCPF 3168	A	A. Verheul (University Hospital, Utrecht, The Netherlands)
NIH 37	A	A. Verheul
NU-2	A	J. Murphy
P0	ND <sup>c</sup>	A. Casadevall
R1064	A	K. J. Kwon-Chung
SB4a	A>D	A. Casadevall
SB4b	A>D	A. Casadevall
SB4c	D	A. Casadevall
SB6a	D	A. Casadevall
SB6b	A/D	A. Casadevall
TYCC 38-602	A	K. J. Kwon-Chung

<sup>a</sup> B>C, activity with B greater than activity with C.<sup>b</sup> Submitted to us as serotype A but chemically and serologically identical to serotype D.<sup>c</sup> ND, not determined.

were used as the desired output data. Each file was submitted individually, via the Complex Carbohydrate Research Center's web site (<http://www.crcr.uga.edu>) by using a web browser (such as Netscape Navigator), for analysis by the trained neural network. A report page containing the percentage of each SRG in the sample data was generated by the neural network. The output of the trained neural network was compared to the data obtained independently by PeakFit.

The initial experiments used the standard backpropagation training algorithm described by Rumelhart and McClelland (53). However, a modification to the definition of the error function improved the performance of the ANN. This modification included an additional mathematical term that required that the sum of the SRG ratios add up to 100%. The final ANN was trained with this modification. The trained ANN was cross-validated (53) with <sup>1</sup>H NMR data which were not part of the training set. The purpose of cross-validation was to determine when to terminate training.



TABLE 2. Proton NMR chemical shifts of the SRG triads of *C. neoformans* GXMs

SRG	No. <sup>a</sup>	Chemical shift (ppm)		
		Ma <sup>b</sup>	Mb	Mc
M1	32	5.241 ± 0.002	5.193 ± 0.003 <sup>c</sup>	5.193 ± 0.002 <sup>c</sup>
M2	56	5.244 ± 0.002	5.306 ± 0.004	5.175 ± 0.003
M3	45	5.208 ± 0.003	5.325 ± 0.005	5.138 ± 0.004
M4	7	5.243 ± 0.003	5.301 ± 0.005	5.075 ± 0.002
M5	26	5.217 ± 0.002	5.192 ± 0.003 <sup>d</sup>	5.192 ± 0.003 <sup>d</sup>
M6	47	5.260 ± 0.002	5.192 ± 0.002 <sup>e</sup>	5.117 ± 0.002

<sup>a</sup> Number of GXMs analyzed to determine the chemical shifts of the SRG.

<sup>b</sup> Nomenclature is as presented in Fig. 1.

<sup>c</sup> In the M1 SRG, Ma and Mb are degenerate.

<sup>d</sup> In the M5 SRG, Mb and Mc are degenerate and overlap with Mb and Mc of the M1 SRG.

<sup>e</sup> In the M6 SRG, Mb is degenerate with Mb and Mc of the M1 and M5 SRGs.

## RESULTS

**SRGs.** The complete assignment of the proton NMR chemical shifts for all the sugar residues present in GXMs selected from a broad range of *C. neoformans* isolates has been described previously (11, 12, 56, 58, 59). Based on the analysis of these data, we identified six recurring triad structures. The first mannosyl residue from the nonreducing end of a triad substituted at O-2 with β-D-glucuronic acid was always labeled Ma (see Fig. 1). For the development of the *C. neoformans* chemotyping system, the six structural triads were designated M1 through M6 (Fig. 2). The proton chemical shifts of the mannosyl residues characteristic of the six triads were obtained from our previous studies. The six triads, M1 through M6, and their characteristic chemical shifts (Table 2), constitute the SRG region of GXM. The development of the chemotyping system for *C. neoformans* was predicated on our ability to identify and quantify the SRGs for GXM by using high-field proton NMR.

**Identification of SRGs by using proton NMR.** The 1D proton spectrum of each GXM was recorded as illustrated for *C. neoformans* isolate 150 in Fig. 4a (1D full spectrum with SRG region marked). The SRG region of the proton spectrum was selected; resolution enhancement and expansion of the data were done to simplify the identification process (Fig. 4b). The presence of SRGs M1 through M6 in the GXM proton spectrum was determined by using the chemical-shift assignments listed in Table 2. Two SRGs, M2 and M3 were identified by using the data for the GXM of *C. neoformans* isolate 150 (Fig. 4b) and the data in Table 2. Similarly, 106 isolates were investigated and the SRG distributions in individual GXMs were determined (Table 3).

**Quantitation of SRGs: PeakFit.** The original proton NMR data for the SRG region were analyzed by using PeakFit as described in Materials and Methods. The PeakFit analysis using the FELIX ASCII file generated from the primary proton NMR data for the SRG region recorded for the GXM from *C. neoformans* isolate 150 is given in Fig. 5. The area for each identifiable resonance appearing in the proton NMR spectrum (Fig. 4) was assigned to a particular SRG based on its characteristic proton chemical shifts, identified as described above (Table 2). The area data depicted in Fig. 5 were used to calculate the percent occurrence of the SRGs present in the GXM from *C. neoformans* isolate 150 (Table 3). The analysis was repeated for all the GXMs available, and the results were tabulated (Table 3). The PeakFit data were normalized to 100%.

