Pergamon

Bioorganic & Medicinal Chemistry, Vol. 4, No. 11, pp 1979–1988, 1996

PII: S0968-0896(96)00179-4
0968-0896/96 $15.00+0.00

How do Antibodies and Lectins Recognize Histo-Blood Group Antigens? A 3D-QSAR Study by Comparative Molecular Field Analysis (CoMFA)

Anne Imberty,[a]* Rosella Mollicone,[b] Emmanuel Mikros,[c] Pierre-Alain Carrupt,[d] Serge Pérez[e] and Rafael Oriol[b]

[a]*Laboratoire de Synthèse Organique, Centre National de la Recherche Scientifique, 2 rue de la Houssinière, F-44072 Nantes cedex 03, France*
[b]*INSERM U178, Hôpital Paul-Brousse, 94807 Villejuif, France*
[c]*Department of Pharmacy, University of Athens, Panespistimiopoli, Zografou, GR-15771 Athens, Greece*
[d]*Institut de Chimie Thérapeutique, BEP, Université de Lausanne, CH-1015 Lausanne, Switzerland*
[e]*Ingeniérie Moléculaire, Institut National de la Recherche Agronomique, BP 1627, F-44316 Nantes cedex 03, France*

Abstract—The cross-reaction patterns of nine antibodies and three lectins against 12 H type 2 related oligosaccharides have been analysed by means of 3D-QSAR study. Three-dimensional descriptors of the molecular properties have been used in comparative molecular field analysis (CoMFA). Three different alignments were considered for the oligosaccharides. One, based on the superimposition of the oligosaccharide core, could be correlated to most of the antibody activities. A second alignment, based on a superimposition of the fucose residue, had to be taken into account for explaining the binding properties of *Ulex europaeus* isolectin I. Analysis of the QSAR data gives indications on the carbohydrate epitopes essential for antibody recognition and yields some insights about the nature of the molecular recognition. This study complements previous biochemical estimates of the H type 2 related oligosaccharide binding areas (Mollicone, R.; Cailleau, A.; Imberty, A.; Gane, P.; Pérez, S.; Oriol, R. *Glycoconj. J.* **1996**, *13*, 263–271). Copyright © 1996 Elsevier Science Ltd

Introduction

Monoclonal antibodies (MAb) and lectins reacting with blood group carbohydrate antigens have been the subject of many studies from the team of R. U. Lemieux.[1–5] The serological cross-reaction patterns of ABO antibodies[6] and Lewis antibodies[7] have been thoroughly studied. Recently, the results obtained with anti-H antibodies submitted to two International Workshops on Monoclonal Antibodies Against Human Red Blood Cell and Related Antigens[8,9] (Paris 1987 and Lund 1990) allowed one to describe the main cross-reaction patterns of H type 2 antibodies.[10] Similar results have been reported, by another group, on 10 different anti-H type 2 MAbs.[11]

From these data and from the computer simulation of the lowest energy conformation of histo-blood group oligosaccharides,[12] it was possible to separate the antibodies in at least seven families, each one recognizing a different area of the carbohydrate.[10] Such interpretation, based on a qualitative and empirical evaluation, was straightforward for the antibodies displaying narrow specificity. For the antibodies with a larger specificity and with a panel of different binding intensities for different epitopes, a quantitative statistical treatment of the information should be undertaken in order to extract all the information contained in the immunology studies.

The quantitative structure–activity relationships (QSAR)[13,14] approach is the appropriate tool for such study. For exploring a recognition process between the ligand and the antibodies, 3D-QSAR methods are now available, such as the comparative molecular field analysis (CoMFA)[15] which uses three-dimensional descriptors of the molecules. Even though these tools are widely used in theoretical chemistry, pharmacology and drug design, they have never been used in the glycobiology field, the unique exception being a study of the sweetness of halogenated sucrose analogues.[16,17] In this case the CoMFA method, with the inclusion of a molecular lipophilicity potential has been able not only to rationalize the data but also to have a predictive power.

The aim of the present work is to compare the results obtained with different antibodies in order to obtain further insight on the binding area of the carbohydrate antigens and on the nature of the recognition, i.e. the relative importance of steric, electrostatic, and lipophilic contributions.

3D-QSAR Strategy

QSAR (quantitative structure–activity relationships) methods[13,14] have been widely used in the last 30 years, producing multilinear equations between several

physico-chemical parameters of molecules and their biological activities. More recently, methods such as the comparative field analysis or the molecular similarity analysis allowed one to take into account the three-dimensional aspects of host–ligand recognition.

The CoMFA method[15] allows the comparison of molecules through their three-dimensional aspects. The molecular properties taken into account are classically the steric and electrostatic fields. These fields are defined by all the values of one potential (either steric or electrostatic) at the positions corresponding to the nodes of a grid (or lattice) around the molecule. Since these two molecular fields cannot take into account all the complex intermolecular forces between ligands and receptors, the use of an additional field, based on molecular lipophilicity, usually enriches the modeling of the interaction.[17,18] The CoMFA method generates a linear relation between the fields and the biological activity of the molecule. The fields are then compared. In such a method, the alignment of the molecules in the model is of a prime importance.

The PLS (partial least square) analysis[19,20] allows one to take into account many molecular descriptors and is therefore well suited for the three-dimensional descriptors issued from CoMFA analysis. The many variables are then represented by a smaller number of components, or vectors. In order to evaluate the correlation between the model and the activity, several coefficients have to be checked. The squared correlation coefficient, r^2, corresponds to the fraction of the variance explained by the QSAR study. For evaluating the predictive value of a model, it is more appropriate to do a cross-validation study in which each compound is taken out of the pool. Its activity is then compared to the calculated one. For the model to have a reasonable statistical significance, q^2 (the cross-validated r^2) should be high (in practice more than 0.4) and s (the predicted error) should be low.

Finally, if a model is considered to be valid, the CoMFA method allows one to display graphically some aspects of the analysis. For example, the regions of each field which have the lowest and highest deviation in the study can be represented using a contour map procedure. Such an approach is used in general for designing a drug with higher affinity for the probe. In the present case, the contour maps are used to understand the differences in the recognition phenomenon occurring for different families of lectins and antibodies.

