A New Method for Predicting the Alignment of Flexible Molecules and Orienting Them in a Receptor Cleft of Known Structure

Maria Cristina De Rosa* and Anders Berglund†

Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, U.K.

Received September 2, 1997

It is not always easy to align flexible compounds with each other or with their binding cleft on a biological macromolecule, and the alignment of nine partly flexible molecules has now been studied. These compounds are heme analogues having either two or three flexible propionate side chains attached to a porphyrin core, and one compound is a close analogue of natural heme. The noncovalent interactions of each compound were predicted using a new version of the program Grid which can take account of the flexibility of the propionate side chains. The Grid results were then analyzed by hierarchical principal component analysis, and this allowed the molecules to be oriented with respect to each other. It also allowed each analogue to be correctly aligned with the receptor cleft for heme in myoglobin, because the alignment of natural heme in that cleft is already known. Factors influencing the predicted alignment are also considered.

Introduction

The heme-binding pocket of myoglobin (Mb) interacts very favorably with its natural prosthetic group (protoporphyrin IX, 8,13-divinyl-3,7,12,17-tetramethyl-21*H*,-23*H*-porphine-2,18-dipropionic acid), and these interactions fall into two main classes: (i) the binding between His F8 imidazole and the iron of heme; (ii) the hydrophobic attraction between the nonpolar atoms of the site and the π system and the alkyl groups of the heme. Three edges of the porphyrin ring, containing the apolar substituents, are plunged into the protein and therefore are not well-exposed to the solvent. On the contrary, the fourth edge, holding the two propionate side chains, is directed toward the protein surface where the carboxy oxygens can interact with water.

The multiple interactions which determine the protoporphyrin IX binding to apomyoglobin have been described in detail.[1,2] They seem to suggest a specific molecular recognition, but several experimental results suggest that heme binding to globins may not be so critically influenced by the detailed protein structure.

In the naturally occurring single-site hemoglobin mutants Hb Milwaukee and Hb Bristol, Val E11, a buried hydrophobic residue situated within van der Waals contact of the heme, is replaced in the β-chains by Glu and Asp, respectively,[3,4] but this very drastic alteration does not displace the heme. Varadarajan et al.[5] replaced Val E11 with Glu, Asp, and Arg in human myoglobin, thus demonstrating again that globins are able to tolerate such drastic changes without displacement of ligand. The role of the distal histidine residue has been studied by replacing His E7 in sperm whale myoglobin with a number of different amino acid residues and measuring ligand binding parameters.[6−8] Each of the distal histidine substitutions resulted in a decreased affinity for dioxygen,[7] but the heme remained in its cleft. In the case of hemoglobin, the heme iron of Hb M Boston (Hisα E7 → Tyr) and M. Saskatoon (His βE7 → Tyr) is stabilized in the ferric state by substitution of the distal histidine with tyrosine in the heme pockets of the α- and β-subunits.[9,10] According to Hargove et al.[11] who measured the rate constants for CO-heme binding to 35 recombinant apomyoglobins with substitutions in the heme pocket, the heme was never displaced.

The reaction of native hemoglobin with several protoheme derivatives has been reported.[12] Gibson and Antonini[13] showed that meso-, deutero-, and hematohemes react at pH 9.1 with apohemoglobin. Rose and Olson[14] reported similar results at pH 7.2. Chu and Bucci[15] showed that dimethyl ester CO-heme reacts with apohemoglobin at the same rate as CO-protoheme. Single-crystal structure determination showed, in erythrocruorin, a reversed heme orientation,[16] and heme rotational disorder has been detected by crystallography in one other case.[17] In solution, however, heme disorder may be more readily characterized by ^{1}H NMR spectroscopy. La Mar et al.[18−20] showed that the initial product of the reaction of apomyoglobin with hemin was a mixture of two interconverting protein forms which differ in the orientation of the heme by a 180° rotation about the α,γ-meso axis. The dominant component (about 90%) has the same orientation as observed by X-ray crystallography although the oxygen affinity differed significantly for the two heme orientations.[21] Larger degrees of equilibrium disorder were found in insect hemoglobin,[22] fish myoglobin,[23] and mammalian myoglobin reconstituted with hemins chemically modified at the 2- and 4-positions.[18] NMR studies revealed that the reversed orientation was predominant in solution for the monomeric hemoglobins of *Glycera dibranchiata*[24] and *Chironomus thummi thummi*.[25] Heme insertion isomers of *C. thummi thummi* hemoglobin

* Address for correspondence: Centro Chimica dei Recettori e delle Molecole Biologicamente Attive del CNR, Università Cattolica del S. Cuore, Largo F. Vito 1, 00168 Roma, Italy. Phone: + 39 6 3057612. Fax: + 39 6 3053598. E-mail: cristina@medea.ccr.rm.cnr.it.

† Present address: Research Group for Chemometrics, Department of Organic Chemistry, Umeå University, S901 87 Sweden.

S0022-2623(97)00582-7 CCC: $15.00
Published on Web 01/31/1998

have been observed.[26] Finally, Hauksson et al.[27] showed by ^{1}H NMR experiments that each of a series of synthetic hemins with different substitution patterns of the propionate side chains can bind to sperm whale apomyoglobin.

These experimental results all demonstrate that the high affinity of hemes for globins is not due to exceptionally specific interactions between the macromolecule and the ligand but that it can be maintained when the structure of either partner is changed. This lack of specificity may be a general feature of molecular recognition whenever several related ligands bind to a receptor cleft.

For a better understanding of the molecular interactions involved in the binding of heme and its analogues to Mb, a molecular modeling study using the program Grid[28] has been carried out. The Grid program calculates intermolecular interaction energies between a molecule and a series of probes each of which represents a specific chemical group. The probes are characterized by their steric, electrostatic, and hydrogen-bonding properties and by their hybridization, and they have been selected for the present work in order to mimic chemical groups which are present in the receptor cleft of myoglobin. There is a "flexibility" option in version 15 of Grid. When this option is turned on, the flexible side chains of the target molecules can move in response to he probe. They move toward the probe when there is attraction and away from it when there is repulsion, thus simulating the response of the target to a change in its environment.