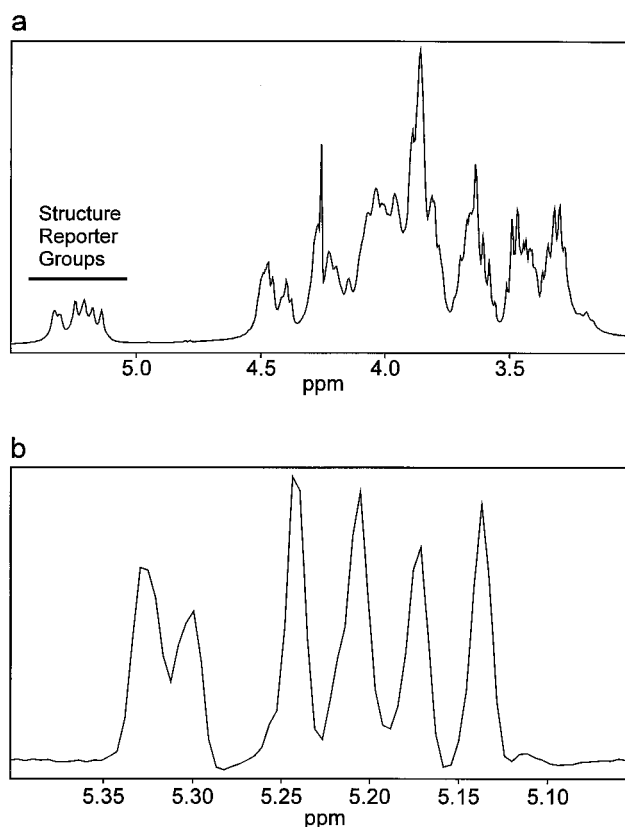


FIG. 4. Full (a) and resolution-enhanced (b) 1D proton NMR spectrum of *C. neoformans* isolate 150.

**Order of preferences for entering SRG data in Table 3.** The PeakFit data were used to determine the order of preference for entering the quantitative distribution of the SRGs in Table 3. The first order of preference was the identification of SRG M1. The data were entered in descending order of area assigned to SRG M1. When SRGs M5 and M6 were also present, they were listed without regard to their quantitative appearance in the GXM. M5 and M6 were not used in determining the order of preference here or in later entries, even in those few cases where one of them was present as the primary component, and this data set was defined as chemotype 1 (Chem1) (Table 3). The data for the ensuing set of entries were selected based on the concurrent appearance of M1 and M2, with the relative area of M1 determining the order of preference (Table 3). This data set was defined as Chem2 (Table 3). The third order of preference was based on the concurrent appearance of M1, M2, and M3, with the relative area of M1 determining the order of entry. This data set was defined as Chem3 (Table 3). The fourth order of preference was based on the sole appearance of M2, and this data set was defined as Chem4 (Table 3). The fifth order of preference was based on the appearance of M2 and M3, with the relative area of M2 determining the order of entry, and this data set was defined as Chem5 (Table 3). The sixth order of preference was based on the appearance of M2, M3, and M4, with the relative area of M2 determining the order of entry. Only two GXMs met these criteria, and this data set was defined as Chem6 (Table 3). The seventh order of preference was based on the appearance of M3, and this data set was defined as Chem7 (Table 3). The eighth, and final, order of preference was based on the appear-

TABLE 3. Assignment of SRGs M1 through M6 to chemotypes 1 through 8 by using PeakFit and the ANN for the isolates described in Table 1<sup>a</sup>

Isolate	M1	M2	M3	M4	M5	M6	SRG	Total
Chem1								
P0	100 [72]				[11]	[16]	M1	[99]
24067 d1	100 [75]				[8]	[22]	M1	[105]
24067 d2	100 [73]				[8]	[20]	M1	[101]
24067 e	100 [71]			[6]	[9]	[17]	M1	[103]
24067 g	100 [75]				[9]	[18]	M1	[104]
24067 f1	100 [62]		[8]	[14]	[10]	[14]	M1	[108]
24067 f2	100 [71]		[23]		[7]		M1	[101]
24067 c	100 [76]				[10]	[18]	M1	[104]
24067 a	100 [71]			[5]	[13]	[16]	M1	[105]
24067 b	100 [74]				[8]	[20]	M1	[102]
D529	79 [72]				[8]	21 [22]	M1	[102]
1254	79 [72]					21 [20]	M1	[92]
9375	74 [71]				10 [10]	16 [21]	M1	[102]
M0050	70 [67]		[18]			30 [32]	M1	[117]
D52	65 [66]					35 [35]	M1	[101]
3168	68 [72]				8 [8]	24 [21]	M1	[101]
127	64 [72]				14 [8]	22 [22]	M1	[102]
J9a	59 [59]					41 [42]	M1	[101]
F12	57 [66]				11 [8]	32 [24]	M1	[98]
NIH 37	56 [35]	[19]			11 [4]	33 [49]	M1	[107]
J9b	53 [46]				27 [30]	20 [14]	M1	[90]
B3502	52 [53]					48 [45]	M1	[98]

*Continued on following page*

TABLE 3—Continued

Isolate	M1	M2	M3	M4	M5	M6	SRG	Total
J9c	50 [46]				29 [26]	21 [15]	M1	[87]
J22b	49 [47]	[8]			22 [4]	29 [48]	M1	[107]
J22a	39 [33]	[18]			18 [3]	43 [53]	M1	[107]
430	38 [49]				44 [36]	18 [12]	M1	[97]
CDC92-159	38 [39]	[10]			28 [8]	34 [30]	M1	[77]
ATCC 28958	34 [34]	[16]			19 [3]	47 [54]	M1	[107]
ATCC 48184	32 [20]	[46]			16 [3]	52 [47]	M1	[116]
CDC92-232	31 [30]					69 [65]	M1	[95]
M0024	30 [32]					70 [64]	M1	[96]
98	27 [32]				20 [2]	53 [54]	M1	[89]
ATCC 34875	27 [25]	[32]			25 [3]	48 [55]	M1	[115]
Chem2								
J9d	63 [66]	18 [1]				19 [22]	M1>M2 <sup>b</sup>	[89]
SB4a	47 [24]	28 [16]				25 [40]	M1>M2	[80]
M0025	46 [5]	29 [40]	[9]			25 [16]	M1>M2	[70]
SB6a	32 [29]	24 [20]			14 [7]	30 [34]	M1>M2	[90]
132	31 [19]	23 [27]			2 [2]	44 [47]	M1>M2	[95]
110	31 [23]	21 [30]			6 [3]	42 [53]	M1>M2	[112]
SB4c	28 [27]	28 [23]			12 [6]	36 [34]	M1>M2	[90]
M0095	19 [10]	59 [54]				22 [24]	M2>M1	[88]
Chem3								
M0051	32 [16]	29 [37]	18 [3]			21 [34]	M1>M2>M3	[90]
M0010	28 [12]	25 [63]	6 [3]		17 [4]	24 [30]	M1>M2>M3	[112]

