

# Evaluation of the Utility of NMR Structures Determined from Minimal NOE-Based Restraints for Structure-Based Drug Design, Using MMP-1 as an Example

Xuemei Huang, Franklin Moy, and Robert Powers\*

Department of Biological Chemistry, Wyeth Research, 85 Bolton Street, Cambridge, Massachusetts 02140

Received July 17, 2000; Revised Manuscript Received September 6, 2000

**ABSTRACT:** The application of deuterium labeling and residual dipolar coupling constants in combination with other structural information has demonstrated the potential for significantly expanding the range of viable protein targets for structural analysis by NMR. A previous study by Clore et al. [(1999) *J. Am. Chem. Soc.* 121, 6513–6514] demonstrated that a significant improvement in the overall protein structure occurs with the combination of residual dipolar coupling constants and minimal tertiary long-range distance restraints. The analysis of NMR protein structures determined with minimal structural information is extended with a particular interest in the utility of these structures for a structure-based drug design program. As an example, the catalytic fragment of human fibroblast collagenase (MMP-1) was used to follow the effect of minimal restraint sets on the protein structure and its utility in drug design with a particular interest in the effect on the active site conformation. An MMP-1 structure that was calculated with the maximal number of restraints attainable with the constraint of a deuterated protein was shown to be very similar to a high-quality MMP-1 structure that was calculated from a complete set of restraints. The superposition of the active site backbone atoms for the high-quality and minimal restraint MMP-1 structures yielded an rmsd of 0.68 Å where the size and shape of the S1' pocket are nearly identical. Additionally, an MMP-1–CGS-27023A complex based on a minimal set of NOE-based restraints reliably reproduced the structure of the complex, establishing the usefulness of the structures for drug design.

NMR<sup>1</sup> has been established as a powerful method for determining the structure of proteins and protein–ligand complexes (1, 2). The development of heteronuclear three- and four-dimensional methods for both backbone and side chain assignments, as well as editing of the NOE data, has dramatically extended the size range of proteins amenable to structure determination by NMR to 25 kDa (3). Extending the molecular mass barrier for NMR structures beyond the 25 kDa range has required the incorporation of deuterium isotope labeling in addition to the established use of <sup>15</sup>N and <sup>13</sup>C isotope labeling (2, 4). By substitution of the nonexchangeable protons with deuterons, the relaxation times of heteronuclear signals are prolonged, resulting in narrowed line widths and a dramatic increase in resolution and sensitivity (4). This has effectively increased the size limit of protein NMR where a 64 kDa trp repressor–DNA complex has been successfully pursued (5, 6). It is well established that the relative precision and accuracy of a protein structure determined by NMR methodology are inherently dependent on the number and accuracy of the restraints used to determine the structure (7). The underlying

difficulty with deuterium labeling of proteins is the significant loss of structural information. It has been previously demonstrated that protein structures based strictly on long-range NH–NH distance restraints lead to structures with the essentially correct overall fold, but with a significant reduction in the precision and accuracy of the calculated structures (8). The quality of the protein structures was significantly improved by the use of fully deuterated, selectively methyl-protonated samples in observing methyl–methyl and methyl–NH NOEs in addition to the NH–NH distance restraints (8). Recently, novel approaches have been established that provide a means of obtaining structural information independent of traditional H–H NOEs. Most notable is the measurement of residual dipolar couplings in partially oriented proteins dissolved in lipid bicelle solutions (9). Residual dipolar couplings have been applied in the structure calculations of several proteins (10, 11) and have demonstrated a tremendous impact on structures calculated with minimal NOE restraints (12). Toward this end, other approaches have been implemented that are distinct from NOE-based distance restraints to increase the total number of restraints used in the refinement protocol of protein structures determined by NMR in an effort to improve the quality of NMR structures of high-molecular mass proteins (13). These methods have included the direct refinement against chemical shifts (14, 15), coupling constants (16), a conformational database potential (17, 18), and the radius of gyration (19).

The rational design of protein inhibitors based on structural information has proven to be an extremely valuable method for drug development (for reviews, see refs 20 and 21).

\* To whom correspondence should be addressed: Wyeth Research, 85 Bolton St., Room 222B, Cambridge, MA 02140. Phone: (617) 665-7997. Fax: (617) 665-8993. E-mail: powersr@war.wyeth.com.

<sup>1</sup> Abbreviations: DHPC, dihexanoylphosphatidylcholine; DMPC, dimyristoylphosphatidylcholine; MMP-1, matrix metalloproteinase; NMR, nuclear magnetic resonance; 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; HSQC, heteronuclear single-quantum coherence spectroscopy; TPPI, time-proportional phase incrementation; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhanced spectroscopy.

However, a fundamental component of the structure-based approach to drug development is the inherent reliance on accurate protein–ligand complex structures to guide the chemical design and synthesis. There is considerable interest in attempting to determine global protein folds on the basis of minimal numbers of NOEs to speed the drug discovery process. These attempts have only realized partial success due to the relatively low accuracy of the resulting structure (8). Additionally, high-molecular mass proteins are more likely to be potential targets in a drug discovery program, which poses a problem because of the reduced structural information inherent in deuterium labeling. Previous work by Clore et al. (12) has demonstrated that the addition of residual dipolar coupling constants to structures determined with minimal NOE-based distance restraints significantly improved the overall accuracy and precision of the protein structures, based on the improvement in the backbone rmsd (12).

Here we extend the analysis of NMR protein structures determined with minimal structural information with a particular interest in the utility of these structures for a structure-based drug design program. For this study, the catalytic fragment of human fibroblast collagenase (MMP-1) was used to follow the effect of varying minimal restraint sets on the protein structure. The solution structures of inhibitor-free MMP-1 and MMP-1 complexed with a hydroxamic inhibitor have been previously determined which provides the necessary reference point and structural restraint for such an analysis (22, 23). MMP-1 is a member of the matrix metalloproteinase (MMP) family that is involved in the degradation of the extracellular matrix (24, 25). Because of the widespread implications of MMPs in arthritis, cancer, and cardiovascular disease, these enzymes have become important targets for inhibitor design and synthesis.

## MATERIALS AND METHODS

**NMR Sample Preparation.** Uniformly  $^{15}\text{N}$ - and  $^{15}\text{N}/^{13}\text{C}$ -labeled human recombinant MMP-1 was expressed in *Escherichia coli* and purified as described previously (26). The purified protein was concentrated to 0.5 mM and exchanged into a buffer containing 10 mM deuterated Tris base (pH 6.5), 20 mM NaCl, 5 mM  $\text{CaCl}_2$ , 0.1 mM  $\text{ZnCl}_2$ , 2 mM  $\text{NaN}_3$ , and 5 mM deuterated DTT in a 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  mixture. The liquid-crystalline medium, comprising a mixture (3.25:1) of dimyristoylphosphatidylcholine (DMHC) and dihexanoylphosphatidylcholine (DHPC) bicelles (purchased from Avanti Polar Lipids, Inc.), was prepared and added to a final level of 5% w/v as described previously (27).

**NMR Data Collection.** All spectra were recorded at 35 °C on a Bruker AVANCE 600 MHz spectrometer using a gradient-enhanced triple-resonance  $^1\text{H}/^{13}\text{C}/^{15}\text{N}$  probe. Quadrature detection in the indirectly detected dimensions was carried out with States–TPPI hypercomplex phase incrementation (28).  $^1D_{\text{NH}}$ ,  $^1D_{\text{NC}'}$ ,  $^2D_{\text{HNC}'}$ ,  $^1D_{\text{CaC}'}$ , and  $^3D_{\text{HNH}\alpha}$  residual dipolar couplings were obtained in lipid bicelles and isotropic MMP-1.  $^1J_{\text{NH}}$  coupling was achieved with a two-dimensional IPAP–HSQC experiment to generate two spectra containing either the upfield or downfield  $^{15}\text{N}$  doublet component (29).  $^1J_{\text{NC}'}$  and  $^2J_{\text{HNC}'}$  couplings were obtained from a 2D  $^{13}\text{C}'$ -coupled/ $^{13}\text{C}_\alpha$ -decoupled  $^1\text{H}$ – $^{15}\text{N}$  HSQC

experiment (30). The  $^1J_{\text{CaC}'}$  coupling was obtained from a 3D HNCO experiment without  $^{13}\text{C}_\alpha$  decoupling during  $^{13}\text{C}'$  evolution (31). The  $^3J_{\text{HNH}\alpha}$  coupling was obtained from a 3D HNCA experiment (32). The measured  $^1D_{\text{NH}}$  values ranged from  $-23$  to  $19$  Hz, and the normalized factors (given by  $\gamma_{\text{N}}\gamma_{\text{H}}^{-3}/\gamma_{\text{A}}\gamma_{\text{B}}^{-3}$ , where  $\gamma$  and  $r$  represent gyromagnetic ratios and distances, respectively) employed for  $^1D_{\text{NC}'}$ ,  $^2D_{\text{HNC}'}$ ,  $^1D_{\text{CaC}'}$ , and  $^3D_{\text{HNH}\alpha}$  relative to  $^1D_{\text{NH}}$  were 9.04, 3.04, 5.36, and 2.30, respectively. The magnitudes of the axial and rhombic components of the alignment tensor  $\mathbf{D}_{\text{NH}}$  were obtained by examining the distribution of normalized dipolar couplings which yielded values for  $D_{\text{a}}^{\text{NH}}$  of  $-10.8$  Hz and an  $R$  of 0.59, where  $D_{\text{a}}^{\text{NH}}$  is the axial component of the tensor and  $R$  is the rhombicity defined as the ratio of the rhombic to axial components of the tensor (33).