For the present study, a subset of the data presented in a previous paper[10] has been selected. The binding activity was evaluated as the maximum specific absorbance of the complete dilution curve for each reagent. Because of the statistical nature of the QSAR method, all the binding data could not be used. Only antibodies displaying the largest panel of binding activities, with a modulation of intensity as a function of the oligosaccharide structures, can be used. Therefore, many antibodies have been discarded. For the same reason, the oligosaccharides which are never recognized by the antibodies of interest, such as the A or B trisaccharides, could not be included. Table 1 summarizes the binding activities which have been selected from Mollicone et al.[10] for undertaking a 3D-QSAR study.

Results

Molecular modeling

The conformational behavior of all oligosaccharides presented here, except one, have been the subject of a recent publication.[12] Therefore, only the trisaccharide αFuc1-2βGal1-3βGal (H type 5 trisaccharide, Table 2) was investigated and added to the present study. Because of the absence of the *N*-acetyl group, this trisaccharide is more flexible that the related compound αFuc1-2βGal1-3βGalNAc (H type 4), and

Table 1. ELISA binding of 10 antibodies and three lectins to 12 immobilized oligosaccharide epitopes (adapted from Mollicone et al.[10]). Clusters have been shadowed, the cut-off criteria to define shaded areas being set to 0.5 OD. Six different patterns of recognition labeled from I to VI have been defined, based on both the shape of the clusters, and the CoMFA analysis data

Reagent	Group	H type 2	H type 6	H type 5	H type 1	H type 4	H type 3	Le^x	Le^y	Le^a	Le^b	Le^c	Le dis
MAbs													
061	I	1.5	1.4	0.4	0.2	0.0	0.1	0.9	1.3	0.0	0.0	0.0	0.0
057	I	1.4	1.4	0.0	0.0	0.0	0.0	1.4	1.5	0.5	0.0	0.0	0.0
066	I	0.7	1.5	0.0	0.0	0.0	0.0	1.4	1.5	0.0	0.0	0.0	0.0
1BE12	I	0.3	0.9	0.0	0.0	0.0	0.0	0.8	1.3	0.0	0.0	0.0	0.0
19-OLE	I	0.9	0.7	0.0	0.0	0.0	0.0	0.6	0.8	0.0	0.0	0.0	0.0
1BD6	I	0.6	0.6	0.0	0.0	0.0	0.0	0.8	1.2	0.0	0.0	0.0	0.0
058	II	1.2	1.2	1.1	0.2	0.0	0.1	0.5	1.2	0.0	0.0	0.0	0.0
BNH9	II	1.2	1.1	1.1	0.0	0.0	0.0	0.4	0.9	0.0	0.0	0.0	0.0
8B9	II	1.2	1.4	0.9	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
64/4D8	III	0.7	0.6	0.0	0.7	0.0	0.0	0.5	0.9	0.0	1.0	0.0	0.0
lectins													
Galactia	IV	1.5	1.5	1.4	0.1	1.2	0.6	0.0	0.0	0.0	0.0	0.0	0.0
Ulex	V	1.5	1.5	0.7	0.1	0.0	0.0	0.3	1.4	0.0	0.0	0.0	0.9
Lotus	VI	1.2	1.4	0.0	0.0	0.0	0.0	0.5	0.6	0.0	0.0	0.0	1.5

Table 2. Structures of the H and Lewis related oligosaccharides used for coating the ELISA plates[a] and for the QSAR studies[b]

Name	Chemical structure
H type 2	αFuc1-2βGal1-4βGlcNAc-R
H type 6	αFuc1-2βGal1-4βGlc-R
H type 5	αFuc1-2βGal1-3βGal-R
H type 1	αFuc1-2βGal1-3βGlcNAc-R
H type 4	αFuc1-2βGal1-3βGalNAc-R
H type 3	αFuc1-2βGal1-3αGalNAc-R
Le^x	βGal1-4(αFuc1-3)βGlcNAc-R
Le^y	αFuc1-2βGal1-4(αFuc1-3)βGlcNAc-R
Le^a	βGal1-3(αFuc1-4)βGlcNAc-R
Le^b	αFuc1-2βGal1-3(αFuc1-4)βGlcNAc-R
Le^c (Type 1 precursor)	βGal1-3βGlcNAc-R
Le disaccharide	αFuc1-4βGlcNAc-R

[a] R = (CH_2)8-CO-NH-BSA (Chembiomed, Alberta Research Council, Edmonton, Canada).
[b] R = H.

three main families of conformers can be identified. The most populated (42%) one is centered about $\Phi_{FUC12GAL} = -89°$, $\Psi_{FUC12GAL} = -99°$, $\Phi_{GAL13GAL} = -77°$, $\Psi_{GAL13GAL} = 136°$, and corresponds to FamI, the main conformational family of histo-blood group oligosaccharides.[12] The conformational family corresponding to the previously described secondary minimum, FamII, represents for this trisaccharide a population of 24% and is centered about $\Phi_{FUC12GAL} = -94°$, $\Psi_{FUC12GAL} = -175°$, $\Phi_{GAL13GAL} = -87°$, $\Psi_{GAL13GAL} = 67°$. A third family is predicted to occur for this trisaccharide centered around about $\Phi_{FUC12GAL} = -75°$, $\Psi_{FUC12GAl} = -88°$, $\Phi_{GAL13GAL} = -87°$, $\Psi_{GAL13GAL} = 67°$ with a population of 31%. This latter conformation cannot exist in the related histo-blood group oligosaccharides because of a steric hindrance between the central galactose residue and the *N*-acetyl group.