Continued on following page

TABLE 3—Continued

Isolate	M1	M2	M3	M4	M5	M6	SRG	Total
M0062	25 [18]	17 [39]	8 [2]		20 [3]	30 [38]	M1>M2>M3	[100]
SB6b	23 [18]	34 [33]	15 [3]			28 [42]	M2>M1>M3	[96]
M0046	23 [12]	18 [56]	4 [3]		30 [3]	25 [29]	M1>M2>M3	[103]
M0027	23 [2]	46 [49]	31 [17]				M2>M3>M1	[68]
M0013	19 [2]	38 [80]	15 [15]		6 [1]	22 [7]	M2>M1>M3	[105]
AW	8 [14]	47 [37]	9 [2]		12 [5]	24 [25]	M2>M1>M3	[83]
Chem4 MU-1		100 [89]	[14]				M2	[103]
201	[5]	100 [25]	[85]	[6]	[10]	[5]	M2	[136]
A24064		100 [92]					M2	[92]
118		100 [87]					M2	[87]
SB4b		77 [72]				23 [8]	M2	[80]
9759a		76 [63]				24 [16]	M2	[79]
9759		75 [68]				25 [12]	M2	[80]
M0029		74 [76]				26 [8]	M2	[84]
J11b		63 [76]	[12]		17 [1]	22 [8]	M2	[97]
A104		58 [57]				42 [35]	M2	[92]
Chem5 R1064		91 [83]	9 [16]				M2>M3	[99]
M0053		85 [84]	15 [13]				M2>M3	[97]
ATCC 62066		82 [86]	18 [13]				M2>M3	[99]
F113		81 [86]	19 [17]				M2>M3	[103]
B4538		76 [81]	11 [15]			13 [6]	M2>M3	[102]

Continued on following page



TABLE 3—Continued

Isolate	M1	M2	M3	M4	M5	M6	SRG	Total
TYCC 38-602		75 [76]	25 [22]				M2>M3	[98]
194		75 [75]	25 [19]				M2>M3	[94]
F27		75 [89]	25 [14]				M2>M3	[103]
F117		73 [86]	27 [8]				M2>M3	[94]
ATCC 6352		71 [84]	29 [25]				M2>M3	[109]
196		70 [76]	30 [12]				M2>M3	[88]
NU-2		69 [78]	31 [20]				M2>M3	[98]
A289		69 [69]	31 [26]				M2>M3	[95]
CDC92-64		67 [56]	33 [35]				M2>M3	[91]
ATCC 90126		65 [66]	35 [35]				M2>M3	[101]
6		65 [69]	14 [21]			21 [8]	M2>M3	[98]
J11a		64 [78]	12 [7]			24 [10]	M2>M3	[95]
M0048		63 [72]	37 [31]				M2>M3	[103]
371		60 [62]	40 [37]				M2>M3	[99]
CDC92-162		59 [55]	41 [45]				M2>M3	[100]
A288		58 [32]	20 [8]		11 [3]	11 [8]	M2>M3	[53]
NCPF 3168		52 [86]	48 [14]				M2>M3	[100]
CDC92-187		49 [56]	44 [42]			7 [2]	M2>M3	[100]
150		45 [62]	55 [41]				M3>M2	[103]
A271		43 [39]	57 [50]	[7]			M3>M2	[96]
ATCC 32608		39 [18]	46 [70]	[10]		15 [2]	M3>M2	[100]
Chem6 401		24 [4]	60 [55]	16 [30]			M3>M2>M4	[89]

Continued on following page

TABLE 3—Continued

Isolate	M1	M2	M3	M4	M5	M6	SRG	Total
C917		13 [21]	74 [72]	13 [13]			M3>M2=M4	[106]
Chem7								
C24066			100 [93]	[6]			M3	[99]
Houpt 24066			100 [93]	[5]			M3	[98]
C24066W		[16]	100 [86]	[5]			M3	[107]
Kuhl 94-8069			100 [96]				M3	[96]
3939			100 [93]				M3	[93]
3183			100 [94]				M3	[94]
ATCC 24065			100 [96]				M3	[96]
ATCC 24066			100 [95]				M3	[95]
ATCC 32269			100 [92]				M3	[92]
ATCC 34878			100 [90]				M3	[90]
409			100 [94]				M3	[94]
184			100 [96]				M3	[96]
18			100 [94]				M3	[94]
Chem8								
D24067			79 [79]	21 [17]			M3>M4	[96]
298Warty			72 [87]	28 [14]			M3>M4	[101]
ATCC 34880			65 [69]	35 [35]			M3>M4	[104]
298			48 [49]	52 [55]			M3>M4	[104]
34			33 [32]	67 [60]			M3>M4	[92]
C24066T				100 [60]	[30]		M4	[90]

<sup>a</sup> Values are percentages. For each GXM file, values on the first line are the PeakFit results normalized to 100%; values on the second line are obtained directly from the neural network analysis of the SRG region of the proton NMR spectrum and are enclosed in brackets.

<sup>b</sup> M1>M2, the area of M1 is greater than the area of M2. See Results, "Order of preference for entering SRG data in Table 3."