Spectra were processed using the NMRPipe software package (34) and analyzed with PIPP (35). When appropriate, data processing included zero-padding data to a power of 2, linear predicting back one data point of indirectly acquired data to obtain zero phase corrections, and linear prediction of additional points for the indirectly acquired dimensions to increase resolution.

**Structure Calculations.** The structures were calculated using the hybrid distance geometry–dynamical simulated annealing method of Nilges et al. (36) with minor modifications using the program XPLOR (37), adapted to incorporate pseudopotentials for secondary  $^{13}\text{C}_\alpha/^{13}\text{C}_\beta$  chemical shift restraints (14), a conformational database potential (17, 18), residual dipolar coupling constants (38), and the radius of gyration (19). Interproton distance restraints, torsion angle restraints, and stereospecific assignments were the same as those used for the refinement of the inhibitor-free MMP-1 NMR structure (22). Calculations for the MMP-1 structure with the minimal restraint set include only residual dipolar coupling restraints, the radius of gyration,  $\text{C}\alpha/\text{C}\beta$  chemical shift restraints, and H-bond restraints for only regions of secondary structure.  $^{15}\text{N}$ ,  $^{13}\text{CO}$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$ , and NH chemical shifts were utilized as input for the program TALOS to generate chemical shift-based dihedral angle restraints for  $\phi$  and  $\psi$  angles (39). The backbone  $\phi$  and  $\psi$  angle restraints were used when the predicted angles were consistent in all 10 structures from TALOS. Afterward, NH–NH, NH–methyl, methyl–methyl, aromatic–NH, aromatic–methyl, and aromatic–aromatic NOE restraints were added to the structure calculations in a stepwise manner. For the calculation of the MMP-1–CGS-27023A complex, the above restraints were used, amended with the 18 intramolecular restraints observed for CGS-27023A and 25 intermolecular distance restraints chosen on the basis of the same criteria described above (23). The bound conformation for CGS-27023A was generated using QUANTA 97 and CHARMM (Molecular Simulations Inc., San Diego, CA), and the XPLOR topology and parameter files were generated using XPLOR2D, as described previously (23). The MMP-1–CGS-27023A complex structure was determined using the same refinement protocol as described previously (23).

**Computer Models of MMPs.** Molecular modeling was carried out using the Sybyl molecular modeling package from Tripos Inc. (St. Louis, MO) on a Silicon Graphics workstation. The S1' pockets were calculated for each set of MMP-1 structures as well as the MMP-1–CGS-27023A complex.

The S1' pocket were represented as Connolly solvent accessible surfaces and generated with the MOLCAD algorithm from Tripos Inc. using a 1.4 Å water probe.

## RESULTS AND DISCUSSION

**Overall Goal.** The application of deuterium labeling has made a significant contribution to expanding the molecular mass range for proteins amenable to analysis by NMR, but with a detrimental impact on the accuracy and precision of the calculated structures as a result of the loss of NOE-based structural information. The incorporation of residual dipolar coupling constants in addition to other structural restraints not dependent on the presence of  $^1\text{H}$  resonances has partially compensated for the lack of distance restraints. This is evident in the observed decrease in the backbone rmsd values relative to the high-resolution structure and the ensemble of NMR structures, respectively (8, 12). These results clearly establish the inherent value of residual dipolar coupling constants in determining NMR structures, but do not address the utility of these structures in applications such as a structure-based drug design. In particular, residual dipolar coupling constants determined for NH groups only provide structural information related to the conformation of the protein's backbone. This is also true for the torsion angle restraints, secondary  $^{13}\text{C}\alpha/^{13}\text{C}\beta$  chemical shift restraints, and hydrogen bond restraints that can potentially be observed in a deuterium-labeled protein. Possible information for side chain conformations may be provided from the conformational database and radius of gyration target functions in addition to the minimal distance restraints experimentally observed. Furthermore, it is plausible that side chain conformations may be constrained or defined by the effective packing of the protein determined by the backbone conformation.

The utilization of a structure of a protein–ligand complex for drug design necessitates a high-quality structure in the vicinity of the ligand for identifying beneficial interactions between the ligand and protein while also identifying potential interactions with the protein that may be incorporated in further refinements of the ligand. Not unexpectedly, a number of the observed or potential beneficial interactions between the protein and the ligand will arise from contact with side chain atoms as opposed to backbone atoms. Additionally, features such as the shape, size, and electrostatic environment of the protein active site that are dependent on the side chain conformations will also impact synthetic design.

In an effort to address the practical utility of a structure that can be obtained from a deuterium-labeled protein, we have used the structural data previously obtained for MMP-1 and MMP-1 complexed with a hydroxamic acid inhibitor to simulate various minimal restraint sets. MMP-1 is a matrix metalloproteinase with a well-defined active site composed of an S1' pocket that has been effectively used for the design of potent inhibitors (40–42). Our goal is to simulate structural restraints that are realistically obtainable given the current technology for a high-molecular mass protein that is fully deuterium labeled and investigate the impact on the conformation of the active site using MMP-1 as an example. Comparison of the MMP-1 structure calculated with the complete set of NMR restraints with the various structures

determined with minimal restraint sets will allow for the analysis of the consistency of the MMP-1 active site relative to the high-resolution structure and determine the potential impact on drug design.

**Solution Structure of MMP-1 and Comparison to Earlier Structures.** The solution structure of inhibitor-free human fibroblast collagenase as well as the collagenase complexed with a sulfonamide derivative of a hydroxamic acid compound have been determined to high resolution (22, 23). Besides all the NMR restraints used in the structure calculation for inhibitor-free human fibroblast collagenase (22), more than 500 residual dipolar coupling restraints were included in the structure calculation of MMP-1. Among these, 90 dipolar coupling restraints are within the active site residues. A radius of gyration of 16.1 Å was calculated using XPLOR, and was incorporated into the calculations as described previously (19). Inclusion of dipolar coupling restraints in the structure calculation required reevaluation of the NOE restraints used in the previous structure determination. There were no NOE violations larger than 0.2 Å observed after incorporation of the dipolar coupling restraints, so modifications to the NOE restraints used for the previous refinement of inhibitor-free MMP-1 were unnecessary. A summary of the structural statistics for the final 30 simulated annealing (SA) structures of human MMP-1 based on the complete set of structural restraints with the incorporation of the dipolar coupling restraints is provided in Table 1 (set 9). Figure 1E shows the best-fit superposition of the backbone atoms. The inclusion of residual dipolar coupling restraints in the structure refinement increases the coordinate precision from 0.42 to 0.30 Å for backbone atoms for residues 7–137 and 145–163 and from 0.80 to 0.73 Å for all heavy atoms (Table 2, set 9). The rmsd for secondary structure elements also drops from 0.28 to 0.19 Å for backbone atoms. Also, the rmsd for the active residues drops from 0.54 to 0.37 Å. A check of the 30 conformers using PROCHECK-NMR shows that 87.5% of the residues are in the most favored region of the Ramachandran plot where no structure has residues in the disallowed regions (43). The pairwise rmsd between the average structures calculated with and without dipolar coupling restraints is 0.13 Å for backbone residues 7–137 and 145–163, indicating that the two structures are indistinguishable.