The results from Grid have been used as input data for a hierarchical principal component analysis (Hi-PCA)[29,30] in order to align the individual heme analogues with each other, and then, since one heme is a close analogue of the naturally occurring protoporphyrin IX, the whole set of hemes was aligned with the myoglobin. The results were checked by comparison with the observations of Hauksson et al. and suggest that a Grid−Hi-PCA approach could be applied to other systems involving the binding of partly flexible molecules to a receptor.

Methods

Procedure for Building the Targets. The nine different modified hemes studied by Hauksson et al.[27] were considered as ligands of sperm whale apomyoglobin. To model these compounds the coordinates of protoporphyrin IX in myoglobin were obtained from the Protein Data Bank.[31] All the unique orientations obtained by a 90° rotation about the normal axis to the heme plane (axis A in Figure 1) and a 180° rotation about the α,γ-meso axis (axis B in Figure 1) were then considered. The flexible side chains are shown in Figure 1 in the conformations which were observed when heme was bound to Mb in the X-ray structure.

To study exactly the same structures as described by Hauksson et al., the vinyl groups of native heme were replaced by methyl groups, and Sybyl (version 6.2)[32] was then used for structure manipulation of the modified protoporphyrin IX. The nomenclature of Hauksson et al. was used for the different molecules and binding modes, to facilitate the comparison of our theoretical analysis with their observations.

The nine heme analogues studied by Hauksson et al. are shown in Figure 2 with their propionate side chains drawn by the Sybyl builder module. However, version 15 of Grid allows the side chains to move, and the Grid calculations are therefore independent of these initial propionate conformations.

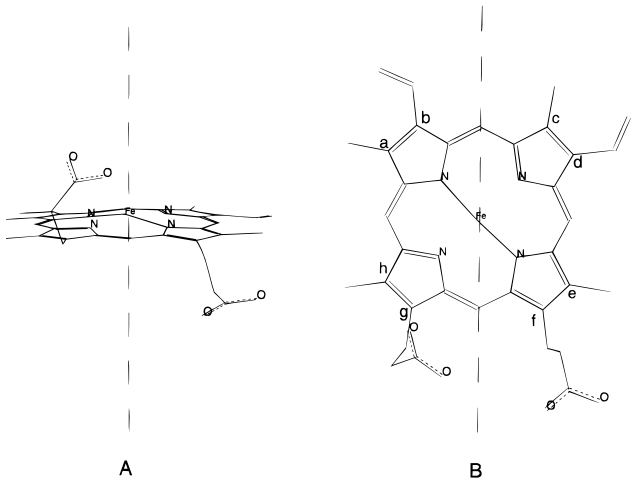

Figure 1. Structure of native protoporphyrin IX. The normal to the porphyrin ring plane (A axis) and the α,γ-meso axis (B axis) are shown. The eight potential sites for the substituents are labeled as a−h.

To rotate the hemes about the α,γ-meso axis (B in Figure 1), two different procedures have been employed. In the first method the FIT procedure of Sybyl was used. FIT is based on the program BMFIT[33] in which the quality of the fit is represented by the root-mean-square deviation computed for selected pairs of matched atoms. For each heme reported in Figure 2, the rotated structure about the α,γ-meso axis was obtained by appropriately superimposing the corresponding N atoms of the porphyrin ring. (This type of α,γ-meso rotation will be called "flipping".) The hemes were thus realigned according to their pyrrole nitrogens, but none of the other porphyrin atoms were perfectly superposed by this procedure, and the iron atom was always badly displaced. The iron of the flipped heme was therefore moved back to its original position with respect to His F8, while the other porphyrin atoms were left in their new positions. After this procedure the flipped molecule was a hybrid, with an unmoved iron, superposed nitrogens, and the other atoms in new places.

A second method for rotating the hemes about the α,γ-meso axis was therefore devised, in which the flipped molecules were built by moving the side chains to appropriate new positions and reattaching them to the ring. This has the advantage of preserving the curvature of the porphyrin and maintaining the iron in its original position, so that the core of each molecule is perfectly superposed.

A total of 52 structures as sketched in Figure 3 were prepared, and all the following computations were carried out in duplicate using each rotation method, so that the importance of exactly aligning the porphyrin cores could be assessed.

Grid Calculations on the 52 Targets. The coordinates of each structure in Figure 3 were carefully checked, and appropriate parameters for the Grid force field were assigned to each atom in each molecule. Ferrous iron was used, and it was assumed that all carboxy groups were fully ionized, so that hemes with two propionate side chains were electrically neutral while hemes with three propionates were monoanions.

In the program Grid, a three-dimensional grid surrounds the target which would be one of the hemes in its original or in one of its rotated orientations. The interaction energy between a probe and each atom of the target was then calculated for each Grid point, and this calculation generated one Grid map. In version 15 of Grid the atoms of the flexible propionate side chains are free to move when the flexibility option is turned on, and their position then always corresponds to the position of lowest interaction energy with the probe at the current grid point.

By contouring the Grid map at various energy levels, one can display favorable interaction areas between the molecule and the probe. Figures 4 and 5 show these Grid contour maps for the N3+ cationic amino group probe in the absence and