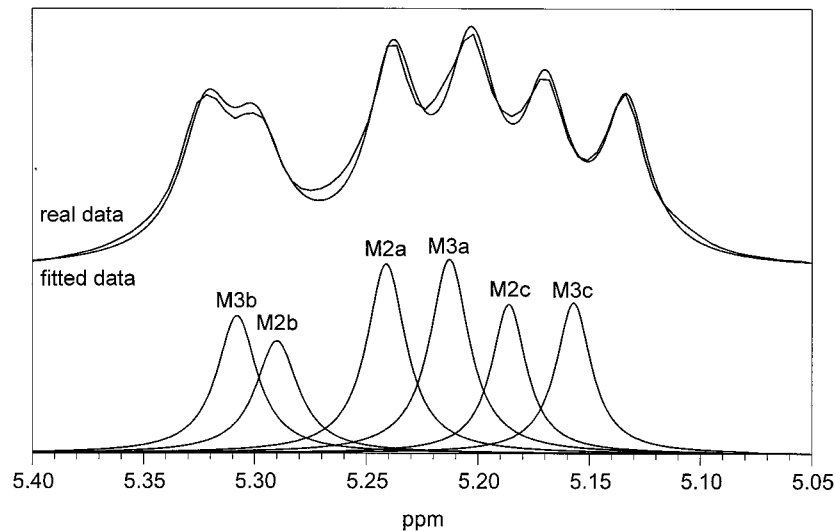


FIG. 5. PeakFit analysis of the SRG region of the 1D proton NMR spectrum of *C. neoformans* isolate 150.

ance of M3 and M4, and this set of data was defined as Chem8 (Table 3). The one archetypal isolate containing M4 only was included in this group.

**Quantitation of SRGs: ANN model.** Several specific ANN models of the feed-forward multistaged backpropagation type were considered (53). The architectures of these models differ in the number of hidden neurons, the complexity of input patterns, and the formulation of the output. The performances of several models were investigated to determine the one most suited for our study. Parameters such as hidden layer size, step size, etc. (53) were optimized experimentally. Fifty different networks were trained in order to determine the optimum parameters. The optimum network consisted of 1,000 input neurons, 5 hidden neurons, and 6 output neurons. A step size of 0.01 was used without a momentum term (53). The 1,000 input neurons covered a small portion of the 1D proton spectrum between 5.0 and 5.5 ppm, representing the mannose anomeric region (SRG). The results of the ANN analysis were tabulated and appear in Table 3 in brackets.

**Chemotype distribution.** The distribution of the eight suggested chemotypes (Chem1 through Chem8) in the 106 isolates of *C. neoformans* investigated in this study is summarized in Table 4.

## DISCUSSION

Latex agglutination tests and enzyme immunoassays for the detection of GXM in cerebral spinal fluid and serum, and

TABLE 4. Suggested chemotypes of *C. neoformans* based on the SRG concept

SRG class	Other SRGs	No.	Chemotype
M1	M6 or M5 and M6	33	Chem1
M1/M2	M6 or M5 and M6	8	Chem2
M1/M2/M3	M6 or M5 and M6 or none	8	Chem3
M2	M6	10	Chem4
M2/M3	None or few M6	26	Chem5
M2/M3/M4	None	2	Chem6
M3	None	13	Chem7
M3/M4 or M4	None	6	Chem8

polyclonal and monoclonal antibodies for serotyping, have been used for diagnostic and epidemiological studies (44). The detection of antigen is diagnostic for infection by *C. neoformans*. The continued presence of high antigen levels indicates a poor prognosis for the patient, whereas antigen clearance is indicative of a successful course of treatment (34). The complete eradication of the organism by current treatment regimens is unlikely, and relapse usually occurs. Comparative investigations of the immunochemistry, immunology, and genetics of *C. neoformans* are intimately linked to GXM. In turn, the results of these investigations are intrinsically related to the chemical structure of GXM at the residue level. The correlation between virulence, the course of the disease, and the chemical structure of GXM has not been studied. The recognition of the importance of GXM structure served as the driving force for this study.

Currently, *C. neoformans* serotypes cannot be further subdivided because the available serological reagents do not have the discriminating power required to recognize structural variation in GXMs (38). This phenomenon is due to the fact that although *O*-acetyl is only a minor contributor to the epitope structure, it is a major determinant of the observed serological activity; de-*O*-acetylated GXMs do not react with serotype-specific factor sera (15, 16, 18). Therefore, the principal variations in structure, which are based on the disposition of the sugar residues in GXM, are not detected. This lack of a strain-typing method has prevented the demonstration of point sources of infection that is required for epidemiological investigations of cryptococcosis (38). Bhattacharjee et al. (18) suggested the occurrence of four discrete model structures, one for each of the serologically defined types shown in Fig. 2 as M1 through M4, and implied that heterogeneity may exist between isolates. Heterogeneity, real or perceived, may be due to the usual pitfalls encountered in the study of complex polysaccharides by classical methods of structural analysis used in earlier studies (such as incomplete acid hydrolysis, differential loss of certain residues over others, low yields by partial acid hydrolysis, and incomplete methylation).

NMR spectroscopy, a nondestructive, highly discriminating analytical method, was used by us to probe the fine structure of de-*O*-acetylated polysaccharides. No secondary chemical reactions, hydrolysis, or derivatization was required. In a series of

in-depth studies, we demonstrated that heterogeneity in GXM structure is the norm. NMR analysis identified sequences originally relegated to one serotype in other types. Structures thought to occur in only one variety were commonly found in the other variety. In addition, two other triads, M5 and M6, were identified, and their structures were characterized by NMR spectroscopy. The six triads, M1 through M6 (Fig. 2), were used as a set of SRGs to develop the chemotyping system. The proton NMR spectra obtained for de-O-acetylated GXMs were resolution enhanced in order to determine precisely the chemical shifts present in a GXM. The chemical-shift data were used to identify which SRGs were present in a particular GXM by comparison to the assignments found in Table 2. Analysis of the SRG region of the original proton NMR spectrum was repeated by using PeakFit. The PeakFit analysis was used to quantify the distribution of SRGs present in a particular GXM. A large selection of GXMs was used to create a database for 106 GXMs (Table 3).