**Structure Calculations Using Minimal NOE Restraints.** In the first set of structure calculations, we have included only the dipolar coupling restraints, together with  $\text{C}\alpha/\text{C}\beta$  chemical shifts, limited  $\phi$  and  $\psi$  torsion angle restraints obtained from TALOS, and H-bond restraints in the well-defined secondary structure region. Backbone assignments for deuterated proteins can be obtained in a manner similar to that used for uniformly  $^{13}\text{C}/^{15}\text{N}$ -labeled proteins.  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ , and  $^{13}\text{CO}$  chemical shifts can be used to quickly and qualitatively identify an  $\alpha$ -helix or  $\beta$ -strand conformation using the Chemical Shift Index (CSI) method (44) after suitable correction for deuterium isotope effects. The chemical shift information can be utilized as input for the program TALOS to generate chemical shift-based dihedral angle restraints for  $\phi$  and  $\psi$  angles (39). H-Bond restraints are readily identified by the observation of slow exchanging amides in 2D  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectra after transferring the protein into a  $\text{D}_2\text{O}$  solution. In addition to the experimental restraints, the MMP-1 structure was refined using both the conformational

Table 1: Structural Statistics for Each Set of Structure Calculations<sup>a</sup>

	set 1	set 2	set 3	set 4	set 5	set 6	set 7	set 8	set 9
rms deviations									
all NOE		0.118 ± 0.027	0.118 ± 0.027	0.065 ± 0.013	0.063 ± 0.013	0.063 ± 0.013	0.055 ± 0.009	0.058 ± 0.014	0.019 ± 0.002
intraresidue			0.139 ± 0.032	0.091 ± 0.020	0.049 ± 0.030	0.039 ± 0.013	0.048 ± 0.026	0.042 ± 0.018	0.008 ± 0.004
sequential ( $ i - j  = 1$ )		0.139 ± 0.032	0.091 ± 0.026	0.082 ± 0.030	0.066 ± 0.024	0.068 ± 0.027	0.063 ± 0.027	0.071 ± 0.028	0.008 ± 0.004
medium-range ( $1 <  i - j  \leq 5$ )		0.079 ± 0.044	0.066 ± 0.026	0.058 ± 0.027	0.052 ± 0.022	0.054 ± 0.023	0.055 ± 0.022	0.059 ± 0.025	0.019 ± 0.003
long-range ( $ i - j  > 5$ )		0.068 ± 0.058	0.068 ± 0.032	0.046 ± 0.014	0.064 ± 0.017	0.042 ± 0.013	0.041 ± 0.008	0.042 ± 0.014	0.022 ± 0.002
H bonds (84)	0.11 ± 0.05	0.094 ± 0.049	0.067 ± 0.029	0.067 ± 0.029	0.059 ± 0.016	0.052 ± 0.008	0.053 ± 0.013	0.060 ± 0.023	0.036 ± 0.004
dihedral restraints (deg) <sup>e</sup>	1.43 ± 0.63	1.28 ± 0.42	0.82 ± 0.038	0.96 ± 0.54	0.64 ± 0.46	0.71 ± 0.51	0.66 ± 0.64	0.54 ± 0.46	0.71 ± 0.07
C $\alpha$ restraints (ppm) (140)	1.43 ± 0.11	1.44 ± 0.08	1.37 ± 0.09	1.36 ± 0.09	1.30 ± 0.08	1.31 ± 0.07	1.29 ± 0.017	1.31 ± 0.07	1.11 ± 0.01
C $\beta$ restraints (ppm) (125)	1.37 ± 0.08	1.35 ± 0.08	1.30 ± 0.07	1.30 ± 0.06	1.26 ± 0.06	1.26 ± 0.04	1.27 ± 0.05	1.26 ± 0.04	1.20 ± 0.03
residual dipolar couplings (Hz) <sup>b</sup>									
<sup>1</sup> D <sub>NH</sub> (Hz) (126)	0.86 ± 0.14	0.86 ± 0.11	0.75 ± 0.06	0.73 ± 0.08	0.77 ± 0.08	0.73 ± 0.08	0.72 ± 0.09	0.75 ± 0.06	0.81 ± 0.02
<sup>1</sup> D <sub>C<math>\alpha</math>C'</sub> (Hz) (133)	1.62 ± 0.17	1.65 ± 0.15	1.42 ± 0.12	1.37 ± 0.10	1.27 ± 0.10	1.22 ± 0.08	1.26 ± 0.10	1.23 ± 0.09	0.64 ± 0.02
<sup>1</sup> D <sub>NC'</sub> (Hz) (106)	0.73 ± 0.08	0.74 ± 0.06	0.66 ± 0.04	0.64 ± 0.04	0.67 ± 0.05	0.65 ± 0.04	0.65 ± 0.04	0.66 ± 0.04	0.55 ± 0.01
<sup>2</sup> D <sub>HNC'</sub> (Hz) (106)	1.35 ± 0.14	1.37 ± 0.13	1.25 ± 0.12	1.16 ± 0.09	1.25 ± 0.11	1.19 ± 0.10	1.21 ± 0.12	1.26 ± 0.10	1.02 ± 0.04
<sup>3</sup> D <sub>HNH<math>\alpha</math></sub> (Hz) (105)	1.36 ± 0.11	1.37 ± 0.10	1.27 ± 0.09	1.26 ± 0.09	1.22 ± 0.10	1.22 ± 0.10	1.23 ± 0.07	1.28 ± 0.09	0.92 ± 0.02
F <sub>NOE</sub> (kcal mol <sup>-1</sup> )		164.4 ± 81.6	134.0 ± 47.9	117.1 ± 42.3	123.0 ± 51.0	112.5 ± 63.2	112.5 ± 39.5	140.0 ± 65.1	49.1 ± 8.9
F <sub>tor</sub> (kcal mol <sup>-1</sup> )	21.5 ± 20.5	16.1 ± 11.1	7.18 ± 6.44	10.6 ± 11.9	5.50 ± 8.62	6.83 ± 9.77	7.53 ± 12.7	4.45 ± 8.17	12.6 ± 2.4
F <sub>repeal</sub> (kcal mol <sup>-1</sup> )	92.1 ± 21.9	98.8 ± 48.9	84.2 ± 15.8	74.2 ± 13.9	74.2 ± 13.9	74.2 ± 10.2	76.2 ± 10.7	79.8 ± 13.9	41.5 ± 3.1
F <sub>L-J</sub> (kcal mol <sup>-1</sup> ) <sup>c</sup>	-384.9 ± 25.9	-391.7 ± 23.6	-422.0 ± 28.5	-448.4 ± 24.9	-467.4 ± 29.0	-492.6 ± 28.1	-485.4 ± 25.3	-511.3 ± 26.6	-621.4 ± 11.1
deviations from idealized covalent geometry									
bonds (Å)	0.005 ± 0.00	0.005 ± 0.00	0.005 ± 0.00	0.004 ± 0.00	0.004 ± 0.00	0.004 ± 0.00	0.004 ± 0.00	0.003 ± 0.00	0.004 ± 0.00
angles (deg)	0.72 ± 0.05	0.72 ± 0.05	0.70 ± 0.03	0.68 ± 0.04	0.69 ± 0.04	0.68 ± 0.04	0.68 ± 0.03	0.59 ± 0.04	0.57 ± 0.001
impropers (deg) <sup>d</sup>	0.69 ± 0.06	0.69 ± 0.07	0.64 ± 0.05	0.64 ± 0.05	0.64 ± 0.05	0.63 ± 0.04	0.64 ± 0.03	0.57 ± 0.07	0.55 ± 0.02
PROCHECK_NMR Ramachandran plot									
most favored	64.7%	64.4%	68.2%	71.1%	73.6%	75.2%	75.8%	73.8%	87.5%
additional allowed	21.3%	21.7%	18.4%	17.5%	15.6%	15.1%	14.4%	16.1%	11.6%
generously allowed	9.3%	8.9%	8.3%	7.6%	7.2%	6.5%	6.8%	7.1%	0.9%
disallowed	4.8%	5.0%	5.0%	3.8%	3.2%	3.2%	3.2%	3.0%	0.0%

<sup>a</sup> Set 1 (the minimal set of constraints) includes five sets of dipolar couplings, the radius of gyration, C $\alpha$ /C $\beta$  chemical shifts, the H-bond for secondary structure elements, Ramachandran, Zn, and Ca constraints, and  $\phi$  and  $\psi$  angle constraints from good TALOS prediction. Set 2 includes set 1 and 142 NH–NH NOE constraints. Set 3 includes set 2 and 205 NH–methyl constraints. Set 4 includes set 3 and 87 methyl–methyl constraints. Set 5 includes set 4 and 72 NH–aromatic NOE constraints. Set 6 includes set 5 and 123 methyl–aromatic NOE constraints. Set 7 includes set 6 and 11 aromatic–aromatic NOE constraints. Set 8 is the same as set 7, with CGS in the active site. Set 9 includes a full set of NOE distance restraints, the residual dipolar coupling restraints, and a radius of gyration of 16.1 Å. <sup>b</sup> The number of dipolar coupling restraints and C $\alpha$ /C $\beta$  chemical shift restraints is given in parentheses. The force constant employed for the <sup>1</sup>D<sub>NH</sub> dipolar coupling restraints is 1.0 kcal mol<sup>-1</sup> Hz<sup>-2</sup>. <sup>c</sup> E<sub>L-J</sub> is the Lennard-Jones–van der Waals energy calculated with the CHARMM empirical energy function and is not included in the target function for simulated annealing or restrained minimization. <sup>d</sup> The overall quality of the structure was assessed using the program PROCHECK\_NMR. <sup>e</sup> A total of a 162 dihedral restraints used for set 1–8, and 413 dihedral restraints were used for set 9.