Comparison of molecular properties

Molecular properties have been exemplified on two molecules, the H type 1 and H type 2 trisaccharides (Fig. 1). As pointed out in previous conformational studies,[10,21,22] the overall shape of the lowest energy conformation of these two disaccharides is very similar. This is due to the similarity in shape of the global minima of the βGal1-4βGlcNac and βGal1-3βGlcNAc disaccharides. Only the *N*-acetyl group of the GlcNAc residue points toward a different direction. If the shapes of these two molecules, and therefore their accessible surfaces calculated with the Connolly method[23] are quite similar, the other molecular properties are not. As illustrated in Figure 1, the two molecules differ in a rather important way, when looking at their molecular properties. The extreme of the electrostatic potential being correlated to the CO—NH dipole, the orientation of the *N*-acetyl group plays an important role on the electrostatic potential. Similarly, the different distribution of methyl groups in the two molecules affects the lipophilicity potential.

This analysis of the molecular properties of two trisaccharides illustrates the variations of type that molecular properties undergo in the panel of molecules of the present study, and therefore the type of variations that will be taken into account in the 3D-QSAR study.

Alignment of conformers

Three different alignments of the 12 oligosaccharides considered here have been tested in the QSAR studies. In the first alignment, all the global minima of the main conformational family (FamI) have been aligned, taking the central disaccharide βGal1-3/4βGlcNAc/GalNAc (or Gal) as a reference. This alignment (Align1) is displayed in Figure 2. The fit of the central rings is excellent and a common shape appears clearly for all the oligosaccharides studied here. A second alignment could be produced for all the oligosaccharides which exhibit a secondary conformational family. The Lewis tri- and tetrasaccharides (Le^x, Le^y, Le^a, and Le^b) which are essentially rigid are not taken into account in this alignment. The result is displayed as Align2 in Figure 2. Because of peculiar results obtained with the fucose-specific lectins, a third alignment has to be considered, using the fucose residues as the reference. In this case, the lowest energy conformers (i.e. FamI) were considered, and the fucose ring of the αFuc1-2βGal linkage was used for the fit. For the oligosaccharide where this residue is missing, i.e. Le^x, Le^a and Lewis disaccharide, the other fucose ring (αFuc1-3/4βGlcNAC) was used. The result, Align3 (Fig. 2) does not correspond to the best superimposition of the overall oligosaccharide shape but allows a fucose residue always to be presented in the same orientation.

QSAR analysis

The ELISA binding values of all the antibodies and lectins listed in Table 1 were tested against the three alignments. Only models with statistical significances are listed in Table 3. The binding activities of 10 antibodies and the lectin from *Galactia tenuiflora* can be satisfactorily correlated to the first model, Align1. In contrast, the second model, Align2, which was not based on the lowest energy conformations of the oligosaccharides but on the secondary minima, could not be correlated to any binding activity. The binding of the fucose-specific lectin from *Ulex europaeus* I could not be correlated to any of these two alignments. However, it yields to a good model with the third one, Align3, which is based on the best superimposition of the fucose rings. The binding of the fucose-specific lectin from *Lotus tetragonolobus* also shows a better correlation with the Align3 model, even though the results are at the limit of the statistical significance.

The relatively small number of compounds studied limits the number of principal components allowed in CoMFA analysis. Based on the regular rules of QSAR (not always respected in published CoMFA work!), a maximum of two independent variables, namely two principal components, can be used with 12 compounds. Moreover, first CoMFA analyses for the antibody 061

display similarly good models for analyses using one field only (independently of the nature of field), a combination of two fields, or using the three fields together (steric, electrostatic, or lipophilic). As already noted,[17,18] this behavior may be due to a number of compounds too small to fully represent all structural variations. In such cases, the interpretation of CoMFA signals in terms of intermolecular interactions should be carefully checked in order to limit misleading interpretations with the systematic examination of CoMFA