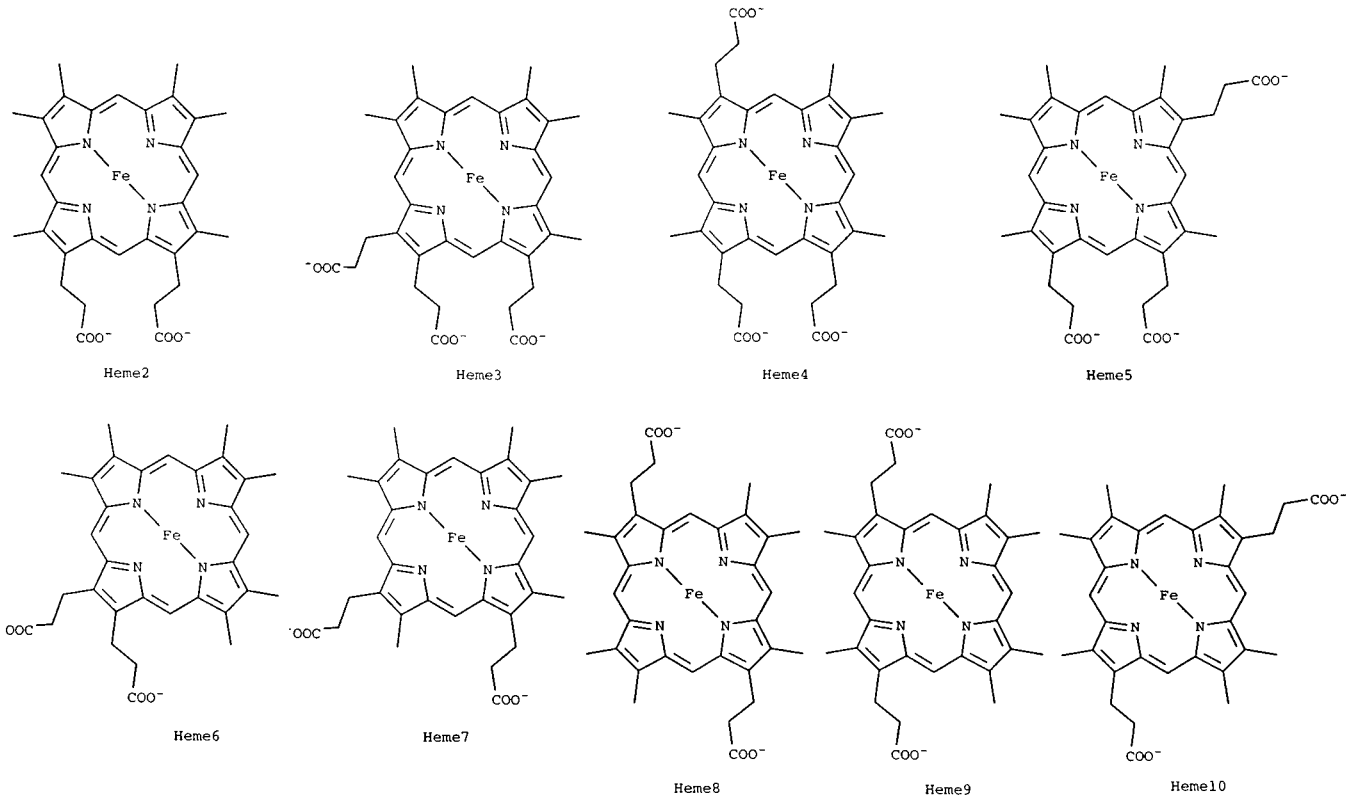

Figure 2. Nine different modified hemes. The numbering of Hauksson[27] is used.

Table 1. List of Amino Acid Residues of Sperm Whale Mb That Have at Least One Atom within 12 Å of the Iron Atom of Heme[a]

residue	no. of residues within 12 Å	probes			
Leu	9	O	N1	C3	
Ile	9	O	N1	C3	
Phe	5	O	N1	C1=	
Ala	3	O	N1	C3	
Val	2	O	N1	C3	
Pro	2	O		C3	
Gly	3	O	N1		
Tyr	2	O	N1	OH	
Thr	4	O	N1	O1	
Ser	2	O	N1	O1	
Gln	1	O	N1	N2	
His	3	O	N1	N:=	N1
Lys	5	O	N1	N3+	
Glu	2	O	N1	O::	CO2−
Asp	2	O	N1	O::	CO2−
Arg	1	O	N1	N2=	N1=

[a] The probes which could simulate these residues are indicated. The carboxylate groups of Asp and Glu could be represented by either O:: or the multiatom CO2- probe.

presence of the flexibility option. Without flexibility of the side chains (Figure 4), there are several weak minima, which are due to hydrogen bond formation between N3+ and the immovable lone pairs on one or other of the carboxy oxygens. On the contrary, in the presence of flexibility (Figure 5), there is a single deeper minimum, since the carboxy groups can move and both can form hydrogen bonds simultaneously. Because of flexibility, therefore, the orientation of the chains does not depend on their initial positions, and the interaction region is symmetrically oriented with respect to the porphyrin ring. In the presence of flexibility, the model is more realistic since the atoms of the chains can move toward or away from the probe, depending on their physical and chemical properties and covalent bonding.

A list of possible probes was prepared (Table 1) corresponding to the functional groups near the heme-binding pocket of myoglobin. All the amino acids having one or more atoms within 12 Å of the heme iron were considered. This table reveals the large number of hydrophobic residues in the cleft but also reflects the presence of polar amino acids and, of course, the polar backbone of the protein. A short list of 10 representative probes was then prepared (Table 2) including the hydrophobic (DRY) probe, and these 10 were used for the computations. The interaction energy between each of the 52 heme structures and each of the 10 probes of Table 2 was computed, giving 520 Grid maps. The Grid program allows one to carry out this calculation in a single step, and a Grid spacing of 1.0 Å was used which ensures a good spatial resolution within a reasonable computational time. The flexibility option was turned on.

Table 2. List of 10 Selected Probes Used for the Computations[a]

		VDRW	*Q*	HD	HA
1, C3	methyl C3 group	1.95	0.0	0	0
2, OH	phenolic hydroxyl group	1.65	−0.15	1	1
3, O1	aliphatic hydroxyl group	1.65	−0.1	1	2
4, N3+	sp3 cationic NH3 group	1.75	0.66	0	3
5, O::	carboxy oxygen atom	1.60	−0.45	0	2
6, N2=	sp2 cationic NH2 group	1.70	0.66	2	0
7, N1=	sp2 cationic NH group	1.65	0.66	1	0
8, N:=	sp2 N with one lone pair	1.65	0.0	0	1
9, N1	amide NH group	1.65	−0.08	1	0
10, DRY	hydrophobic probe	1.70	0.0	2	2

[a] The van der Waals radius (VDRW), charge (*Q*), and number of hydrogen bonds donated (HD) and accepted (HA) are shown for each probe.