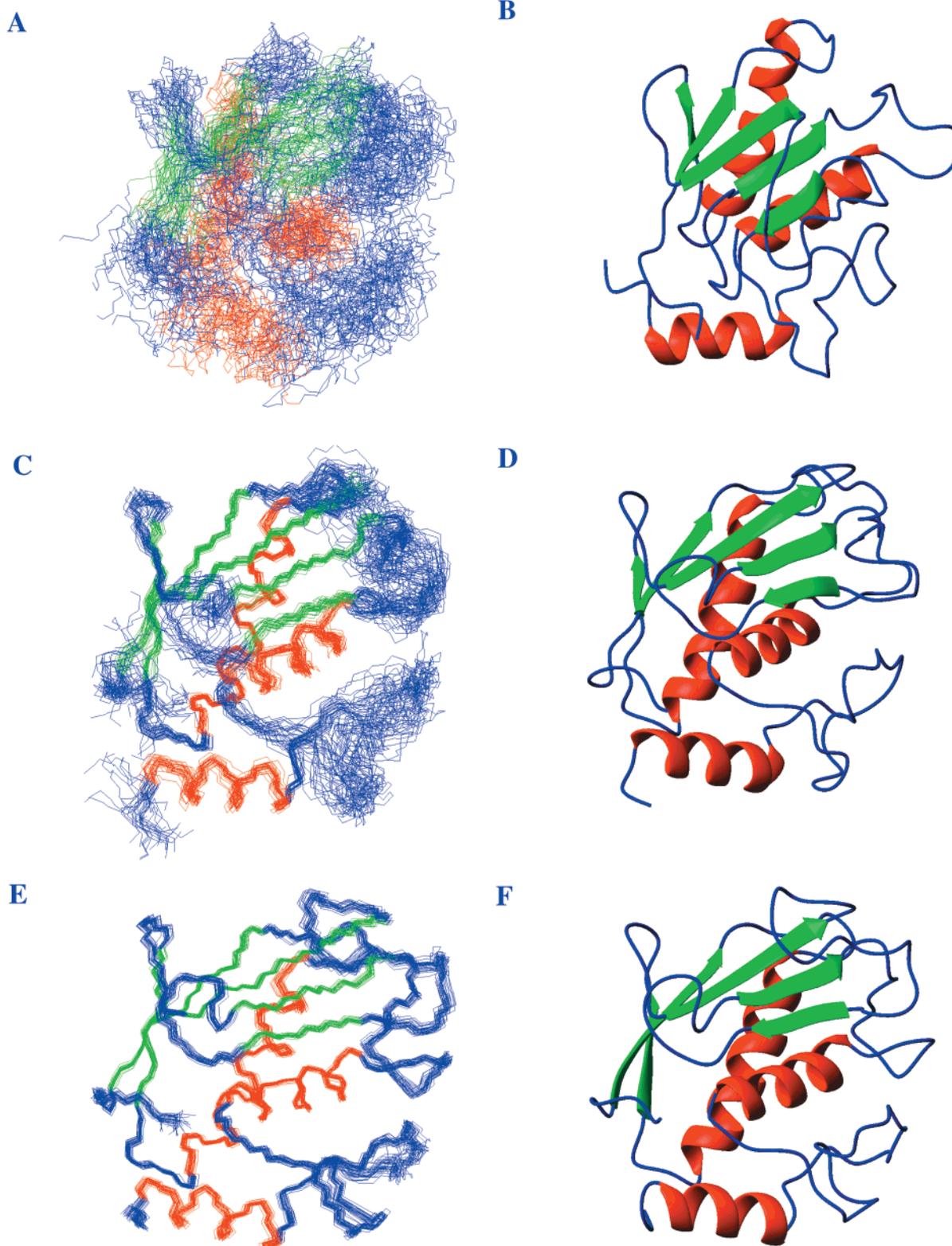


FIGURE 1: Superposition of the 30 final simulated annealing structures (A, C, and E) and ribbon diagram of the MMP-1 structure (B, D, and F). The five  $\beta$ -strands are shown in green, and the three  $\alpha$ -helices are shown in red. Structures A and B were calculated from the minimal set of restraints (set 1). Structures C and D were calculated from the maximal set of restraints (set 7) for a deuterated protein. Structures E and F were calculated from the complete set of restraints (set 9) for MMP-1.

database (17, 18) and the radius of gyration (19) target functions. The calculated structures have well-defined secondary structure elements (Figure 1B); however, due to the lack of NOE distance restraints, the structures do not converge well as shown in Figure 1A. The atomic rms distribution of the 30 simulated annealing structures about

the mean coordinate positions for residues 7–137 and 145–163 is  $4.24 \pm 0.31$  Å for backbone atoms and  $4.93 \pm 0.35$  Å for all atoms (Table 2, set 1).

In the next step, we added 142 NH–NH NOE distance restraints that can be obtained easily from a fully deuterated sample to the structural restraints already defined in set 1.

Table 2: Atomic rms Differences for the Final 30 Simulated Annealing Structures over the Mean Structure Obtained by Averaging the Coordinates of the Individual SA Structures of Each Set

		atomic rms difference (Å) <sup>a</sup>							
		active site residues <sup>b</sup>		active site residues <sup>c</sup>		residues 7–137 and 145–163		secondary structure <sup>d</sup>	
CGS <sup>e</sup>		backbone	all	backbone	all	backbone	all	backbone	all
set 1 <sup>f</sup>		3.76 ± 0.69	4.56 ± 0.75	2.12 ± 0.70	2.90 ± 0.88	4.24 ± 0.31	4.93 ± 0.35	3.00 ± 0.60	3.63 ± 0.70
set 2		3.67 ± 0.69	4.57 ± 0.77	2.13 ± 0.63	2.94 ± 0.86	4.03 ± 0.33	4.71 ± 0.34	2.75 ± 0.53	3.39 ± 0.55
set 3		2.11 ± 0.45	3.03 ± 0.44	0.98 ± 0.31	1.76 ± 0.35	2.29 ± 0.18	3.09 ± 0.20	1.19 ± 0.28	1.93 ± 0.27
set 4		1.92 ± 0.40	2.73 ± 0.37	0.86 ± 0.26	1.56 ± 0.27	2.11 ± 0.17	2.89 ± 0.17	0.83 ± 0.14	1.60 ± 0.13
set 5		1.55 ± 0.28	2.18 ± 0.26	0.58 ± 0.18	1.23 ± 0.21	1.70 ± 0.09	2.31 ± 0.12	0.73 ± 0.10	1.40 ± 0.13
set 6		1.37 ± 0.37	1.91 ± 0.33	0.53 ± 0.12	1.10 ± 0.14	1.49 ± 0.13	2.04 ± 0.13	0.56 ± 0.08	1.16 ± 0.08
set 7		1.31 ± 0.27	1.89 ± 0.31	0.54 ± 0.15	1.07 ± 0.18	1.56 ± 0.12	2.09 ± 0.12	0.55 ± 0.11	1.14 ± 0.09
set 8	0.98 ± 0.42	1.31 ± 0.32	1.83 ± 0.33	0.49 ± 0.16	0.94 ± 0.19	1.49 ± 0.14	2.05 ± 0.15	0.56 ± 0.09	1.15 ± 0.11
set 9		0.37 ± 0.11	0.67 ± 0.13	0.18 ± 0.04	0.61 ± 0.13	0.30 ± 0.03	0.73 ± 0.04	0.19 ± 0.02	0.61 ± 0.04

<sup>a</sup> Defined as the rms difference between the final 30 simulated annealing structures and the mean structure from each set. <sup>b</sup> The residues in the active site correspond to residues 80–85, 112–124, and 134–143. <sup>c</sup> The residues in the active site correspond to residues 80–85 and 112–124. <sup>d</sup> The residues in the regular secondary structures are 13–19 ( $\beta$ 1), 48–53 ( $\beta$ 2), 59–65 ( $\beta$ 3), 82–85 ( $\beta$ 4), 94–99 ( $\beta$ 5), 27–43 ( $\alpha$ 1), 112–124 ( $\alpha$ 2), and 150–160 ( $\alpha$ 3). <sup>e</sup> Only heavy atoms from the CGS-27023A structure were used for the rmsd calculation. <sup>f</sup> Set number defined as shown in Table 1.