The SRG data formed a pattern of eight clusters based on the preferences used to generate Table 3. Chem1 (33 isolates [Table 4]) was determined by the presence of the SRG M1. M1 in Chem1 was almost always found associated with M6 or with M6 and M5. M5 never occurred alone or without M1 in any GXM. Chem2 (eight isolates [Table 4]) was determined by the concurrent presence of SRGs M1 and M2. Chem2 isolates were always associated with M6 or with M5 and M6. Chem3 (eight isolates [Table 4]) was determined by the concurrent presence of SRGs M1, M2, and M3. Chem3 isolates were associated with M6 in two cases and with M5 and M6 in one case; in one instance, M5 and M6 were absent. Chem4 (10 isolates [Table 4]) was determined solely by the presence of M2. In the Chem4 isolates, M6 was found associated with half the isolates and M5 was never detected. Chem5 (26 isolates [Table 4]) was determined by the presence of M2 and M3. Only four isolates in Chem5 were found associated with M6; M5 was never found. Chem6 (two isolates [Table 4]) was characterized by the concurrent presence of M2, M3, and M4; M5 and M6 were absent. Chem7 (13 isolates [Table 4]) was determined by the presence of M3; M5 and M6 were absent. This chemotype consisted of the largest structurally homogeneous set of isolates. Chem8 (6 isolates [Table 4]) was determined by the presence of M3 and M4; M5 and M6 were absent. One of the isolates was an M4 archetype isolate but was included in this chemotype. The data in Table 3 demonstrated that a continuum of SRGs characterized the GXMs of *C. neoformans*. A blurred demarcation between the structural elements previously assigned specifically to one or the other of the two varieties of *C. neoformans* (varieties *neoformans* and *gattii*) is apparent. As the degree of substitution of the core structure with Xyl increased, a concomitant decrease in the occurrence of SRGs M5 and M6 was observed. The identification of GXM archetype isolates was the exception rather than the rule. The 10 isolates in Chem7 that consisted of SRG M3 only were exceptional. However, M3 was frequently associated with Chem3, Chem5, Chem6, and Chem7.

Both the test set and the learning set of NMR data files, which were analyzed previously by PeakFit, were used to evaluate the trained neural network. The quantitative distribution of the SRGs obtained by the neural network (data in brackets) was compared to the data obtained by PeakFit (Table 3). At present there is not an absolute correspondence between the data obtained by PeakFit and those obtained by the neural network. In fact, a comparatively large percentage of M6 triad was retrieved when we performed a validation test of the ANN with a series of GXMs that essentially comprised only M1 triads (24067 isolates) (Tables 1 and 3). This phenomenon may

be due to the overlap of M6b resonance with the degenerate M2b,c resonances (Table 2). Therefore, we know that training needs to be modified in order to remove any remaining ambiguity in the output data obtained by the present ANN.

The results of the standard backpropagation model for the ANN are not shown in Table 3 because of the large deviation of the sum of the ratios for the SRGs from the target of 100%. The sum of the ratios reported by the modified backpropagation model also diverged from the desired 100% value. But this divergence was much less than that observed for the standard backpropagation model. Although the ratios listed in Table 3 for the PeakFit analyses sum to 100%, this was accomplished by the manual scaling of the initial results. We believe that the small amount of residual chemical noise and resonance overlap present in the spectra affects the results of all mathematical analyses of the data.

An advantage of the ANN's analysis of this type of data is that it can be shown mathematically that the further the sums of the ratios diverge from 100%, the less confidence one should place in the results (30). Therefore, a confidence index can be computed for each output. We did not scale the ANN output data algebraically to give a sum of the ratios equal to 100% because we believe that the observed divergence carries important information that could indicate the presence of contaminants and other possible artifacts.

In the future, we will refine the neural network by increasing the size of the learning set and by using NMR data from several archetypal isolates of *C. neoformans* that produce GXMs with a single mannosyl triad in the SRG region of the spectra. We will also conduct experiments using larger segments of the spectrum as input data for the ANN in an attempt to increase the accuracy of the results. We plan to test the accuracy of the trained ANN by using archetype GXMs to prepare synthetic mixtures that contain predetermined ratios of the various SRGs.

The correlation of chemotype, based on the distribution of SRGs, with serotype was not possible because the *O*-acetyl substituent is instrumental in determining the structure of the epitope reactive with serological reagents. The GXM of 24066Tan is composed of M4 only and has no detectable acetyl substitution. Therefore, we can conclude that M3 acetylation fixed the conformation of M3 to produce the same epitope structure created by *O*-4 substitution in M4.

The chemotyping system and the identification of subtypes within each chemotype, based on the quantitative distribution of the SRGs, could lead to (i) a better understanding of how the polysaccharide antigens influence type 1 and type 2 T-cell responses, (ii) a better understanding of how the polysaccharides potentiate opportunistic infections in HIV-infected patients, (iii) the design of effective haptens for the preparation of neoglycoconjugate vaccines, (iv) a better understanding of how virulence is related to molecular structure, (v) the development of a computer-assisted automated system for the chemotyping of *C. neoformans* isolates and for the monitoring of phenotype switching, (vi) the development of a new epidemiological tool for the surveillance of cryptococcosis, and (vii) a possible explanation of the variation in antigen titer determined by common diagnostic tests.