Statistical Analysis. Each Grid map contained 24 025 Grid points, and interaction energies were calculated at each Grid point giving, with 520 Grid maps, a very large array of data. Principal component analysis (PCA)[29] is a statistical projection technique which permits features in such data to be identified as principal components, while still retaining much of the information in the original data array. Hi-PCA[30] provides a good method for handling variables in blocks and has the additional advantage of operating on two levels. On

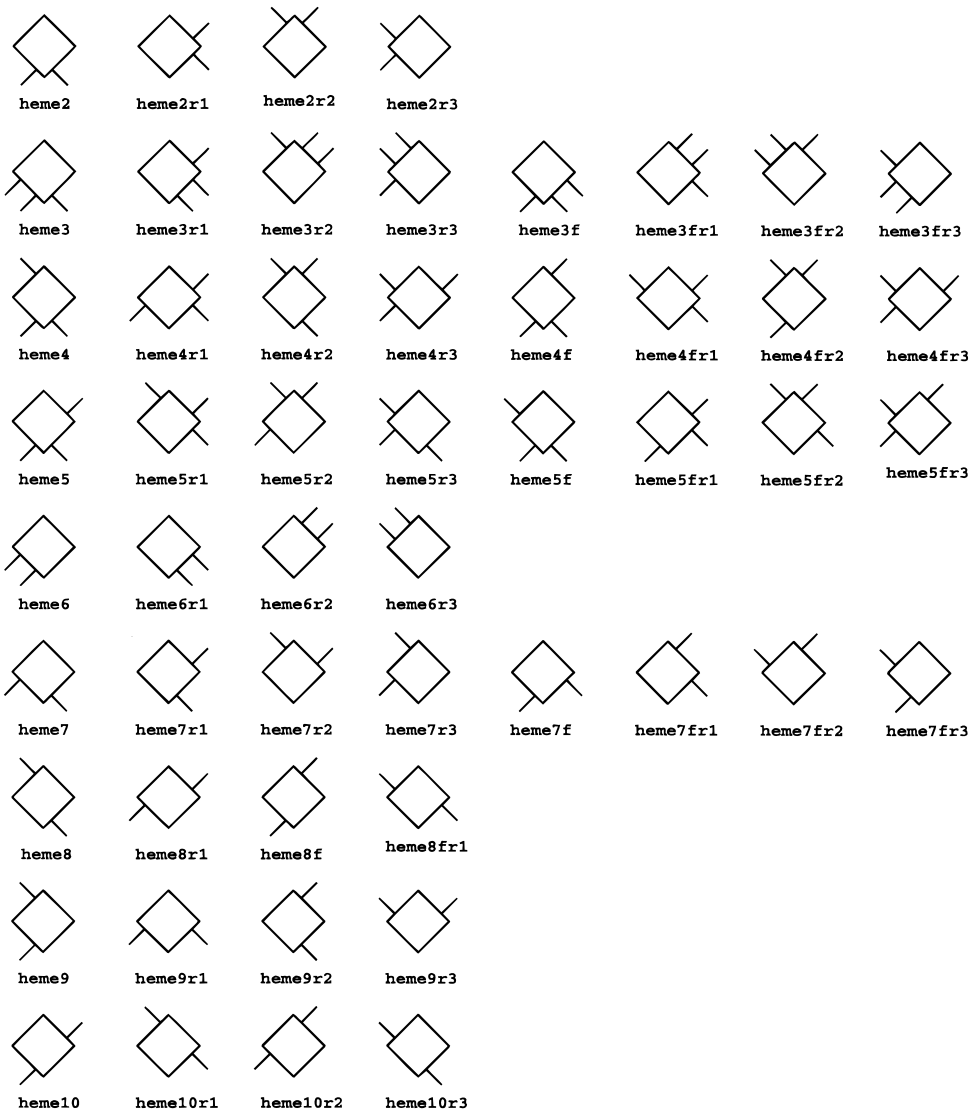

Figure 3. 52 target structures for Grid calculations. The r and f labels indicate the structures obtained by rotation about the normal axis and flipping about the α,γ-meso axis. The numbers 1, 2, and 3 represent rotations of 90°, 180°, and 270°, respectively.

the lower level each block of data may be treated separately taking account of the number of variables and the variance in each individual block. On the higher level the blocks may be related one to the other while still retaining the independent information for each variable. Therefore Hi-PCA allows at one level to assess the importance of each Grid point for each probe, while the relative performance of the individual probes can be determined at the other level.

On both levels plots, loading plots, and score plots provide tools for the analysis of the data set. In the score plot of the upper level it is possible to see how the different molecules are related to each other and to detect misalignments, while the relative importance of each block (i.e., each probe) is revealed by the corresponding upper loading plot.

The output of the Grid calculations on the 52 hemes was therefore analyzed by Hi-PCA. Each probe formed one descriptor block, and each column, in a block, corresponds to a specific position in the 3D-space. A few principal components were found describing the differences among the different heme orientations. The Euclidean distances in Hi-PCA space from heme2 (the reference compound most similar to protoporphyrin IX) to each of the other hemes were calculated (Figure 6), and for each of the eight remaining structures the closest conformation in Hi-PCA space was selected as the predicted binding orientation. This procedure may be compared with other recent approaches,[34] but it is a novel method of aligning and orienting flexible molecules.

Grid Calculations on the Pocket. The program Grid is particularly suitable for the study of proteins[35] and was used in order to investigate the heme-binding pocket of myoglobin after the natural ligand (protoporphyrin IX) had been removed. The hydrophobic probe (called DRY) and the carboxy oxygen probe (O::) were employed, and Grid maps were prepared with the flexibility option of Grid turned off and turned on.

The results are displayed as contour maps showing regions of the empty pocket where the apolar edges of the porphyrin ring, and where the side-chain carboxylate groups, would make favorable interactions (Figures 7−10).

Molecular Mechanics Calculations. The heme structures holding two propionate side chains were bound, each in its predicted orientation, to apomyoglobin. The complexes were minimized using AMBER all-atom force field[36] as implemented in the MSI software package[37] till an rms gradient of 0.01 kcal/mol Å.