Table 3: Atomic rms Differences for the Final 30 Simulated Annealing Structures over the Mean Structure Obtained by Averaging the Coordinates of the Individual SA Structures from the Complete Set of NOE Restraints

		atomic rms difference (Å) <sup>a</sup>							
		active site residues <sup>b</sup>		active site residues <sup>c</sup>		residues 7–137 and 145–163		secondary structure	
		backbone	all	backbone	all	backbone	all	backbone	all
set 1 <sup>d</sup>		4.87 ± 0.97	5.89 ± 1.03	2.52 ± 1.12	3.50 ± 1.27	5.87 ± 0.42	6.73 ± 0.45	3.95 ± 0.76	4.64 ± 0.85
set 2		4.54 ± 0.68	5.62 ± 0.67	2.34 ± 0.90	3.38 ± 1.22	5.53 ± 0.41	6.47 ± 0.43	3.61 ± 0.58	4.31 ± 0.60
set 3		3.56 ± 0.72	4.55 ± 0.68	1.55 ± 0.38	2.43 ± 0.36	4.01 ± 0.31	4.99 ± 0.32	2.11 ± 0.32	2.87 ± 0.30
set 4		2.61 ± 0.76	3.75 ± 0.77	1.06 ± 0.34	2.00 ± 0.41	3.77 ± 0.41	4.76 ± 0.43	1.27 ± 0.20	2.27 ± 0.22
set 5		2.16 ± 0.63	2.99 ± 0.47	0.77 ± 0.24	1.61 ± 0.33	3.07 ± 0.35	3.87 ± 0.37	1.07 ± 0.16	1.97 ± 0.19
set 6		1.80 ± 0.54	2.63 ± 0.50	0.67 ± 0.15	1.42 ± 0.21	2.48 ± 0.19	3.13 ± 0.21	0.81 ± 0.12	1.48 ± 0.14
set 7		1.70 ± 0.36	2.53 ± 0.49	0.68 ± 0.21	1.31 ± 0.24	2.62 ± 0.23	3.27 ± 0.24	0.80 ± 0.15	1.46 ± 0.14
set 8		1.83 ± 0.43	2.57 ± 0.38	0.66 ± 0.20	1.21 ± 0.19	2.52 ± 0.25	3.19 ± 0.25	0.83 ± 0.12	1.52 ± 0.18

<sup>a</sup> Defined as the rms difference between the final 30 simulated annealing structures and the mean structure calculated from the complete set of NOE constraints. <sup>b</sup> The residues in the active site correspond to residues 80–85, 112–124, and 134–143. <sup>c</sup> The residues in the active site correspond to residues 80–85 and 112–124. <sup>d</sup> Set number defined as shown Table 1.

Among these, there are 100 sequential, 32 medium-range, and 10 long-range NOEs mostly from secondary structure elements. However, this did not improve the structure dramatically, mostly due to the lack of tertiary long-range NOE distance restraints (Table 2, set 2). It is important to note that in deuterated protein samples it is not uncommon to observe NH–NH NOEs between NH protons that are more than 5 Å apart. These NOEs were not part of the structural restraints used for the MMP-1 structure calculation since the protein was not deuterated. Also, no attempt was made to simulate potential long-range NH–NH NOEs from the high-resolution MMP-1 structure on the basis of the concern that these NOEs were unsubstantiated and may bias the analysis.

Through the use of deuterated, selectively methyl-protonated samples, NH–methyl and methyl–methyl distance restraints may be easily obtained (8, 45). With the addition of 205 NH–methyl and 87 methyl–methyl NOE restraints to the structural restraints already defined in set 2, the accuracy and precision of the structures improved dramatically (sets 3 and 4 in Table 2). Structural restraint set 3 contains only the addition of NH–methyl NOEs, whereas structural restraint set 4 contains both the NH–methyl and methyl–methyl NOEs. Comparison of the rms distribution for the 30 simulated annealing structures calculated using structural restraint sets 2 and 4 about the mean coordinate positions for residues 7–137 and 145–163 dropped from

4.03 to 2.11 Å for the backbone atoms. For the secondary structure elements, the rmsd dropped from 2.75 to 0.83 Å for the backbone atoms. An indication of the accuracy of the MMP-1 structure calculated using structural restraint set 4 is obtained by comparison with the restrained minimized average structure obtained using the complete set of structural restraints (set 9), where a backbone rmsd difference of 1.27 ± 0.20 Å is obtained for the secondary structure elements (Table 3, set 4). This is a dramatic improvement compared to the structures determined with sets 1 and 2 where the backbone rmsds are 3.95 ± 0.76 and 3.61 ± 0.58 Å for the secondary structure elements, respectively. The importance of methyl NOEs in generating global folds of proteins has been previously described (46). Additionally, the significant impact of methyl NOEs on NMR structures with minimal restraints was also previously observed (8).

It is also interesting to compare the structures determined with structural restraint sets 3 and 4, where set 3 contains only NH–methyl NOEs. Again, the structure based on structural restraint set 3 is vastly improved relative to the structures based on sets 1 and 2 (Tables 2 and 3). While there is clearly an improvement in the MMP-1 structure between sets 3 and 4, the effect is more of a gradual change. This suggests a potential transition in the impact of tertiary long-range distance restraints on the precision and accuracy of the protein structure that occurs with a minimal number of restraints. Set 4 contains 20% more distance restraints

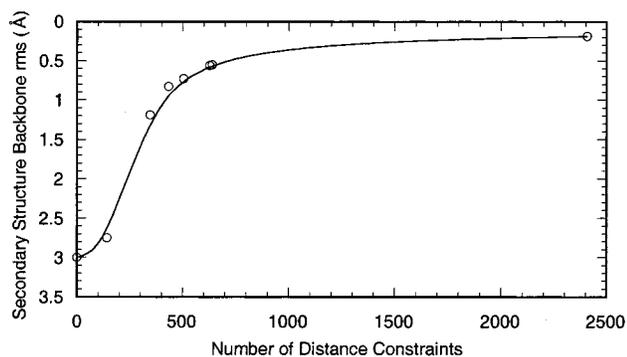


FIGURE 2: Plot of the backbone rmsd (Å) for secondary structure elements vs the number of NOE distance constraints.

(435 compared to 347), but the overall backbone rmsd for residues 7–137 and 145–163 only improves from  $2.29 \pm 0.18$  to  $2.11 \pm 0.17$  Å. The structures for MMP-1 calculated using either set 1 or set 2 contain a minimal number of tertiary long-range restraints (0–10), implying that the packing of the protein is primarily dictated by the imposed restraints from the defined secondary structure regions. The dramatic improvement between structures based on restraint sets 1 and 2 and those based on restraint sets 3 and 4 suggests that a minimal number of restraints are required to establish the general packing of the protein, an effective all-or-none transition. Thereafter, the impact of each additional restraint set may approximate a linear improvement in the quality of the structure based on the percentage change in the total number of restraints; i.e., the structures asymptotically approach the structure based on the complete set of restraints (Figure 2). A secondary smaller transition in the impact of the long-range distance restraints may occur when stereospecific  $\beta$ -methylene and particularly Leu and Val methyl assignments are incorporated into the structure refinement restraint set (47).

From a structural perspective, aromatic residues are also excellent candidates for protonation as well, since these amino acids are frequently important components of the hydrophobic core and better resolved due to ring current shifts. Further MMP-1 structure calculations included 72 NH–aromatic (set 5), 123 methyl–aromatic (set 6), and 11 aromatic–aromatic (set 7) NOEs in combination with the previous restraints. The quality of the structures has continued to improve, especially in the secondary structure regions. Particularly, the backbone rmsd for the ensemble of structures relative to the average has dropped to  $0.55 \pm 0.11$  Å for secondary structure elements. More importantly, the overall fold of the structure is quite similar to that of MMP-1 calculated with the complete set of restraints (set 9, Tables 2 and 3) where the backbone rmsd between the structures is  $0.80 \pm 0.15$  Å. As described previously, the MMP-1 structure appears to gradually improve as a function of the total number of restraints independent of the source of the restraints (methyl vs aromatic). We have chosen to include the NH–methyl and methyl–methyl constraints before NH–aromatic constraints in this study mainly because NH–methyl constraints are readily obtained compared to the aromatic constraints.

