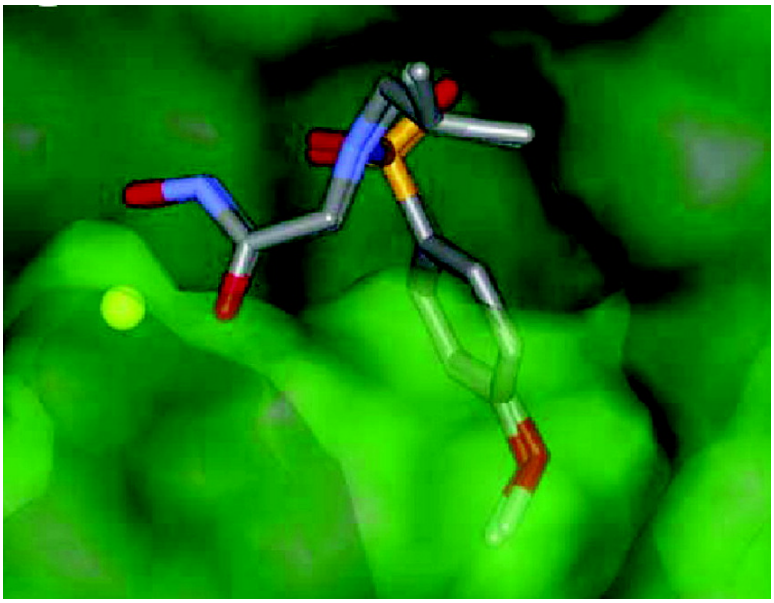
Journal of Medicinal Chemistry

Article

Combining in Silico Tools and NMR Data To Validate Protein–Ligand Structural Models: Application to Matrix Metalloproteinases

Ivano Bertini, Marco Fragai, Andrea Giachetti, Claudio Luchinat, Massimiliano Maletta, Giacomo Parigi, and Kwon Joo Yeo

J. Med. Chem., **2005**, 48 (24), 7544-7559• DOI: 10.1021/jm050574k • Publication Date (Web): 05 November 2005



Combining in Silico Tools and NMR Data To Validate Protein−Ligand Structural Models: Application to Matrix Metalloproteinases

Ivano Bertini,*[,†,‡] Marco Fragai,[†,§] Andrea Giachetti,[†,⊥] Claudio Luchinat,[†,§] Massimiliano Maletta,[†,‡] Giacomo Parigi,[†,§] and Kwon Joo Yeo[†]

Magnetic Resonance Center, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy, Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy, Department of Agricultural Biotechnology, University of Florence, Via Donizetti 6, 50144 Florence, Italy, and ProtEra, Viale delle Idee 22, 50019 Sesto Fiorentino, Italy

Received June 17, 2005

A combination of in silico tools and experimental NMR data is proposed for relatively fast determination of protein−ligand structural models and demonstrated from known inhibitors of matrix metalloproteinases (MMP). The ^{15}N ^{1}H heteronuclear single quantum coherence (HSQC) spectral assignment and the 3D structure, either X-ray or NMR, are needed. In this method, the HSQC spectrum with or without the ligand is used to determine the interaction region of the ligand. Docking calculations are then performed to obtain a set of structural models. From the latter, the nuclear Overhauser effects (NOEs) between the ligand and the protein can be predicted. Guided by these predictions, a number of NOEs can be detected and assigned through a HSQC NOESY experiment. These data are used as structural restraints to reject/refine the initial structural models through further in silico work. For a test protein (MMP-12, human macrophage metalloelastase), a final structure of a protein−ligand adduct was obtained that matches well with the full structural determination. A number of structural predictions were then made for adducts of a similar protein (MMP-1, human fibroblast collagenase) with the same and different ligands. The quality of the final results depended on the type and number of experimental NOEs, but in all cases, a well-defined ligand conformation in the protein binding site was obtained. This protocol is proposed as a viable alternative to the many approaches described in the literature.