Results and Discussion

Statistical Results. Hierarchical PCA models were calculated separately for each of the two alignment methods. Both procedures suggested that the neutral hydrogen bond-donating probes (OH, O1, and N1 of Table 2) were the most important for each computed component.[38] Using the second alignment procedure for

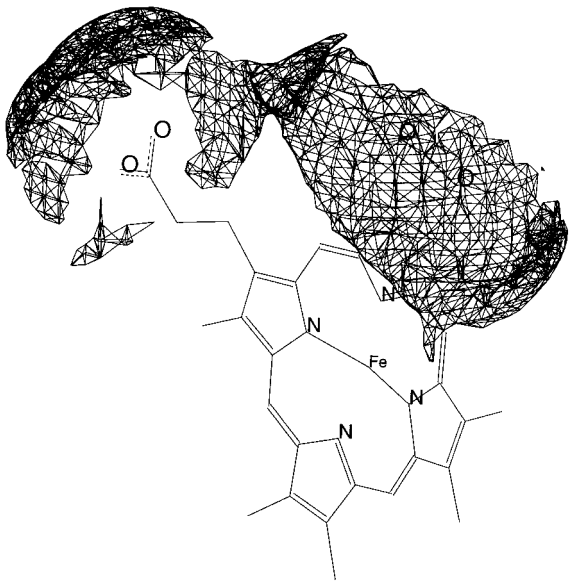

Figure 4. Grid contour map for heme2 at −5.0 kcal/mol with the N3+ probe. Heme2 is the molecule with the same propionate pattern as native heme. For this computation the position of the side chains was fixed, and the interaction area is therefore localized around the carboxylate groups. A Grid spacing of 0.5 Å was used. The contour maps are displayed using Sybyl.

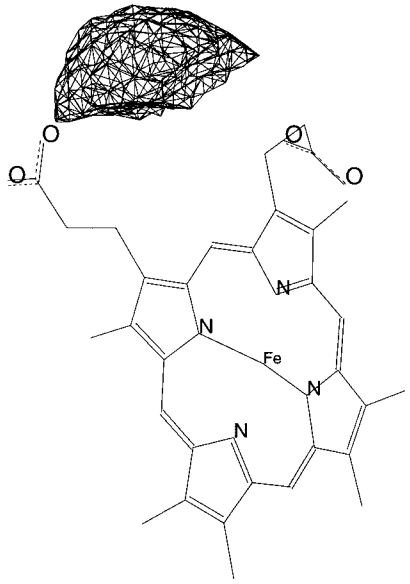

Figure 5. Grid contour map at −15.0 kcal/mol for the N3+ probe with heme2. The side chains were treated as being flexible for this computation. The interaction area is therefore symmetrically placed with respect to the two side chains, and a much stronger interaction is predicted. See text.

the hemes and the first four Hi-PCA components, which explained totally 80% of the variance (28%, 28%, 12%, 12%), the distance plot shown in Figure 6 was obtained.

The first alignment gave a similar plot, showing that the choice of alignment method was not a critical factor, but we preferred the second procedure because the curvature of the heme was correctly maintained. The distances in Figure 6 are measured in Hi-PCA space, and heme2, which represents the propionate pattern of native heme in its observed alignment to myoglobin, is placed at the center of the plot. For each of the eight synthetic hemes, the conformation nearest to the center in Hi-PCA space is most similar to heme2 and therefore represents the predicted orientation which should bind preferentially to apomyoglobin.

Eight orientations of heme3 have been considered and eight orientations of heme4. When these possibilities are combined, one has 64 orientation possibilities for these two molecules alone. When the eight orientations of heme5, the four orientations of heme6, the eight orientations of heme7, the four orientations of heme8, the four orientations of heme9, and the four orientations of heme10 are all considered, there would be 1 030 576 possible ways of aligning the molecules with each other. It would be virtually impossible to consider all of these combinations when one is trying to compare the binding properties of the molecules to myoglobin. However, the present analysis allows the odds to be dramatically improved.

Inspection of Figure 6 shows that heme3 and heme3f are much better than the remaining six orientations of this molecule; similarly heme4 and 4f are also good, and two good orientations can be identified for each of the other compounds giving a reduced number of only 256 possible alignments. More detailed inspection of Figure 6 shows that heme3f is closer to the center than heme3. The difference is small, but comparison with the observations of Hauksson et al. shows that in every case the closest orientation (in the Hi-PCA space of Figure 6) corresponds to an observed mode of binding. In other words one unique alignment of the compounds has been correctly predicted in the present example.

Grid Contour Maps of the Myoglobin Pocket. A Grid contour map at −0.8 kcal/mol was calculated (Figure 7) from the apomyoglobin structure for the hydrophobic (DRY) probe without the flexibility option. The map shows four main areas of interaction (A, B, C, D) which are due to the several hydrophobic side chains of the pocket. These four regions are occupied when protoporphyrin IX binds to apomyoglobin, by the more hydrophobic parts of the protoporphyrin molecule (Figure 8). The hydrophobic contours from apomyoglobin cover much of the porphyrin ring and its hydrocarbon side chains but avoid the iron and its nitrogens and the two polar carboxyl groups, as would be expected.

Somewhat different results are obtained when the flexibility option is activated in version 15 of Grid. This option allows the flexible side chains of the protein to move in response to the probe, and it may be seen (Figure 9) that areas C and D are then enhanced. On the other hand, area A is diminished because polar side chains such as His 97 and His 93 can move into this region, and the overall contrast between the hydrophilic and hydrophobic regions is even more clearly displayed.

The interaction energy between a carboxy oxygen probe (O::) and the empty pocket was also calculated. In Figure 10 a contour map at −12.5 kcal/mol with the flexibility option is shown. This map was prepared from apomyoglobin, but the figure shows the heme molecule superimposed on the contours generated by the empty pocket. The strongest energy of interaction between the O:: probe and the amino acids of the pocket is related to the formation of hydrogen bonds donated by the ND1 hydrogen of His 64, NH1 of Arg 45, and OG1 of Thr 67.