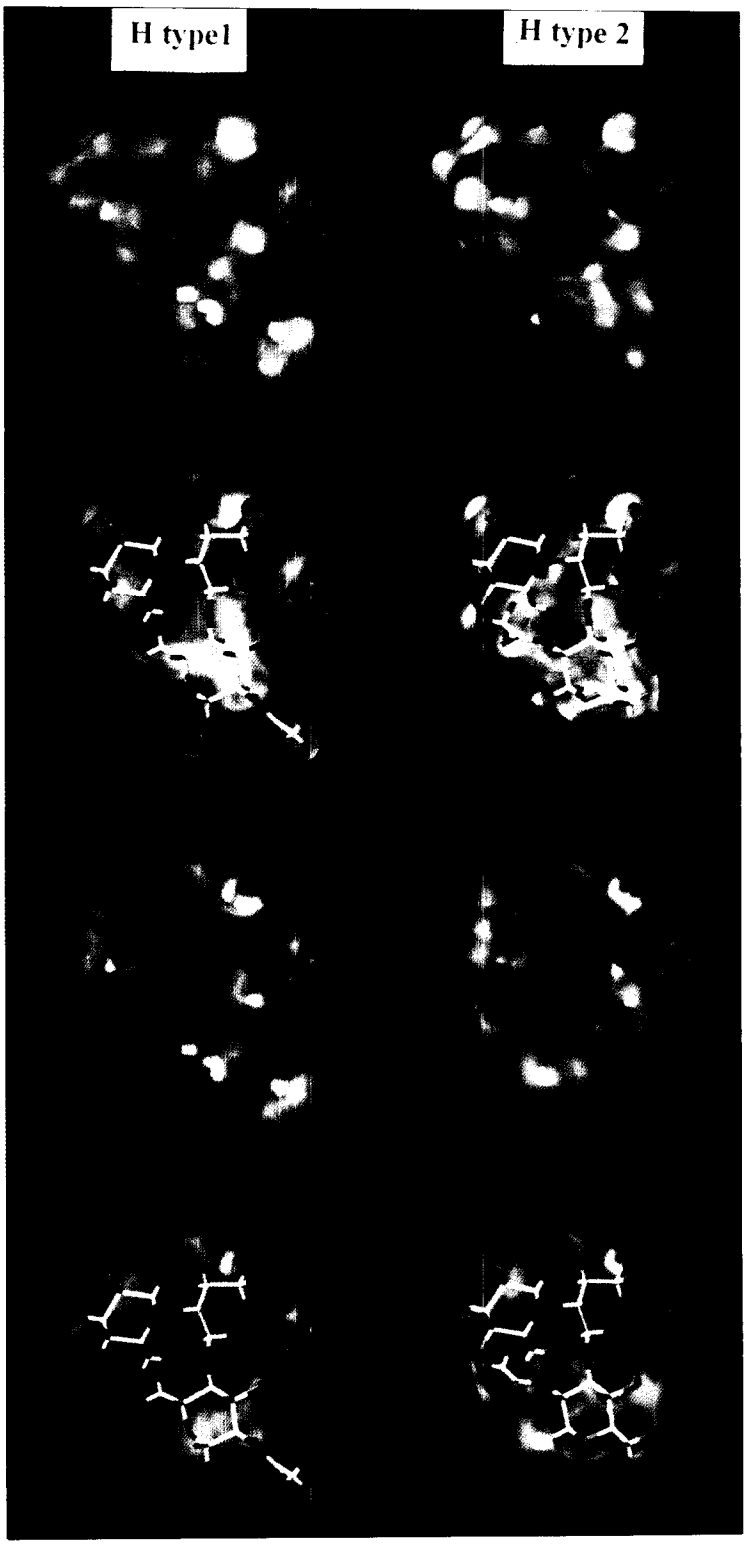

Figure 1. Molecular properties of H type 1 and H type 2 trisaccharide. Top two rows: electrostatic potential displayed on the Connolly surface. Bottom two rows: lipophilicity potential. In each case, the front face and the back face of the molecules are displayed.

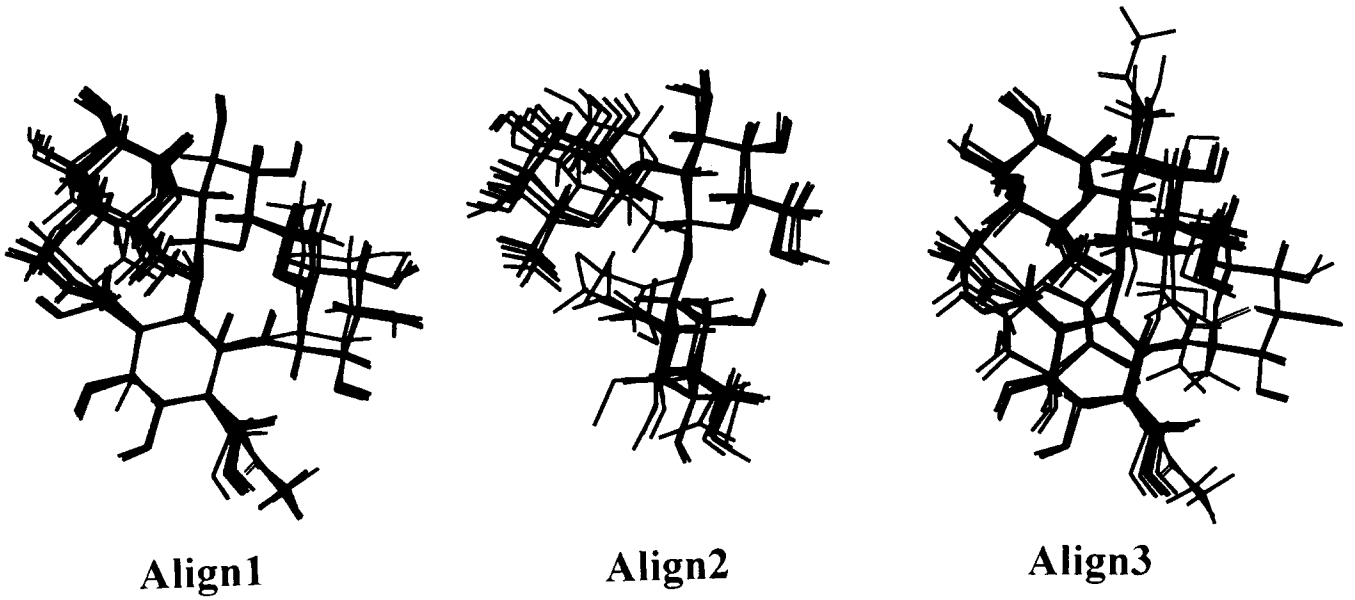

Figure 2. The three alignments: Align1, Align2, and Align3 used in the QSAR study.

models with separated fields and with all possible combinations of fields.

For all models, the mixing of the two fields (as illustrated with the case of the CoMFA analyses with steric and electrostatic fields) or the three fields did not improve the statistics. Furthermore, in many models, the quality of the statistics was similar in models using each field alone. These observations suggest that the structural variations modulating the recognition are described well only by the presence or the absence of a field in some regions of space. A limited structural variation and a restricted series of compounds are generally responsible for this yes/no-type answer from CoMFA models. Although this type of answer could also be obtained with simpler QSAR techniques like the Free–Wilson analysis,[24] 3D-QSAR techniques allow the discrimination between several alignments and the proposal of three-dimensional topological relations between pharmacophoric elements.

For the antibodies of group I, the best q^2 values are obtained with electrostatic or lipophilic fields, the steric fields giving, in general, significant models with lower quality. For the antibodies of group II, the q^2 are lower but the effects of each field appears similar to those of group I. In one case (antibody BNH9), the use of a lipophilic field improves the statistics. For the antibodies of group III, only the electrostatic or lipophilic field gives a model, whereas for the antibodies of group IV, the electrostatic field is suitable to derive a model. For group V, each field gave a significant model. However, in any case the mixing of different fields led to an enhancement of the model quality.

These results confirm that the variation of the recognition is controlled by the presence or the absence to a structural motif in given regions of space, entirely encoded by the variation of each field.

Graphical Analysis of CoMFA

The most useful way to analyse the CoMFA results is the visualization of the variation of each field associated with the variation of the biological activity. Here, the contouring of statistical fields (standard deviations multiplied by the coefficients in the CoMFA equation) were displayed for each analysis with a single field, and superimposed on each family (Fig. 3). It should be noted that the QSAR study is only valid for the portions of the oligosaccharides which display some variations in the table. It is therefore possible to draw some conclusions about the presence or absence of an *N*-acetyl group, but not on the presence or absence of the center galactose residue, since this sugar has never been altered in the present study.