Introduction

Rational drug design strategies must rely on the availability of high-throughput methods to experimentally determine the structure of candidate drug-target complexes.[1] The obtained structural information is then used to improve and optimize the candidate drug in a cyclic procedure. Obtaining three-dimensional macromolecular structures is still a time-consuming task. X-ray structure determination is becoming a high-throughput method,[2] but the method requires the easy availability of protein crystals that are suitable for soaking with the various candidate drugs. NMR is also a high-throughput technique in drug discovery,[3,4] but its power lies mostly in the earlier phases of the process, i.e., in the first screening of a relatively large number of compounds. NMR quickly provides information on binding affinity and on the region of interaction of the candidate drug with the target molecule.[5]

NMR is of course also able to determine the three-dimensional structure of the adduct, but the procedure is time-consuming.[6] Moreover, obtaining a 3D structure depends on the full assignment of thousands of intraprotein nuclear Overhauser effect spectroscopy (NOESY) cross-peaks, while the only relevant ones are the few intermolecular cross-peaks between protein and ligand signals. In silico prediction of the structure of the adduct through docking programs, while valuable in the early ligand design phases, is not reliable at this stage.[7−9] Independently of the docking program used, in many cases more than one binding poses are found that do not significantly differ in predicted binding energies.

The availability of a fast and reliable method able to provide a molecular model based on few experimental restraints is an ambitious goal for overcoming these problems. Recently, several efforts have been performed in this direction.[10−13] For instance, a suite of NMR experiments has been recently proposed as a tool to provide structural information on protein−ligand adducts,[12] through intermolecular NOEs detected in selectively labeled proteins. The method is applicable to very large proteins once their three-dimensional structure is known.

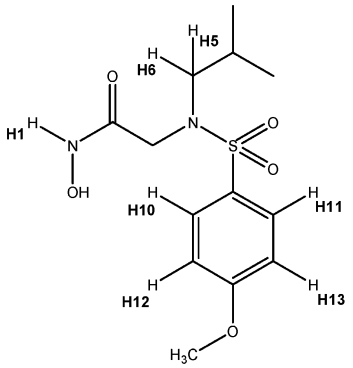
For smaller proteins, it is worth to investigate whether a few NOEs may be obtained even without selective labeling of the proteins. We propose here a combined use of computational tools and a small number of experimental NMR restraints as an efficient way of selecting the correct binding pose among those proposed by docking programs. The experimental restraints are (i) the heteronuclear single quantum coherence (HSQC) chemical shifts to select the region of interest on the

* To whom correspondence should be addressed. Tel: +39-055-4574270. Fax: +39-055-4574271. E-mail: bertini@cerm.unifi.it.
† Magnetic Resonance Center, University of Florence.
‡ Department of Chemistry, University of Florence.
§ Department of Agricultural Biotechnology, University of Florence.
⊥ ProtEra.

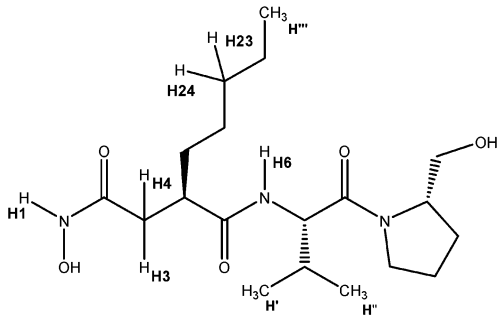
10.1021/jm050574k CCC: $30.25
Published on Web 11/05/2005