In addition to the observed improvement in the backbone rmsd as a function of the increased amount of structural information, similar trends can be seen in the standard

structural statistics (Table 1). Of particular note is the improvement in the Lennard-Jones–van der Waals energy ( $F_{L-J}$ ) from  $-384.9 \pm 25.9$  to  $-621.4 \pm 11.1$  kcal mol<sup>-1</sup> from set 1 to 9, respectively. This is consistent with an improvement in the overall fold and packing of the MMP-1 structure. Similar results are seen in the energies for other target functions (Table 1). Another point of interest is the improvement in the Ramachandran plot where the percentage of residues in the most favored region increases from 64.7 to 87.5%. This increase occurs despite the fact that the conformational database target function is used throughout the structure calculations, implying that while the use of the conformational database target function is beneficial it does not bias the resulting protein structure. Comparison of the structural statistics for MMP-1 calculated with the complete set of restraints with the values for the structure calculated with set 7 indicates that the structure with the complete set of restraints is of higher quality. This is not an unexpected result, but the data also suggest that the structure based on restraint set 7 is reasonably comparable to the high-quality MMP-1 structure. The rmsd data and structural statistics clearly indicate that a reasonably accurate overall structure for MMP-1 can be calculated on the basis of a minimal set of restraints that may be obtained from a deuterated protein. These results are consistent with the previous observation of structures based on minimal NOE-based structural information (8). Nevertheless, the real test of the utility of the MMP-1 structures will be the reproducibility of the active site conformation obtained from the minimal restraint sets relative to the high-quality MMP-1 structure. More importantly will be how the structures accurately predict the binding of MMP-1 inhibitors.

*Active Site of MMP-1 and the S1' Pocket.* The MMPs have a well-defined active site centered about a catalytic Zn chelated by three histidine residues. Previous structural analysis of the MMP family of enzymes has identified distinct substrate binding sites on both the left and right sides of the catalytic Zn (40). Most efforts in the design of inhibitors to MMP-1 have focused on the right side binding site with particular interest in the S1' pocket. The shape and size of the S1' pocket vary significantly between the various MMP enzymes and are a potential source of designing selectivity for a given MMP inhibitor. The current paradigm in the development of MMP inhibitors is to design specificity based on the postulation that a broad-spectrum MMP inhibitor would provide a higher level of exposure to toxic side effects. Therefore, an important aspect for the utility of the MMP-1 structure based on minimal structural information is the accuracy of the conformation of the active site and particularly the S1' pocket.

A comparison of the rms difference for residues comprising the MMP-1 active site and the relative size and shape of the S1' pocket among the various restraint data sets will indicate the relative accuracy of the MMP-1 active site. The size and shape of the S1' pocket for the MMP-1 structures determined from the various structural information data sets were calculated from Connolly solvent accessible surfaces. The S1' pocket for the high-quality MMP-1 structure calculated from the complete set of structural information is illustrated in Figure 3C. A visible comparison of the S1' pocket for MMP-1 calculated from the complete set of restraints and the minimal restraint set 1 clearly indicates a

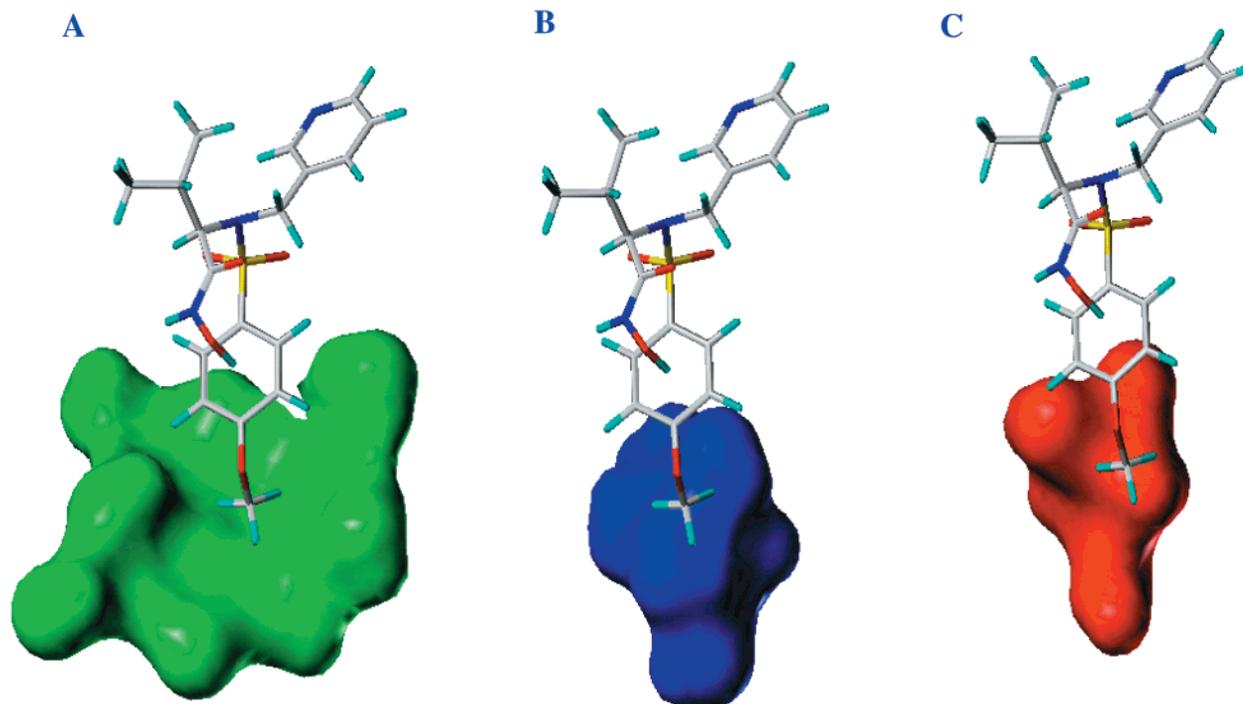


FIGURE 3: S1' pocket of MMP-1 calculated from a minimal set of restraints (set 1, A), a maximum set of restraints (set 7, B) for a deuterated protein, and a complete set (set 9, C) of restraints for MMP-1 with CGS-27023A docked in for comparison.

poor reproduction of the S1' pocket. The S1' pocket for the MMP-1 structure based on minimal restraint set 1 is significantly broader than that in the high-quality MMP-1 structure (see Figure 3A, set 1). It is readily apparent that from a drug design perspective the MMP-1 structure based on data set 1 would be misleading since it would suggest the availability of space for additional modification of the ligand that is not apparently available in the high-quality structure. The situation improves drastically with the incorporation of minimal tertiary long-range distance restraints (Figure 3B, set 7). Comparison of the S1' pocket between the high-quality MMP-1 structure and the MMP-1 structure based on minimal structural information, including a minimal number of tertiary long-range distance restraints, indicates that the S1' pockets are nearly identical. Clearly, from a drug design perspective the MMP-1 structure based on structural restraint set 7 would be very valuable, especially in light of the importance of the S1' pocket in the design of MMP inhibitors. The MMP-1 structure based on structural restraint set 7 would accurately predict appropriate modifications to bound ligands for optimization of utilization of the available space in the S1' pocket comparable to the high-quality MMP-1 structure. Obtaining S1' pockets with similar sizes and shapes for MMP-1 with both a complete and a minimal restraint data set is very encouraging for the utility of NMR structures based on minimal structural information.

Further analysis of the MMP-1 active site residues suggests additional support for the utility of structures based on minimal restraints. Table 2 shows the rmsd for the MMP-1 active site residues (80–85, 112–124, and 134–143) of the 30 simulated annealing structures about the mean structure determined for each ensemble calculated for each data set. Similarly, Table 3 shows the rmsd for the MMP-1 active site residues of the 30 simulated annealing structures calculated for each data set about the mean structure of MMP-1 determined from the complete set of NOE calcula-

tions. Like the results observed for the overall MMP-1 structure, the accuracy and precision in the active site improve dramatically with the increase in the number of NOE distance restraints. The total number of NOEs per residue in the active site increases from 0.82 per residue in set 2 to 5.21 in set 7, with a total of 90 dipolar coupling constants used to define the active site. Again, the largest change occurs with the first incorporation of tertiary long-range distance restraints. The rmsd for the active site residues is  $3.67 \pm 0.69$  Å for set 2 where only 23 NOEs were used to define the active site. A total of 146 NOEs are used to define the active site residues in set 7 with the incorporation of methyl and aromatic restraints. However, even with the additional methyl and aromatic restraints (set 7), the accuracy in the active site is relatively low compared to that in the secondary structure regions,  $1.70 \pm 0.36$  Å compared to  $0.80 \pm 0.15$  Å for backbone atoms (Table 3).