The differences among the groups that have been described in Table 3, are here displayed in a different way. For example, the antibodies of group I and group II have quite a similar behavior. For all of them, one important feature is the red region appearing at the left of the figure, indicating a negative influence of a bulky substituent in this place. Group I presents in addition a favorable influence of a bulky substituent on the bottom right. This is clearly correlated with the nature 1–3 or 1–4 of the core disaccharide (type 1 precursor or type 2 precursor) which orients the *N*-acetyl group to one side of the trisaccharide or to the other. For the same reason, the electrostatic field gives a favorable influence of the CO—NH dipole on the right. For both I and II groups, another region of interest occurs around the fucose residue linked to the position 2 of galactose (also called the H fucose). This fucose appears to have a positive influence for the

binding mainly by its lipophilicity. Favorable steric interactions also play a role, which could only be visualized by varying the green isocontour level for graphic output (not shown here).

The antibody of group III has a drastically different behavior. The 3D-QSAR study concentrates in a region around the H fucose which has a strongly positive influence, both lipophilic and electrostatic, on binding (the influence of steric field is not represented here because its q^2 value is not statistically significant). The lectin from *G. tenuiflora* again exhibits a different pattern. The H fucose also has a positive influence, but smaller than in group III (only the influence of electrostastic field has been represented).

Binding activities of lectins of group V are correlated with an alignment based on a fucose residue. Since all the oligosaccharides considered here but one (the Le^c disaccharide) have a fucose in the same position in this alignment, this position presents very limited variations and it is not surprising that it does not appear in the field variations that are isocontoured here. As in group I, the nature of the 1–3 or 1–4 central disaccharide has a strong steric and lipophilic influence.

All these different observations can be summarized as follows

- Group I and II: importance of the 1–4 nature of the core disaccharide, through steric, electrostatic,

Table 3. Partial least square analysis (PLS) of the CoMFA study. q^2 is the proportion of the original variance which is explained by using the QSAR during the cross-validation test, N being the number of components giving the optimum value of q^2. s is the standard error on the prediction. r^2 is the final proportion of the original variance explained by the study

Receptor (and group)		1 Field				2 Fields Ste + Ele		3 Fields Ste + Ele + Lipo	
		N	q^2	s	r^2	N	q^2	N	q^2
	Ste	2	0.70	0.21	0.87				
061	Ele	2	0.87	0.12	0.97	2	0.87	2	0.87
(I)	Lipo	2	0.87	0.12	0.97				
	Ste	2	0.58	0.26	0.89				
057	Ele	1	0.85	0.15	0.96	2	0.80	2	0.78
(I)	Lipo	2	0.77	0.16	0.96				
	Ste	2	0.45	0.30	0.85				
066	Ele	1	0.74	0.24	0.88	2	0.70	2	0.65
(I)	Lipo	2	0.60	0.23	0.90				
	Ste	2	0.24	—	—				
1BE12	Ele	1	0.59	0.21	0.81	1	0.53	2	0.42
(I)	Lipo	2	0.41	0.20	0.84				
	Ste	2	0.78	0.11	0.93				
19-OLE	Ele	2	0.88	0.09	0.95	2	0.87	2	0.89
(I)	Lipo	2	0.88	0.05	0.98				
	Ste	2	0.60	0.16	0.89				
1BD6	Ele	1	0.83	0.11	0.93	2	0.79	2	0.74
(I)	Lipo	2	0.65	0.13	0.92				
	Ste	2	0.50	0.23	0.86				
058	Ele	1	0.58	0.26	0.79	1	0.56	2	0.57
(II)	Lipo	2	0.57	0.29	0.75				
	Ste	2	0.47	0.22	0.86				
BNH9	Ele	1	0.44	0.30	0.71	2	0.44	2	0.48
(II)	Lipo	2	0.50	0.19	0.89				
	Ste	2	0.39	—	—				
8B9	Ele	2	0.36	—	—	2	0.35	2	0.40
(II)	Lipo	2	0.51	0.21	0.88				
	Ste	2	0.38	—	—				
64/4D8	Ele	2	0.60	0.13	0.92	2	0.60	2	0.60
(III)	Lipo	2	0.60	0.14	0.90				
	Ste	1	0.28	—	—				
Galactia	Ele	2	0.58	0.20	0.87	2	0.43	1	0.36
(IV)	Lipo	1	0.35	—	—				
	Ste	2	0.70	0.13	0.96				
Ulex[a]	Ele	2	0.73	0.20	0.92	2	0.78	2	0.72
(V)	Lipo	2	0.45	0.11	0.97				
	Ste	2	0.27	—	—				
Lotus[a]	Ele	2	0.22	—	—	2	0.28	2	0.27
(VI)	Lipo	2	0.08	—	—				

[a] CoMFA study performed using the Align3 model.