All nine molecules, each in its predicted orientation, were superimposed on the contours calculated from

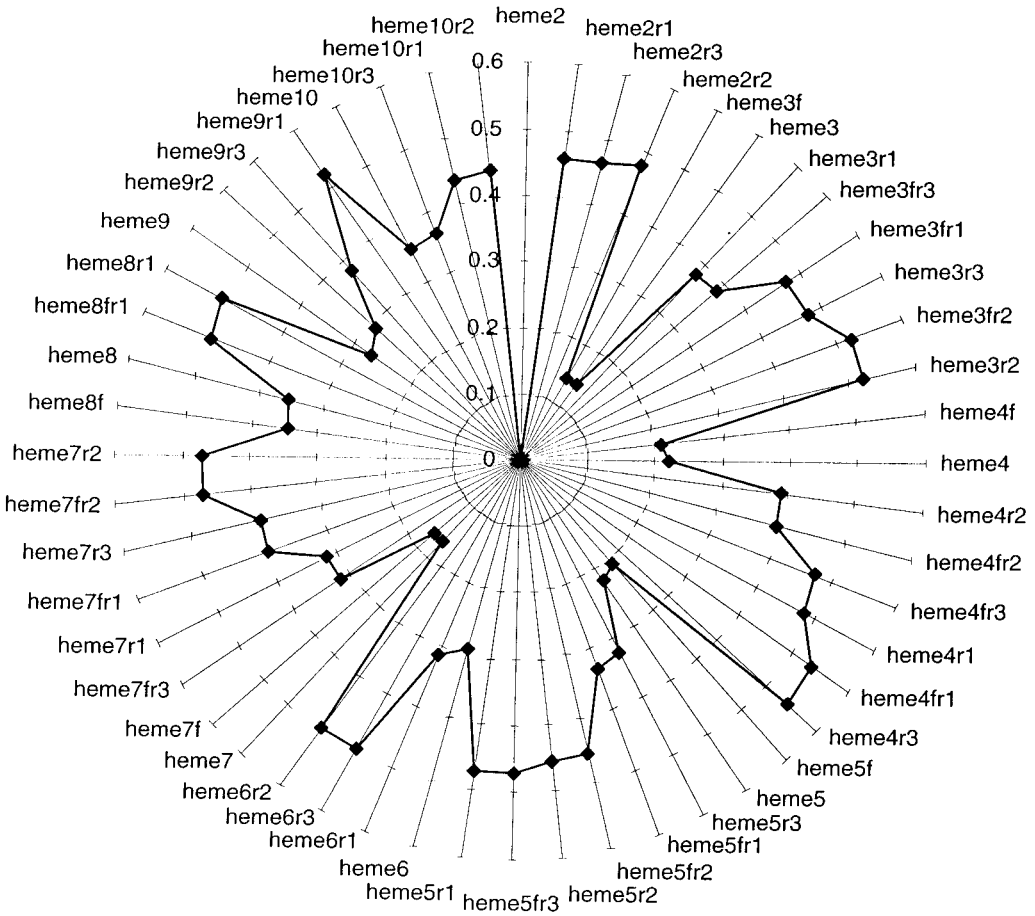

Figure 6. Distance plot between heme2 and the rest of the 51 molecules. The distances are calculated on the first four components.

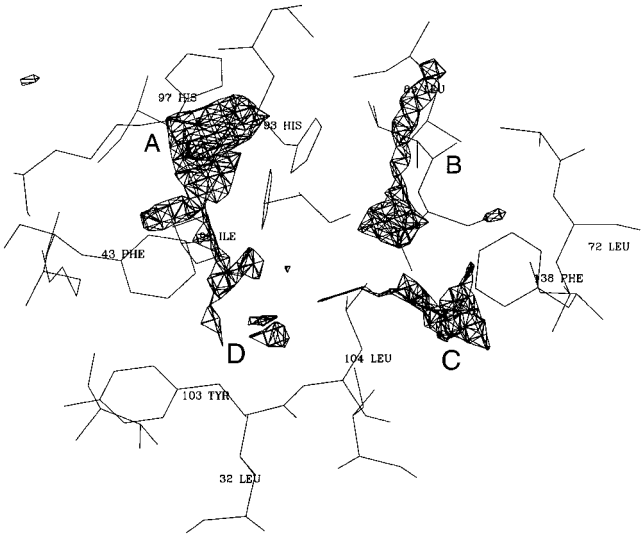

Figure 7. Grid contour maps at −0.8 kcal/mol for the DRY probe at the empty myoglobin heme pocket. The side chains of the amino acids were fixed.

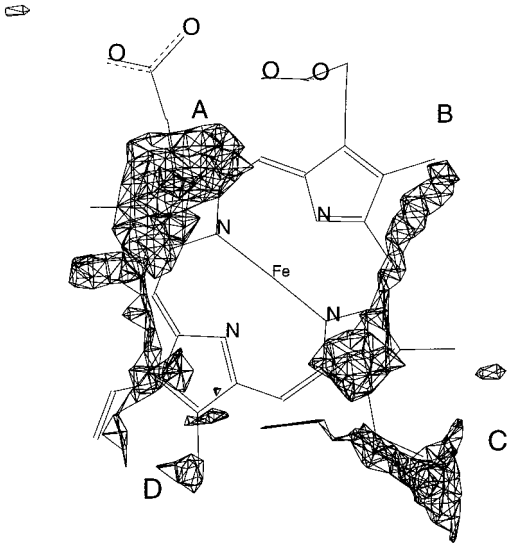

Figure 8. Native heme superimposed on the grid contour map of Figure 7.

apomyoglobin. The heme structures were obtained by molecular mechanics calculations (two propionates) or by dihedral torsions of the chains (three propionates). In every case it may be seen that the hydrophobic contours cover much of the hydrophobic surface of the ligand but that the carboxy oxygens can protrude outside the hydrophobic region of the protein. Moreover the superimposition of each molecule with the carboxy oxygen contours demonstrates that the heme oxygens can always be located in a favorable place. These results suggest why the predicted orientations are compatible with the observations of Hauksson et al. It is because the hydrophobic parts of the hemes tend to go to hydrophobic regions in the myoglobin cleft and the polar parts generally tend to avoid those regions, while the carboxy groups prefer regions where carboxy binding is specifically predicted by Grid.