Dynamic analysis of MMP-1 has previously demonstrated that residues 134–143 are highly flexible, as is evident because of the low order parameters ( $S^2 < 0.6$ ) even in the presence of an inhibitor (26). This observation suggests that while residues 134–143 compose part of the MMP-1 active site it may contribute minimal information for inhibitor binding and design. Clearly, in the case of the MMP-1 structures calculated on the basis of minimal structural information, residues 134–143 are poorly defined because of this flexibility. To some extent, the poorly defined structure for residues 134–143 may be beneficial information for drug design, since these residues would probably not be a focal point to drive modification of the ligand. Conversely, targeting these residues on the basis of an apparently better defined structure could potentially be wasted effort because of the inherent mobility of these residues. Excluding residues 134–143 from the active site definition results in a significant drop in the rmsd from  $1.70 \pm 0.36$  to  $0.68 \pm 0.21$  Å for the backbone atoms (set 7, Table 3). This actually results

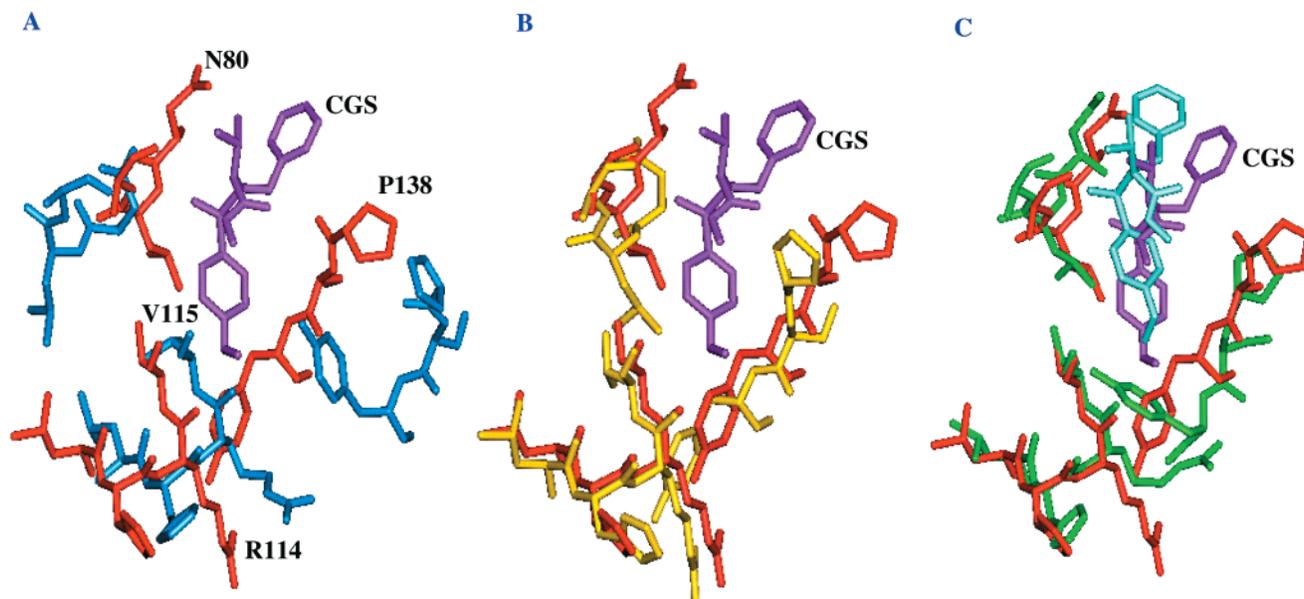


FIGURE 4: Best fit superposition of the backbone of the MMP-1 structure for residues in the active site together with CGS-27023A. The active residues that are displayed are 80–82, 112–115, and 138–140 which play a critical role in MMP-1 activity. (A) Comparison of the active site for structure sets 1 (cyan) and 9 (red) with CGS-27023A (magenta) docked in. (B) Comparison of the active site for structure sets 7 (yellow) and 9 (red). (C) Comparison of the active site for structure set 9 (red) and the MMP-1–CGS-27023A complex calculated with a complete set of constraints (green).

in an improvement in the active site region compared to the secondary structure region. Figure 4 shows the best fit superposition of the active site residues for the high-quality MMP-1 structures calculated with the complete set of structural restraints with structures based on minimal structural restraints (Figure 4A, set 1, and Figure 4B, set 7). The results are similar to those from the comparison of the size and shape of the S1' pocket. While there is some similarity in the positioning of the active site residues between the high-quality MMP-1 structure and minimal data set 1, the resemblance is slight, as is apparent by the structural overlap (Figure 4A) and large rms difference in the backbone atoms ( $2.34 \pm 0.90$  Å). Again, a significant improvement is observed when the high-quality MMP-1 structure is compared with the MMP-1 structure based on minimal structural information, including a minimal number of tertiary long-range distance restraints (Figure 4B, set 7). The reasonable degree of similarity and relatively low backbone rmsd for the active site residues between the high-quality MMP-1 structure and the MMP-1 structure based on data set 7 again suggest a valuable structure for drug design.

**Impact on the Drug Design Process.** The fundamental information that is utilized in a structure-based drug design approach is the relative location of potential positive interactions between the protein and small molecule. Essentially, the structure of the protein–ligand complex identifies regions of the protein that are proximal to the ligand and may interact with the small molecule with the addition of an appropriate substituent. Additionally, the structure may also suggest possible modifications to the ligand where existing interactions to the protein are suboptimal. Even with the available structural information, improving the affinity of a ligand on the basis of the predicted modifications does not always yield the desired result because of the inherent mobility of the protein and how it accommodates the modified ligand. As a result, the drug design process is an iterative approach where a new structure is determined for each modified ligand bound

to the protein (21). The impact on the drug design process of a lower-resolution structure is the higher variability in the precise location of a desirable interaction. Accordingly, a larger set of modified ligands is required to effectively explore the potential available space in the protein's active site. Clearly, there is a point where the low quality of the protein structure results in such a poorly defined active site as to render the structure virtually useless in the further design of ligands. The resolution at which a low-quality structure becomes unproductive in the drug design process is difficult to definitively identify, but the utility of a low-resolution structure is primarily based on maintaining the fundamental characteristics of the protein's active site. Comparison of the MMP-1 active sites that were determined with data sets 2 and 7 and the complete restraint set illustrates this point (Figures 3 and 4).

The MMP-1 structure based on data set 7 still maintains the fundamental characteristics of the protein's active site relative to the high-resolution structure, which is not the case for the MMP-1 structure calculated with data set 2. Therefore, comparison of the backbone rmsd for the ensemble of MMP-1 structures calculated with data sets 2 and 7 provides some insight into general limits defining the utility of the structure based on minimal restraints (Table 2). Additionally, comparison of the backbone rmsd for the MMP-1 active site residues for the structure based on minimal restraint set 7 ( $0.54 \pm 0.15$  Å) with the structure calculated with the full set of constraints ( $0.18 \pm 0.04$  Å) provides some indication of the precision of the atom position in each structure. Similarly, the accuracy of the atom positions may be ascertained from the backbone rms difference between the high-resolution structure and the structure based on minimal restraint set 7 ( $0.68 \pm 0.21$  Å). The difference in accuracy and precision between the two MMP-1 structures may influence the range of potential modifications to the ligand, but the structure based on minimal restraints still identifies the key interactions with a bound ligand and the available

space to explore in the active site. Fundamentally, the impact of the lower-resolution structure on the drug design process would be to increase the number of compounds needed to test any potential improvement in ligand affinity. Nevertheless, the information from a lower-resolution structure would be invaluable in the absence of a high-resolution structure. In fact, the lower-resolution structure may also have some advantages over a high-resolution structure.

An inevitable downside of using a high-resolution structure for drug design is viewing the structure as a static image where the exact position of individual atoms is perceived to be precisely known. The "softer" definition of the active site for a lower-resolution structure may maintain potential beneficial modifications to the ligand that would be obviously eliminated due to poor contact in a high-resolution structure. Essentially, the lower-resolution structure may incorporate, to some extent, the inherent mobility and flexibility of a protein structure in the drug design process. Recent NMR and X-ray structures of MMP-1 and MMP-13 complexed with distinct inhibitors have demonstrated the potential for side chains in the active site to undergo conformational changes to accommodate bound inhibitors (23, 48, 49). These changes in side chain conformation could not readily be predicted from previous structures and have a significant impact on ligand binding.

A potential caveat for the utility of the MMP-1 structure based on the minimal restraint set is its use in a de novo approach to drug design. Given the inherent difficulties in de novo drug design, the added complication of the lower-resolution structure would probably result in a low success rate (50). This is simply based on the fact that the presence of a bound ligand in the protein's active site provides an abundance of information that is beneficial to the drug design process.