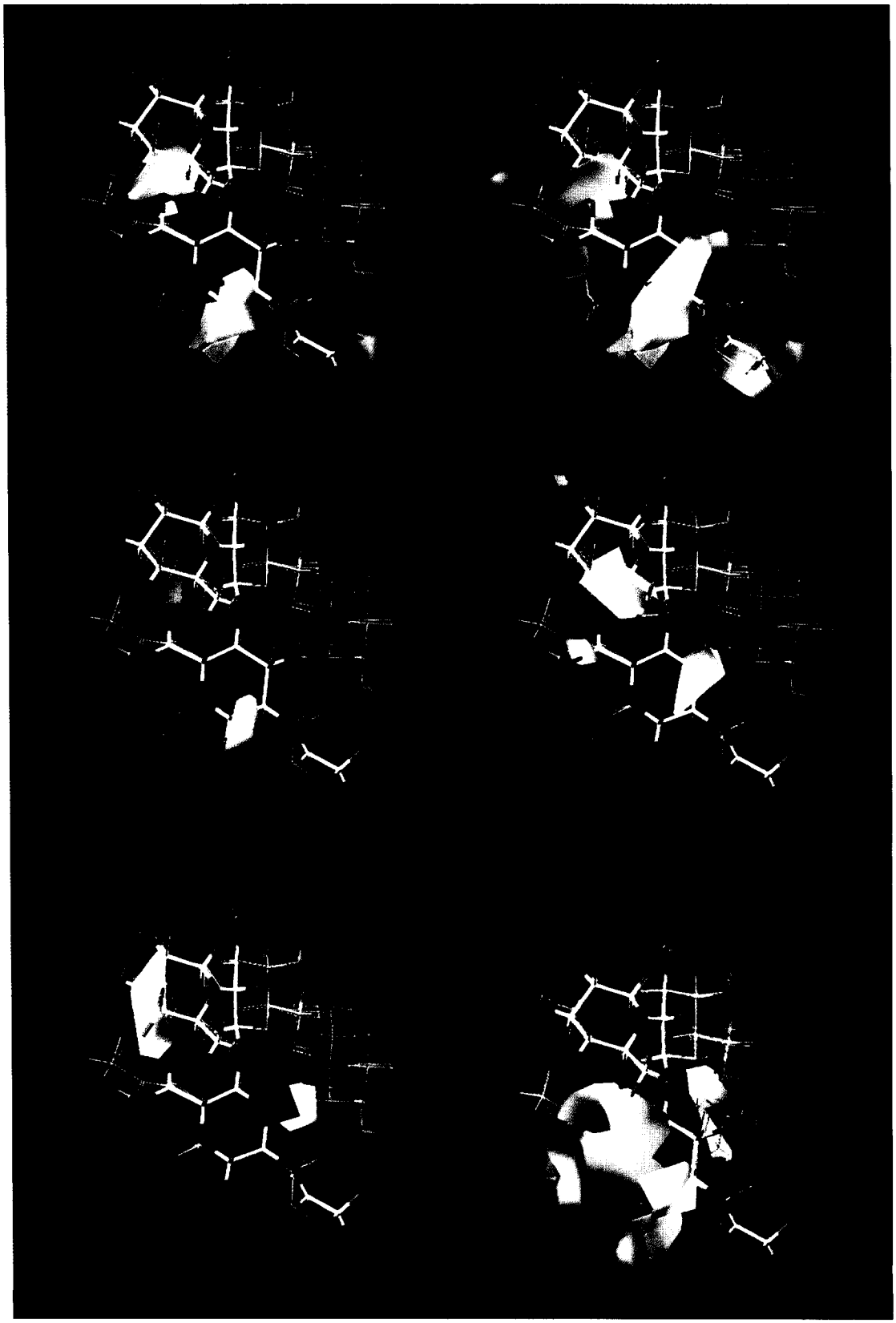

Figure 3. Graphic results of the CoMFA study: standard deviation of each field in the one-field CoMFA analysis. Isocontour of 0.012 and −0.012 are displayed for each field. For the steric field: green, positive deviation (regions where a bulky group would enhance the binding); red, negative deviation (regions where a less bulky group would enhance the binding). For the electrostatic field: white, positive deviation (regions where a positive charge will enhance the binding); magenta, negative deviation (regions where a negative charge will enhance the binding). For the lipophilicity field: yellow, positive deviation (regions where a hydrophobic group will enhance the binding); cyan, negative deviation (regions where a hydrophilic group will enhance the binding).

and lipophilic potential. The existence of positive or negative regions near the same positions similar for each field confirms the yes/no character of these CoMFA analyses.

- Group III and IV: dominated by the influence of the H fucose.
- Group V: necessity of modifying the alignment and importance of the 1–4 nature of the core disaccharide.

Discussion

The results presented here can be compared to those previously obtained using a method consisting of 'mapping' the epitope by defining the part necessary for binding.

The clustering in families proposed in the present work is slightly different from the one resulting from the previous study of H blood group epitope recognition.[10] This is because all the oligosaccharides tested biologically were not taken into account in the present study. However, the main families are conserved and the binding topologies proposed previously can be compared with the present simulations. In the previous studies, we were willing to show the binding area on each oligosaccharide derived from the H type 2 trisaccharide. In each case, the H fucose and the galactose were part of the binding region, and variations could occur on the *N*-acetyl group recognition, depending upon its orientation and upon the antibody. We arrived at the same conclusions here but with more details concerning the positive or negative influence of the substituents and about the type of forces involved in the recognition.

Many data have been collected by the group of Lemieux for the recognition of blood group determinants by fucose-specific lectins, either the lectin I of *U. europaeus*[25] or the lectin from *G. tenuiflora*.[26] Recently, a comparison of the binding site of these two lectins, has been performed by means of chemical mapping.[27,28] When looking at the region which is mainly screened in the present study, i.e. the GlcNAc/GalNAc orientation and substituents, they suggest that the OH in position 6 of GlcNAc of the H type 2 trisaccharide is important for the recognition by *Ulex*, but not for the recognition by *Galactia*. This is in agreement with our results (Fig. 3). The region around the hydroxymethyl of H type 2 GlcNAc has no important field isocontour for *Galactia* (group IV), whereas it displays an important negative steric influence and a requirement of hydrophilicity for *Ulex* (group V).

The ideas developed by Lemieux[29] about protein–carbohydrate interactions highlighted the role of three main factors: the key polar interactions, the peripheral polar interactions, and the lipophilic interactions. The regions of the oligosaccharides involved in these three types of interaction were generally determined by means of chemical mapping. Our approach also allows the carbohydrate regions which play a role in the interaction to be displayed, but with different descriptors (only the lipophilic descriptor is common to both approaches). Since it is difficult to compare directly the hydrogen bonding ability with the electrostatic or steric field, the present approach should be considered as a complementary method, yielding a different view of the interactions.

Conclusion

This work confirms that the blood group carbohydrate antigens present several 'micro-epitopes' which means that they will not be recognized from the same side by different antibodies. These epitopes are conformational-dependent and in all the cases studied here, only the lowest energy conformation is recognized. We can now postulate that different kinds of forces are involved and, for example, that the same fucose residue could have a steric role in the recognition by one antibody and a more lipophilic role in the recognition by another antibody. These points are important, for example, when trying to design glycomimics for 'blocking' anti-blood group antibodies.