#### ACKNOWLEDGMENTS

This research was supported in part by Public Health Service grant AI-31769 from the National Institutes of Health. The NMR spectrometers were purchased with the help of the National Science Foundation and the Georgia Research Alliance.

## REFERENCES

- Akay, M. 1992. Noninvasive diagnosis of coronary artery disease using a neural network algorithm. *Biol. Cybern.* **67**:361-367.
- Alpsan, D. 1994. Auditory evoked potential classification by unsupervised ART 2-A and supervised fuzzy ARTMAP networks, p. 3512-3515. *In Proceedings of IEEE International Conference on Neural Networks*, Orlando, Fla. IEEE Neural Network Council, New York, N.Y.
- Andrea, T. A., and H. Kalayeh. 1991. Applications of neural networks: quantitative structure-activity relationships of dihydrofolate reductase inhibitors. *J. Med. Chem.* **34**:2824-2836.
- Andreasen, H., H. Bohr, J. Bohr, S. Brunak, T. Bugge, R. M. J. Cotterill, C. Jacobsen, P. Kusk, and B. Lautrap. 1990. Analysis of secondary structure of the human immunodeficiency virus proteins by computer modelling based on neural network methods. *J. Acquired Immune Defic. Syndr.* **3**:615-622.
- Apolloni, B., G. Avanzini, N. Cesa-Bianchi, and G. Ronchini. 1990. Diagnosis of epilepsy via backpropagation, p. 571-574. *In Proceedings of the 1990 International Joint Conference on Neural Networks*, vol. 2.
- Armentrout, S. L., J. A. Reggia, and M. Weinrich. 1994. A neural model of cortical map reorganization following a focal lesion. *Artif. Intellig. Med.* **6**:383-400.
- Asada, N., K. Doi, H. MacMahon, S. M. Montner, M. L. Giger, C. Abe, and Y. Z. Wu. 1990. Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung disease: pilot study. *Radiology* **177**:857-860.
- Astion, M. L., and P. Wilding. 1992. The application of backpropagation neural networks to problems in pathology and laboratory medicine. *Arch. Pathol. Lab. Med.* **116**:995-1001.
- Astion, M. L., and P. Wilding. 1992. Application of neural networks to the interpretation of laboratory data in cancer diagnosis. *Clin. Chem.* **38**:34-38.
- Avanzolini, G., P. Barbini, and G. Gnudi. 1990. Unsupervised learning and discriminant analysis applied to identification of high risk postoperative cardiac patients. *Int. J. Bio-Med. Comput.* **25**:207-221.
- Bacon, B. E., and R. Cherniak. 1995. Structure of the *O*-deacetylated glucuronoxylomannan from *Cryptococcus neoformans* serotype C as determined by 2D <sup>1</sup>H NMR spectroscopy. *Carbohydr. Res.* **276**:365-386.
- Bacon, B. E., R. Cherniak, K. J. Kwon-Chung, and E. S. Jacobson. 1996. Structure of the *O*-deacetylated glucuronoxylomannan from *Cryptococcus neoformans* Cap70 as determined by 2D <sup>1</sup>H NMR spectroscopy. *Carbohydr. Res.* **283**:95-110.
- Barski, L. L., R. S. Gaborski, and P. G. Anderson. 1993. A neural network approach to the histogram segmentation of digital radiographic images, p. 375-380. *In C. H. Dagli, L. I. Burke, B. R. Fernandez, and J. Ghosh (ed.), Intelligent engineering systems through artificial neural networks. Proceedings of the Artificial Neural Networks in Engineering Conference*, vol. 3. ASME Press, New York, N.Y.
- Baxt, W. G. 1990. Use of an artificial neural network for data analysis in clinical decision-making: the diagnosis of acute coronary occlusion. *Neural Comput.* **2**:480-489.
- Belay, T., and R. Cherniak. 1995. Determination of antigen binding specificities of *Cryptococcus neoformans* factor sera by enzyme-linked immunosorbent assay. *Infect. Immun.* **63**:1810-1819.
- Belay, T., R. Cherniak, E. B. O'Neill, and T. R. Kozel. 1996. Serotyping of *Cryptococcus neoformans* by dot enzyme assay. *J. Clin. Microbiol.* **34**:466-470.
- Belay, T., R. Cherniak, T. R. Kozel, and A. Casadevall. 1997. Reactivity patterns and epitope specificities of anti-*Cryptococcus neoformans* monoclonal antibodies by enzyme-linked immunosorbent assay and by dot enzyme assay. *Infect. Immun.* **65**:718-728.
- Bhattacharjee, A. K., J. E. Bennett, and C. P. J. Glaudemans. 1984. Capsular polysaccharides of *Cryptococcus neoformans*. *Rev. Infect. Dis.* **6**:619-624.
- Bottone, E. J., I. F. Salkin, N. J. Hurd, and G. P. Wormser. 1987. Serogroup distribution of *Cryptococcus neoformans* in patients with AIDS. *J. Infect. Dis.* **156**:242.
- Bulmer, G. S., and M. D. Sans. 1968. *Cryptococcus neoformans*. III. Inhibition of phagocytosis. *J. Bacteriol.* **95**:5-8.
- Chaka, W., A. F. M. Verheul, V. V. Vaishnav, R. Cherniak, J. Scharringa, J. Verhoef, H. Snippe, and I. M. Hoepelman. 1997. *Cryptococcus neoformans* and cryptococcal glucuronoxylomannan, galactoxylomannan, and mannoprotein induce different levels of tumor necrosis factor alpha in human peripheral blood mononuclear cells. *Infect. Immun.* **65**:272-278.
- Cherniak, R., L. C. Morris, B. C. Anderson, and S. A. Meyer. 1991. Facilitated isolation, purification, and analysis of glucuronoxylomannan of *Cryptococcus neoformans*. *Infect. Immun.* **59**:59-64.
- Cherniak, R., L. C. Morris, S. A. Meyer, and T. B. Mitchell. 1992. Glucuronoxylomannan of *Cryptococcus neoformans* obtained from patients with AIDS. *Carbohydr. Res.* **249**:405-413.
- Cherniak, R., L. C. Morris, and S. A. Meyer. 1992. Glucuronoxylomannan of *Cryptococcus neoformans* serotype C: structural analysis by gas-liquid chromatography-mass spectrometry and by <sup>13</sup>C nuclear magnetic resonance spectroscopy. *Carbohydr. Res.* **225**:331-337.