The importance of a bound ligand is further exemplified when the structure of MMP-1 complexed with CGS-27023A was determined from minimal structural information (set 8). The MMP-1–CGS-27023A complex was previously determined by NMR (23). CGS-27023A simply docked into the MMP-1 active site pocket based on the superposition of the MMP-1 backbone atoms is illustrated in Figure 4A,B. The docking of CGS-27023A into the MMP-1 structure based on data set 1 indicates the low level of reproduction of interactions between MMP-1 and CGS-27023A. A significant improvement is observed for CGS-27023A docked into the MMP-1 structure based on data set 7, but some errors in the structure are still apparent (Figure 4B) where the discrepancies in the active site are mostly from residues 138–140. When the MMP-1 structure is calculated in the presence of CGS-27023A utilizing 18 additional intramolecular NOEs for CGS-27023A and 25 intermolecular NOEs between CGS-27023A and MMP-1, these errors are removed and the resulting structure is strikingly similar to the previous structure of the MMP-1–CGS-27023A complex (Figure 4C). In particular, the experimental interactions observed by NMR between MMP-1 and CGS-27023A are reproduced and the binding of CGS-27023A in MMP-1 is comparable to that in the MMP-1–CGS-27023A complex determined from a complete set of structural information. It is important to note that the additional NOEs used in the calculation of the MMP-1–CGS-27023A complex using minimal restraint set 8 are consistent with NOEs that would be observable in a

deuterated protein sample. In addition to the CGS-27023A intramolecular NOEs, which would all be observable in a deuterated sample, only the CGS-27023A–NH, CGS-27023A–methyl, and CGS-27023A–aromatic NOEs were used in the structure calculation. These results indicate that the added restraint information that can be obtained from a protein–ligand complex provides a significant improvement in determining the structure of the active site. Additionally, these results suggest that the MMP-1–CGS-27023A complex determined from a minimal set of structural restraints would be an effective tool for a structure-based approach to modifying CGS-27023A and other ligands to improve inhibitor affinity and selectivity. Furthermore, the results described herein for MMP-1 strongly suggest that structures calculated for high-molecular weight proteins using deuterium labeling and minimal structural restraints similar to the ones utilized for the MMP-1 calculations would generate valuable structures of reasonable accuracy for drug design and other biological applications. In addition, the results with MMP-1 also suggest that utilizing minimal NOE-based structural information to rapidly determine protein structures compared to the large time commitment required to completely analyze all the NOE data may be a viable approach to expediting the drug discovery process.

## REFERENCES

1. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
2. Clore, G. M., and Gronenborn, A. M. (1998) *Trends Biotechnol.* 16, 22–34.
3. Clore, G. M., and Gronenborn, A. M. (1991) *Prog. Nucl. Magn. Reson. Spectrosc.* 23, 43–92.
4. Gardner, K. H., and Kay, L. E. (1998) *Annu. Rev. Biophys. Biomol. Struct.* 27, 357–406.
5. Shan, X., Gardner, K. H., Muhandiram, D. R., Rao, N. S., Arrowsmith, C. H., and Kay, L. E. (1996) *J. Am. Chem. Soc.* 118, 6570–6579.
6. Shan, X., Gardner, K. H., Muhandiram, D. R., Kay, L. E., and Arrowsmith, C. H. (1998) *J. Biomol. NMR* 11, 307–318.
7. Clore, G. M., and Gronenborn, A. M. (1991) *Science* 252, 1390–1399.
8. Gardner, K. H., Rosen, M. K., and Kay, L. E. (1997) *Biochemistry* 36, 1389–1401.
9. Tjandra, N., and Bax, A. (1997) *Science* 278, 1111–1114.
10. Cai, M., Huang, Y., Zheng, R., Wei, S. Q., Ghirlando, R., Lee, M. S., Craigie, R., Gronenborn, A. M., and Clore, G. M. (1998) *Nat. Struct. Biol.* 5, 903–909.
11. Bewley, C. A., Gustafson, K. R., Boyd, M. R., Covell, D. G., Bax, A., Clore, G. M., and Gronenborn, A. M. (1998) *Nat. Struct. Biol.* 5, 571–578.
12. Clore, G. M., Starich, M. R., Bewley, C. A., Cai, M., and Kuszewski, J. (1999) *J. Am. Chem. Soc.* 121, 6513–6514.
13. Clore, G. M., and Gronenborn, A. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 5891–5898.
14. Kuszewski, J., Qin, J., Gronenborn, A. M., and Clore, G. M. (1995) *J. Magn. Reson., Ser. B* 106, 92–96.
15. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1995) *J. Magn. Reson., Ser. B* 107, 293–297.
16. Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M., and Clore, G. M. (1994) *J. Magn. Reson., Ser. B* 104, 99–103.
17. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) *J. Magn. Reson., Ser. B* 112, 79–81.
18. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1997) *J. Magn. Reson.* 125, 171–177.
19. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1999) *J. Am. Chem. Soc.* 121, 2337–2338.
20. Gubernator, K., and Boehm, H. J. (1998) *Methods Princ. Med. Chem.* 6, 15–36.

21. Kubinyi, H. (1998) *Curr. Opin. Drug Discovery Dev.* 1, 4–15.
22. Moy, F. J., Chanda, P. K., Cosmi, S., Pisano, M. R., Urbano, C., Wilhelm, J., and Powers, R. (1998) *Biochemistry* 37, 1495–1504.
23. Moy, F. J., Chanda, P. K., Chen, J. M., Cosmi, S., Edris, W., Skotnicki, J. S., Wilhelm, J., and Powers, R. (1999) *Biochemistry* 38, 7085–7096.
24. Woessner, J. F., Jr. (1991) *FASEB J.* 5, 2145–2154.
25. Ries, C., and Petrides, E. (1995) *Biol. Chem. Hoppe-Seyler* 376, 345–355.
26. Moy, F. J., Pisano, M. R., Chanda, P. K., Urbano, C., Killar, L. M., Sung, M. L., and Powers, R. (1997) *J. Biomol. NMR* 10, 9–19.
27. Ottiger, M., and Bax, A. (1998) *J. Biomol. NMR* 12, 361–372.
28. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson.* 85, 393–399.
29. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* 131, 373–378.
30. Wans, Y.-X., Marquardt, J. L., Wingfield, P., Stahl, S. J., Lee-Huang, S., Torchia, D., and Bax, A. (1998) *J. Am. Chem. Soc.* 120, 7385–7386.
31. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson.* 96, 432–440.
32. Cai, M., Wang, H., Olejniczak, E. T., Meadows, R. P., Gunasekera, A. H., Xu, N., and Fesik, S. W. (1999) *J. Magn. Reson.* 139, 451–453.
33. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson.* 133, 216–221.
34. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* 6, 277–293.
35. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson.* 95, 214–220.
36. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) *Protein Eng.* 2, 27–38.
37. Brünger, A. T. (1993) *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
38. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) *J. Magn. Reson.* 131, 159–162.
39. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* 13, 289–302.
40. Zask, A., Levin, J. I., Killar, L. M., and Skotnicki, J. S. (1996) *Curr. Pharm. Des.* 2, 624–661.
41. Morphy, J. R., Millican, T. A., and Porter, J. R. (1995) *Curr. Med. Chem.* 2, 743–762.
42. Browner, M. F. (1995) *Perspect. Drug Discovery Des.* 2, 343–351.
43. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* 8, 477–486.
44. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* 4, 171–180.
45. Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., and Kay, L. E. (1996) *J. Mol. Biol.* 263, 627–636.
46. Metzler, W. J., Wittekind, M., Goldfarb, V., Mueller, L., and Farmer, B. T., II (1996) *J. Am. Chem. Soc.* 118, 6800–6801.
47. Clore, G. M., Bax, A., and Gronenborn, A. M. (1991) *J. Biomol. NMR* 1, 13–22.
48. Lovejoy, B., Welch, A. R., Carr, S., Luong, C., Broka, C., Hendricks, R. T., Campbell, J. A., Walker, K. A., Martin, R., Van Wart, H., and Browner, M. F. (1999) *Nat. Struct. Biol.* 6, 217–221.
49. Moy, F. J., Chanda, P. K., Chen, J. M., Cosmi, S., Edris, W., Levin, J. I., Wilhelm, J., and Powers, R. (2000) *J. Mol. Biol.* 302 (3), 673–691.
50. Waszkowycz, B. (1998) in *Advances in Drug Discovery Techniques* (Harvey, A. L., Ed.) pp 143–164, Wiley, Chichester, U.K.

BI001658S