Whereas the fucose-specific lectins have a high affinity for H type 2 containing oligosaccharides, they have another scheme of recognition than the antibodies. This is specially true for *U. europaeus* isolectin I and for *L. tetragonolobus* lectin.

Because of the low sensitivity of the experimental method used to produce the assay data, no quantitative predictions can be made in the present study. Nevertheless, the 3D-QSAR results presented here complete the chemical (or biochemical) mapping previously published[10] since they provide a different approach for visualizing the binding area of carbohydrate antigens. This method is in its infancy in the carbohydrate field but it opens a new avenue in glycobiology and glycotherapy. When applied to carbohydrate-based drugs with biological activities evaluated using more sensitive methods, this approach provides a new tool for the design of glycomimetics.

Experimental

Material

Monoclonal antibodies (MAbs) 057, 058, 061, 066 were described in Chester et al.;[30] 1BE12 and 1BD6 in Pancino et al.;[31,32] 19-OLE named 34W6 in Oriol et al.;[8] BNH9 in Delsol et al.;[33] 64/4D8 in Brodin et al.;[34] and 8B9 in Mollicone et al.[10]

Lectins: *G. tenuiflora* was affinity purified on H type 2-Synsorb (Chembiomed, Alberta Research Council, Edmonton, Canada) and labeled with biotin.[10] Affinity-purified, biotin-labeled, *U. europaeus* I and *L. tetrago-*

nolobus lectins were obtained from Vector Laboratories (Burlingame, California, U.S.A.).

Experiments

ELISA tests were performed with synthetic oligosaccharide-BSA compounds (Chembiomed, Alberta Research Council, Edmonton, Canada) as described in Mollicone et al.[10] Briefly, microtiter plates with 96 flat bottom wells were coated overnight at room temperature with a solution at 1 mg mL^{-1} of the synthetic oligosaccharide hapten-BSA in pH 8 PBS. The antigen solution was removed and the plates were incubated 1 h at 37 °C with gelatin 3% in PBS. After washing, twofold serial dilutions of monoclonal antibodies containing gelatin (0.5%) and Tween 20 (0.02%) were added and incubated overnight. The plates were washed and all the MAbs selected for this study were reacted with biotin-labeled secondary anti-mouse Ig antibodies for 2 h, washed again, and peroxidase-labeled streptavidin diluted 1 to 1000 was added for 3 h. After a final wash the coloured reaction was developed with *o*-phenylene diamine (OPD) at 1 mg mL^{-1} in citrate buffer 0.02 M, pH 5, containing 3 mL mL^{-1} of H_2O_2 at 30%. After 3 min the reaction was stopped with 4 N HCl and the optical density at 54 nm was read in a microplate reader (MR 600, Dynatech).

Direct tests with biotin-labeled lectins were performed in the same way without the primary and secondary antibodies. Negative controls were performed with BSA-coated plates and results for BSA-coated wells were substracted from the results of the hapten-BSA-coated wells for each lectin and antibody.

Molecular modeling

The CICADA method[35] was used for a complete exploration of the potential hypersurface, together with the MM3 force-field.[36] A dielectric constant value of 80 was used to mimic the effect of the water environment. The resulting conformations were clustered in families and the lowest energy conformation of each conformational family was determined. Detailed information about the whole procedure and its application to oligosaccharides has been published previously.[12,37]

Molecular properties calculations

Molecular properties have been displayed using the MOLCAD[38] module in the SYBYL software.[39] Its purpose is the creation of molecular surfaces and the display of molecular properties on top of them. The Connolly surface[23] has been chosen here. Calculations of several molecular properties: electrostatic potential, lipophilicity potential, and hydrogen bonding sites capacities can be calculated for a given conformation. The electrostatic potential is calculated using the Coulombic equation. The lipophilicity potential used here is the default one of the SYBYL package: it is based on a modified Fermi type of function, with Crippen atomic partial values.[40]

Alignment of low energy conformers

All the alignments of conformers were performed with the FIT procedure of the SYBYL software. Only the ring atoms and glycosidic oxygen atoms were taken into account. Unless otherwise specified the central disaccharide was used and the inversion of the ring orientation of GlcNAc (or GalNAc) was considered, depending upon the nature (1–3 or 1–4) of the linkage.

QSAR

The CoMFA procedure of the SYBYL software was used for the QSAR study. The grid size was automatically adjusted by the program to accommodate all the molecules of a given alignment. Intervals of 1.5 Å were chosen for the lattice. The steric field and electrostatic field of the SYBYL software were used, which are based on a Lennard–Jones function and a Coulombic potential, respectively. The molecular lipophilicity potential was calculated by a recently developed method.[41] This new approach is based on the atomic lipophilicity values of Broto and Moreau[42] used with a distance equation derived from the Fermi one.

The analysis was conducted with the partial least square method. First a cross-validation was performed. Two numbers of components were tested (1 and 2). The value yielding to the best correlation was kept for the final calculations. After the final analysis a graphic display of the field was done. One low and one high standard deviation of each field were visualized by displaying the isocontour values corresponding to −0.012 and 0.012 variation values.

Miscellaneous

All the calculations were performed on Silicon Graphics workstations. Drawings were done with the SYBYL software.

Acknowledgment

The authors thank Dr Patrick Gaillard for his help in the interpretation of CoMFA results.

References

1. Lemieux, R. U. In *IUPAC Frontiers of Chemistry*; Laidler, K. J. Ed.; Pergamon: New York, 1982; pp 3–23.

2. Lemieux, R. U.; Venot, A. P.; Spohr, U.; Bird, P.; Mandal, G.; Morishima, N.; Hindsgaul, O.; Bundle, D. R. *Can. J. Chem.* **1985**, *63*, 2664.