- Coker, R. J. 1992. Cryptococcal infection in AIDS. *Int. J. Sex. Transm. Dis. AIDS* **3**:168-172.
- Diamond, R. D. 1995. *Cryptococcus neoformans*, p. 2331-2340. *In G. L. Mandell, J. E. Bennett, and R. G. Douglas, Jr. (ed.), Principles and practice of infectious diseases*. Churchill Livingstone, New York, N.Y.
- Dismukes, W. E. 1988. Cryptococcal meningitis in patients with AIDS. *J. Infect. Dis.* **157**:624-628.
- Dong, Z. M., and J. W. Murphy. 1995. Intravascular cryptococcal culture filtrate (CneF) and its major component, glucuronoxylomannan, are potent inhibitors of leukocyte accumulation. *Infect. Immun.* **63**:770-778.
- Echaz, J., and G. Vachtsevanos. 1994. Neural network detection of anti-epileptic drugs from a single EEG trace, p. 346-351. *In Proceedings of the IEEE Electro/94 International Conference*.
- Ersoy, O. K., and D. Hong. 1990. Parallel, self-organizing, hierarchical neural networks. *IEEE Trans. Neural Netw.* **1**:167-178.
- Evans, E. E. 1950. The antigenic composition of *Cryptococcus neoformans*. I. A serologic classification by means of the capsular and agglutination reactions. *J. Immunol.* **64**:423-430.
- Fromtling, R. A., H. J. Shadomy, and E. S. Jacobson. 1982. Decreased virulence in stable, acapsular mutants of *Cryptococcus neoformans*. *Mycopathologia* **79**:23-29.
- Ikeda, R., T. Shinoda, Y. Fukazawa, and L. Kaufman. 1982. Antigenic characterization of *Cryptococcus neoformans* serotypes and its application to serotyping of clinical isolates. *J. Clin. Microbiol.* **16**:22-29.
- Kaufman, L., and E. Reiss. 1992. Serodiagnosis of fungal diseases, p. 507-528. *In N. R. Rose, E. C. de Macario, J. L. Fahey, H. Friedman, and G. M. Penn (ed.), Manual of clinical laboratory immunology*, 4th ed. American Society for Microbiology, Washington, D.C.
- Kovacs, J. A., A. A. Kovacs, M. Polis, W. C. Wright, V. J. Gill, C. U. Tuazon, E. P. Gelmann, H. C. Lane, R. Longfield, G. Overturf, A. M. Macher, A. S. Fauci, J. E. Parrillo, J. E. Bennett, and H. Masur. 1985. Cryptococcosis in the acquired immunodeficiency syndrome. *Ann. Intern. Med.* **103**:533-538.
- Kozel, T. R., and E. C. Gotschlich. 1982. The capsule of *Cryptococcus neoformans* passively inhibits phagocytosis of the yeast macrophages. *J. Immunol.* **129**:1675-1680.
- Kozel, T. R., W. F. Gulley, and J. Cazin, Jr. 1977. Immune response to *Cryptococcus neoformans* soluble polysaccharide: immunological unresponsiveness. *Infect. Immun.* **18**:701-707.
- Kwon-Chung, K.-J., and J. E. Bennett. 1992. Medical mycology, p. 397-446. Lea and Febiger, Malvern, Pa.
- Kwon-Chung, K. J., I. Polacheck, and J. E. Bennett. 1982. Improved diagnostic medium for separation of *Cryptococcus neoformans* var. *neoformans* (serotypes A and D) and *Cryptococcus neoformans* var. *gattii* (serotypes B and C). *J. Clin. Microbiol.* **15**:535-537.
- Kwon-Chung, K. J., A. Varma, and D. H. Howard. 1990. Ecology of *Cryptococcus neoformans* and prevalence of its two varieties in AIDS and non-AIDS associated cryptococcosis, p. 103-113. *In H. V. Bossche and D. W. R. Mackenzie (ed.), Mycoses in AIDS patients*. Plenum Press, New York, N.Y.
- Levitz, S. M. 1991. The ecology of *Cryptococcus neoformans* and the epidemiology of cryptococcosis. *Rev. Infect. Dis.* **13**:1163-1169.
- Levitz, S. M., A. Tabuni, H. Kornfeld, C. C. Reardon, and D. T. Golenbock. 1994. Production of tumor necrosis factor alpha in human leukocytes stimulated by *Cryptococcus neoformans*. *Infect. Immun.* **62**:1975-1981.
- Meyer, B., T. Hansen, D. Nute, P. Albersheim, A. Darvill, W. York, and J. Sellers. 1991. Identification of the <sup>1</sup>H-NMR spectra of complex oligosaccharides with artificial neural networks. *Science* **251**:542-544.
- Mitchell, T. G., and J. R. Perfect. 1995. Cryptococcosis in the era of AIDS—100 years after the discovery of *Cryptococcus neoformans*. *Clin. Microbiol. Rev.* **8**:515-548.
- Murphy, J. W. 1989. Immunoregulation in cryptococcosis, p. 319-344. *In E. Kurstak (ed.), Immunology of fungal infections*. Marcel Dekker, Inc., New York, N.Y.
- Orendi, J. M., H. S. L. M. Nottet, M. Visser, A. F. M. Verheul, H. Snippe, and J. Verhoef. 1994. Enhancement of HIV-1 replication in peripheral blood mononuclear cells by *Cryptococcus neoformans* is monocyte-dependent but tumor necrosis factor-independent. *AIDS* **8**:423-429.
- Pettoello-Mantovani, M., A. Casadevall, T. R. Kollmann, A. Rubinstein, and H. Goldstein. 1992. Enhancement of HIV-1 infection by the capsular polysaccharide of *Cryptococcus neoformans*. *Lancet* **339**:21-23.
- Pinner, R. W., R. A. Hajjeh, and W. G. Powderly. 1995. Prospects for preventing cryptococcosis in persons infected with human immunodeficiency virus. *Clin. Infect. Dis.* **21**:S103-S107.
- Powell, D. A., V. Turula, J. A. de Haseth, H. van Halbeek, and B. Meyer. 1994. Sulfate detection in glycoprotein-derived oligosaccharides by artificial neural network analysis of Fourier-transform infrared spectra. *Anal. Biochem.* **220**:20-27.
- Radomski, J. P., H. van Halbeek, and B. Meyer. 1994. Neural network-based recognition of oligosaccharide <sup>1</sup>H-NMR spectra. *Nat. Struct. Biol.* **1**:217-218.
- Rinaldi, M. G., D. J. Drutz, A. Howell, M. A. Sande, C. B. Wofsy, and W. K. Hadley. 1986. Serotypes of *Cryptococcus neoformans* in patients with AIDS. *J. Infect. Dis.* **153**:642.
- Rippon, J. W. 1988. Medical mycology: the pathogenic fungi and the patho-