Conclusions

The present work addresses three important questions which can arise when a set of partly flexible

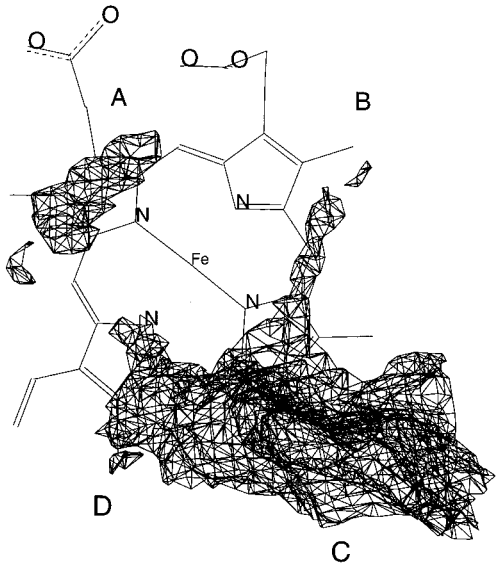

Figure 9. Native heme superimposed on a −1.0 kcal/mol map of the empty myoglobin pocket using the DRY probe. The amino acid side chains were treated as being flexible, and the large hydrophobic region is clearly displayed.

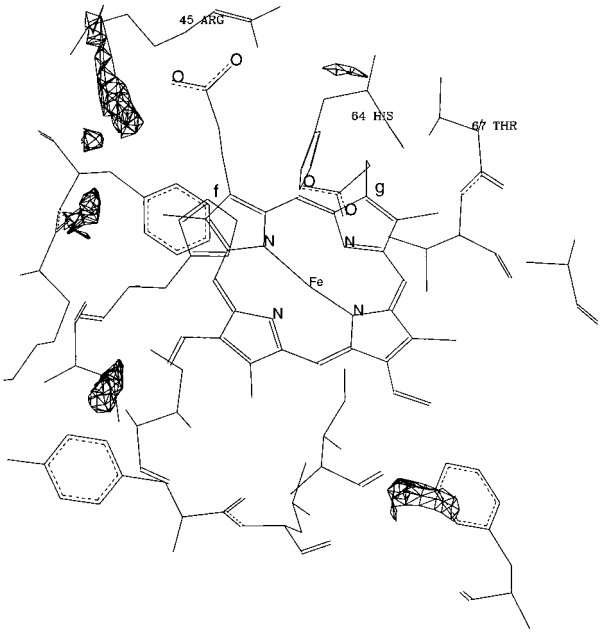

Figure 10. Grid contour maps at −12.5 kcal/mol for the O:: probe at the empty myoglobin pocket. The side chains of the amino acids were treated as being flexible, and the heme structure is superimposed.

molecules are being compared with each other as possible ligands for a biological macromolecule: (1) How can one take account of molecular flexibility? (2) How can one align the molecules with each other? (3) How can one align the set of molecules with their macromolecular binding cleft?

In the case of the nine heme analogues studied experimentally by Hauksson et al., it has been shown that Grid maps can be analyzed by Hi-PCA in order to answer all three questions. Of course this particular example may turn out to be a special case, but the findings suggest that the Grid−Hi-PCA approach may be of more general applicability.

Acknowledgment. The authors are grateful to Dr. P. J. Goodford for his suggestions, helpful discussions, and review of this manuscript. We are greatly indebted to him as well as to the Laboratory of Molecular Biophysics. M.C.D.R. warmly thanks Prof. B. Giardina of the Università Cattolica of Rome for support which made this work possible and Consiglio Nazionale delle Ricerche for a grant.

Supporting Information Available: Two figures showing each heme, in its predicted orientation, superimposed on the DRY and O:: Grid maps generated by apomyoglobin (3 pages). See any current masthead page for ordering information.

References

(1) Dickerson, R. E.; Geis, I. *Hemoglobin*; Benjamin/Cummings: Menlo Park, CA, 1983.
(2) Hargove, M. S.; Olson, J. S. The Stability of Holomyoglobin is Determined by Heme Affinity. *Biochemistry* **1996**, *35*, 11310−11318.
(3) Perutz, H. F.; Pulsinelli, P. D.; Ranney, H. M. Structure and Subunit Interaction of Hemoglobin M Milwaukee. *Nature (London) New Biol.* **1972**, *237*, 259−264.
(4) Steadman, J. H.; Yates, A.; Huehns, E. R. Idiopathic Heinz Body Anaemia: Hb-Bristol (*β*67 (E11) Val → Asp). *Br. J. Haematol.* **1970**, *18*, 435−446.
(5) Varadarajan, R.; Lambright, D. J.; Boxer, S. G. Electrostatic Interactions in Wild-Type and Mutant Recombinant Human Myoglobins. *Biochemistry* **1989**, *28*, 3771−3781.
(6) Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Rohfls, R. J.; Mathews, A. J.; Olson, J. S. Discrimination between Oxygen and Carbon Monoxide and Inhibition of Autoxidation by Myoglobin. *J. Biol. Chem.* **1989**, *264*, 3057−3060.
(7) Rohlfs, R. J.; Mathews, A. J.; Carver, T. E.; Olson, J. S.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G. The Effects of Amino Acid Substitution at Position E7 (Residue 64) on the Kinetics of Ligand Binding to Sperm Whale Myoglobin. *J. Biol. Chem.* **1990**, *265*, 3168−3176.
(8) Ikeda-Saido, M.; Hori, H.; Anderson, L. A.; Prince, R. C.; Pickering, I. J.; George, G. N.; Sanders, C. R., II; Lutz, R. S.; McKelvy, E. J.; Mattera, R. Coordination Structure of the Ferric Heme Iron in Engineered Distal Histidine Myoglobin Mutants. *J. Biol. Chem.* **1992**, *267*, 22843−22852.
(9) Pulsinelli, P. D.; Perutz, M. F.; Nagel, R. L. Structure of Hemoglobin M Boston, a Variant with a Five-Coordinated Ferric Heme. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 3870−3874.
(10) Nagai, K.; Yoneyama, Y.; Kitagawa, T. Characteristics in Tyrosine Coordinations of Four Hemoglobins M Probed by Resonance Raman Spectroscopy. *Biochemistry* **1989**, *28*, 2418−2422.
(11) Hargove, M. S.; Barrick, D.; Olson, J. S. The Association Rate Constant for Heme Binding to Globin is Independent of Protein Structure. *Biochemistry* **1996**, *35*, 11293−11299.
(12) Hill, R.; Holden, M. F. The Preparation and Some Properties of the Globin of Oxyhaemoglobin. *Biochem. J.* **1926**, *20*, 1326.
(13) Gibson, Q. H.; Antonini, E. Rates of Reaction of Native Human Globin with Some Hemes. *J. Biol. Chem.* **1963**, *238*, 1384−1388.
(14) Rose, M. Y.; Olson, J. S. The Kinetic Mechanism of Heme Binding to Human Apohemoglobin. *J. Biol. Chem.* **1983**, *258*, 4298−4303.
(15) Chu, A. H.; Bucci, E. Effect of Polyanions on the Kinetics of the Reaction of Apohemoglobin with Carbonmonoxy Heme. *J. Biol. Chem.* **1979**, *254*, 3772−3776.
(16) Steigemann, W., Weber, E. Structure of Erythrocruorin in Different Ligand States Refined at 1.4 Å Resolution. *J. Mol. Biol.* **1979**, *127*, 309−338.
(17) Moffat, K. Structure−Function Relationships in Hemoglobins: Scientific Aspects. *Texas Rep. Biol. Med.* **1981**, *40*, 191−198.
(18) La Mar, G. N.; Toi, H.; Krishnamoorthi, R. Proton NMR Investigation of the Rate and Mechanism of Heme Rotation in Sperm Whale Myoglobin: Evidence for Intramolecular Reorientation About a Heme Twofold Axis. *J. Am. Chem. Soc.* **1984**, *106*, 6395−6401.
(19) La Mar, G. N.; Davis, N. L.; Parish, D. W.; Smith, K. M. Heme Orientational Disorder in Reconstituted and Native Sperm Whale Myoglobin. *J. Mol. Biol.* **1983**, *168*, 887−896.
(20) La Mar, G. N.; Emerson, S. D.; Lecomte, J. T. J.; Pande, U.; Smith, K. M.; Craig, G. W.; Kehres, L. A. Influence of Propionate Side Chains on the Equilibrium Heme Orientation in Sperm Whale Myoglobin. Heme Resonance Assignments and Structure Determination by Nuclear Overhauser Effect Measurements. *J. Am. Chem. Soc.* **1986**, *108*, 5568−5573.
(21) Livingston, D. J.; Davis, N. L.; La Mar, G. N.; Brown, W. D. Influence of Heme Orientation on Oxygen Affinity in Native Sperm Whale Myoglobin. *J. Am. Chem. Soc.* **1984**, *106*, 3025.