3. Hindsgaul, O.; Khare, D. P.; Bach, M.; Lemieux, R. U. *Can. J. Chem.* **1985**, *63*, 2653

4. Spohr, U.; Morishima, N.; Hindsgaul, O.; Lemieux, R. U. *Can. J. Chem.* **1985**, *63*, 2659.

5. Lemieux, R. U.; Hindsgaul, O.; Bird, P.; Narasimhan, S.; Young, W. W. *Carbohydr. Res.* **1988**, *178*, 293.

6. Oriol, R.; Samuelsson, B.; Messeter, L. *J. Immunogenet.* **1990**, *17*, 279.

7. Good, A. H.; Yau, O.; Lamontagne, L. R.; Oriol, R. *Vox Sang.* **1992**, *62*, 180.

8. Oriol, R.; Gane, P.; Rouger, P.; Mollicone, R. *Rev. Franç. Transf. Immunohématol.* **1987**, *30*, 671–679.

9. Mollicone, R.; Dalix, A. M.; Gane, P.; Nemec, M.; Oriol, R. In *Second International Workshop on Monoclonal Antibodies against Human Red Blood Cells and Related Antigens*; Chester, M. A.; Johnson, U.; Lundblad, A.; Löw, N.; Messeter, L.; Samuelssson, B. Eds.; Lund: Sweden, 1990; p 85.

10. Mollicone, R.; Cailleau, A;, Imberty, A.; Gane, P.; Pérez, S.; Oriol, R. *Glycoconj. J.* **1996**, *13*, 263–271.

11. Furukawa, K., Welt, S.; Yin, B. W. T.; Feickert, H. J.; Takahashi, T.; Ueda, R.; Lloyd, K. O. *Mol. Immunol.* **1990**, *27*, 723.

12. Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconj. J.* **1995**, *12*, 331.

13. Hansch, C. *Acc. Chem. Res.* **1969**, *2*, 232.

14. Martin, Y. C. In *Quantitative Drug Design*; Marcel Dekker: New York, 1978.

15. Cramer, III R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.

16. Meyer, C. Ph.D. Thesis, Université de Nantes, France, 1994.

17. Carrupt, P.-A.; Gaillard, P.; Billois, F.; Weber, P.; Testa, B.; Meyer, C.; Pérez, S. In *Methods and Principles in Medicinal Chemistry*; Pliska, V.; Testa, B.; van de Waterbeened, H. Eds., VCH: Weinheim; pp 195–217.

18. Gaillard, P.; Carrupt, P. A.; Testa, B.; Schambel, P. *J. Med. Chem.* **1996**, *39*, 126.

19. Wold, S.; Ruhe, A.; Wold, H.; Dunn, III W.J. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735.

20. Dunn, III W. J.; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. *Quant. Struct.-Act. Relat.* **1984**, *3*, 131.

21. Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. *Can. J. Chem.* **1980**, *58*, 631.

22. Biswas, M.; Rao, V. S. R. *Biopolymers* **1980**, *19*, 1555.

23. Connolly, M. L. *Science* **1983**, *221*, 709.

24. Kubinyi, H. In *QSAR: Hansch Analysis and Related Approaches*; VCH: Weinheim, 1993.

25. Hindsgaul, O.; Norberg, T.; Le Pendu, J.; Lemieux, R. U. *Carbohydr. Res.* **1982**, *109*, 109.

26. Cromer, R.; Spohr, U.; Khare, D. P.; Le Pendu, J.; Lemieux, R. U. *Can. J. Chem.* **1992**, *70*, 1511.

27. Du, M.-G.; Spohr, U.; Lemieux, R. U. *Glycoconj. J.* **1994**, *11*, 443.

28. Lemieux, R. U. In *Complex Carbohydrates in Drug Research*, Bock, K.; Clausen, H. Eds.; Munksgaard; Copenhagen, 1994; pp 188–197.

29. Lemieux, R. U. *Chem. Soc. Rev.* **1989**, *18*, 347.

30. Chester, M. A.; Johnson, U.; Lundblad, A.; Löw, B.; Messeter, L.; Samuelsson, B. In *Proceedings of the Second International Workshop on Monoclonal Antibodies against Human Red Blood Cells and Related Antigens*; Lund: Sweden; p 209.

31. Pancino, G.; Charpin, C.; Osinaga, E.; Betaille, B.; Le Roy, M.; Calvo, F.; Roseto, A. *Cancer Res.* **1990**, *50*, 7333.

32. Pancino, G.; Toubert, M. E.; Osinaga, E.; Chatelet, F.; Leroy, M.; Schlageter, M. H.; Desroys du Roure, F.; Calvo, F.; Teillac, P.; Najean, Y. *Int. J. Cancer* **1991**, *47*, 221.

33. Delsol, G.; Blancher, A.; Al Saati, T.; Ralfkiaer, E.; Lauritzen, A.; Bruigeres, L.; Brousset, P.; Rigal-Huguet, F.; Mazerolles, C.; Robert, A.; Chittal, S. M. *Br. J. Cancer.* **1991**, *64*, 321.

34. Brodin, T.; Chester, M. A.; Karlsson, K. A.; Messeter, L.; Zopf, D.; Lundblad, A. *Glycoconj. J.* **1987**, *4*, 399.

35. Koca, J. *J. Mol. Struct. (Theochem)* **1994**, *308*, 13.

36. Allinger, N. L.; Yuh, Y. H.; Li, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

37. Koca, J.; Pérez, S.; Imberty, A. *J. Comp. Chem.* **1995**, *16*, 296.

38. Waldherr-Teschner, M.; Goetze, T.; Heiden, W.; Knoblauch, M.; Vollhardt, H.; Brickmann, J. In *Advances in Scientific Visualization*; Post, F. H.; Hin, A. J. S. Eds.; Springer: Heidelberg, 1992; pp 58–67.

39. SYBYL V6.2, Tripos Associates, 1699 S. Hanley Road, Suite 303, St Louis MO 63144, U.S.A.

40. Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163.

41. Gaillard, P.; Carrupt, P. A.; Testa, B.; Boudon, A. *J. Comp.-Aided Mol. Design* **1994**, *8*, 83.

42. Broto, P.; Moreau, G.; Vadycke, C. *Eur. J. Med. Chem.* **1984**, *19*, 61.

(*Received in U.S.A.* 15 *February* 1996; *accepted* 17 *May* 1996)