- genic actinomycetes, 3rd ed., p. 582–609. W. B. Saunders Co., Philadelphia, Pa.
53. **Rumelhart, D. E., and J. L. McClelland.** 1987. Parallel distributed processing: explorations in the microstructure of cognition, vol. 1 and 2. MIT Press, Cambridge, Mass.
  54. **Sellers, J., W. York, P. Albersheim, A. Darvill, and B. Meyer.** 1990. Identification of the mass spectra of partially methylated alditol acetates by artificial neural networks. *Carbohydr. Res.* **207**:C1–C5.
  55. **Shadmehr, R., and D. Z. D'Argenio.** 1990. A neural network for nonlinear Bayesian estimation in drug therapy. *Neural Comput.* **2**:216–225.
  56. **Sheng, S., and R. Cherniak.** 1997. Structure of the  $^{13}\text{C}$ -enriched *O*-deacetylated glucuronoxylomannan of *Cryptococcus neoformans* serotype A determined by NMR spectroscopy. *Carbohydr. Res.* **301**:33–40.
  57. **Shimizu, R. Y., D. H. Howard, and M. N. Clancy.** 1986. The variety of *Cryptococcus neoformans* in patients with AIDS. *J. Infect. Dis.* **154**:1042.
  58. **Skelton, M. A., H. van Halbeek, and R. Cherniak.** 1991. Complete assignment of the  $^1\text{H}$  and  $^{13}\text{C}$ -NMR spectra of the de-*O*-acetylated glucuronoxylomannan from *Cryptococcus neoformans* serotype B. *Carbohydr. Res.* **221**:259–268.
  59. **Skelton, M. A., R. Cherniak, L. Poppe, and H. van Halbeek.** 1991. Structure of the de-*O*-acetylated glucuronoxylomannan from *Cryptococcus neoformans* serotype D, as determined by 2D NMR spectroscopy. *Magn. Reson. Chem.* **29**:786–793.
  60. **Sundstrom, J. B., and R. Cherniak.** 1994. Polysaccharide antigens of the capsule of *Cryptococcus neoformans*. *Infect. Immun.* **62**:1507–1512.
  61. **Thomsen, J., and B. Meyer.** 1989. Pattern recognition of  $^1\text{H}$  NMR spectra of sugar alditols using a neural network. *J. Magn. Reson.* **84**:212–217.
  62. **Turner, S. H., R. Cherniak, E. Reiss, and K. J. Kwon-Chung.** 1992. Structural variability in the glucuronoxylomannan of *Cryptococcus neoformans* serotype A isolates determined by  $^{13}\text{C}$  NMR spectroscopy. *Carbohydr. Res.* **233**:205–218.
  63. **Turner, S. H., and R. Cherniak.** 1991. Multiplicity in the structure of the glucuronoxylomannan of *Cryptococcus neoformans*. NATO ASI (Adv. Sci. Inst.) Ser. Ser. H **53**:123–142.
  64. **Valafar, F., H. Valafar, O. K. Ersoy, and R. J. Schwartz.** 1995. Comparative studies of two neural network architectures for modeling of human speech production, p. 2056–2059. *In* Proceedings of the International Conference on Neural Networks, vol. 4.
  65. **van Halbeek, H.** 1984. Structural analysis of the carbohydrate chains of mucin-type glycoproteins by high-resolution  $^1\text{H}$ -NMR spectroscopy. *Biochem. Soc. Trans.* **12**:601–605.
  66. **Vliegthart, J. F. G., L. Dorland, and H. van Halbeek.** 1983. High-resolution  $^1\text{H}$ -nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *Adv. Carbohydr. Chem. Biochem.* **41**:209–374.
  67. **Welisk, R. M., E. T. Starcher, and J. W. Curran.** 1987. Opportunistic disease reported in AIDS patients: frequencies, associations and trends. *AIDS* **1**:175–182.
  68. **Wilson, D. E., J. E. Bennett, and J. W. Bailey.** 1968. Serologic grouping of *Cryptococcus neoformans*. *Proc. Soc. Exp. Biol. Med.* **127**:820–823.