(22) La Mar, G. N.; Smith, K. M.; Gersonde, K.; Sick, H.; Overkemp, M. Proton Nuclear Magnetic Resonance Characterization of Heme Disorder in Monomeric Insect Hemoglobins. *J. Biol. Chem.* **1980**, *255*, 266−270.
(23) Levy, M. J.; La Mar, G. N.; Jue, T.; Smith, K. M.; Pandey, R. K.; Smith, W.; Livingston, D. J.; Brown, W. D. Proton NMR Study of Yellowfin Tuna Myoglobin in Whole Muscle and Solution. *J. Biol. Chem.* **1985**, *260*, 13694−13698.
(24) Cook, R. M.; Wright, P. E. Heme Orientation in the Major Monomeric Hemoglobins of Glycera dibranchiata. *Biochem. Biophys. Acta* **1985**, *832*, 365−372.
(25) La Mar, G. N.; Budd, D. J.; Viscio, D. B.; Smith, K. M.; Langry, K. C. Proton Nuclear Magnetic Resonance Characterization of Heme Disorder in Hemoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *76*, 5755−5759.
(26) Gersonde, K.; Sick, H.; Overkamp, M.; Smith, K. M.; Parish, D. W. Bohr Effect in Monomeric Insect Haemoglobins Controlled by O2 Off-Rate and Modulated by Haem-Rotational Disorder. *Eur. J. Biochem.* **1986**, *157*, 393−404.
(27) Hauksson, J. B.; La Mar, G. N.; Pandey, R. K.; Rezzano, I. N.; Smith, K. M. NMR Study of Heme Pocket Polarity/Hydrophobicity of Myoglobin Using Polypropionate-Substituted Hemins. *J. Am. Chem. Soc.* **1990**, *112*, 8315−8323.
(28) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857. Goodford, P. J. Multivariate Characterization of Molecules for QSAR Analysis. *J. Chemometrics* **1996**, *10*, 107−117.
(29) Jackson, J. E. *A User's Guide to Principal Components*; John Wiley & Sons: New York, 1991.
(30) Wold, S.; Kettaneh, N.; Tjessem, K. Hierarchical Multiblock PLS and PC Models for Easier Model Interpretation and as an Alternative to Variable Selection. *J. Chemometrics* **1996**, *10*, 463−482.
(31) Takano T. Structure of Myoglobin Refined at 2.0 Å Resolution. *J. Mol. Biol.* **1977**, *110*, 569−584.
(32) Sybyl; Tripos Associates, Inc., St. Louis, MO 63144.
(33) Nyburg, S. C. Some Uses of a Best Molecular Fit Routine. *Acta Crystallogr. B* **1974**, *30*, 251−253.
(34) Dunn, W. J., III; Hopfinger, A. J.; Catana, C.; Duraiswami, C. Solution of the Conformation and Alignment Tensors for the Binding of Trimethoprim and its Analogues to Dihydrofolate Reductase: 3D-Quantitative Structure−Activity Relationship Study Using Molecular Shape Analysis, 3-Way Partial Least-Squares Regression, and 3-Way Factor Analysis. *J. Med. Chem.* **1996**, *39*, 4825−4832.
(35) Reynolds, C. A.; Wade, R. C.; Goodford, P. J. Identifying Targets for Bioreductive Agents: Using GRID to Predict Selective Binding Regions of Proteins. *J. Mol. Graph.* **1989**, *7*, 103−108.
(36) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230−252. Heme parameters from Giammona, D. A. Ph.D. Dissertation, University of California at Davis, Davis, CA, 1986.
(37) MSI, 9685 Scranton Rd., San Diego, CA 92121.
(38) Berglund, A.; De Rosa, M. C.; Wold, S. Alignment of Flexible Molecules at their Receptor Site Using 3D Descriptors and Hierarchical-PCA. *J. Comput. Aided Mol. Des.*, in press.

JM9705824