Chart 1

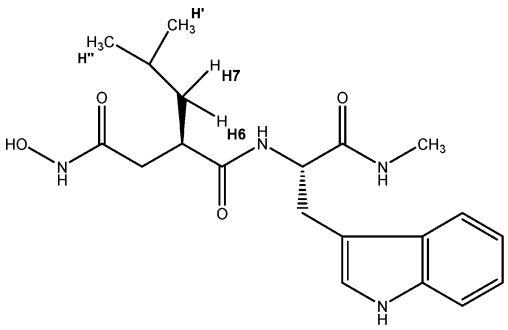
NNGH



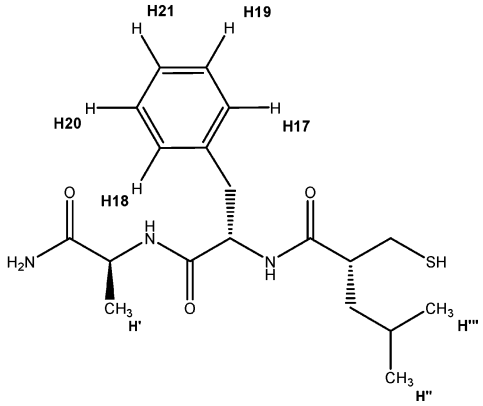
Actinonin



Galardin



SIMP-1



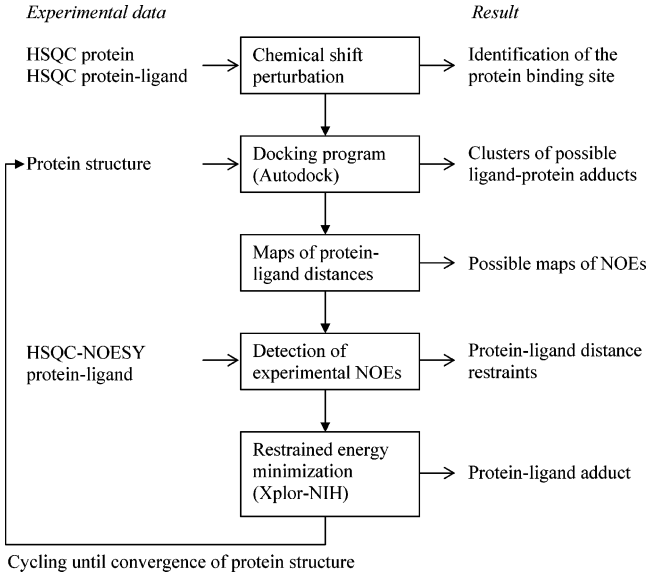
target, and (ii) the few ligand-target NOEs that can be unambiguously identified from ^{15}N NOESY−HSQC experiments. Besides the protein three-dimensional structure, only a singly ^{15}N-labeled protein sample and a preexisting assignment of its ^{15}N ^{1}H HSQC spectrum are required.

The method has been validated by reproducing the known docked conformation of *N*-isobutyl-*N*-[4-methoxyphenylsulfonyl]glycyl hydroxamic acid (NNGH, see Chart 1) bound to matrix metalloproteinase 12 (MMP-12, human fibroblast metalloelastase). The method has been then applied to obtain the docked conformations of NNGH and other three ligands, (3-[[1-[[2-(hydroxymethyl)-1-pyrrolidinyl]carbonyl]-2-methylpropyl]carbamoyl]-octanohydroxamic acid (actinonin), *N*-[(2*R*)-2-(hydroxamidocarbonylmethyl)-4-methylpentanoyl]-L-tryptophan methylamide (galardin), and (2*R*)-2-mercaptomethyl-4-methylpentanoyl-L-phenylalanyl-L-alanine amide (SIMP-1) (see Chart 1) to MMP-1 (human fibroblast collagenase). MMPs belong to a family of zinc-dependent endopeptidases responsible for the metabolism of extracellular matrix proteins,[14−16] and alterations in their levels are implicated in a wide range of pathological states,[17,18] so that these proteins represent attractive drug targets.

Methods

The protocol consists of the following steps, reported in Scheme 1: (a) identification of the protein binding site, (b) calculation of possible protein−ligand adducts, (c) prediction of the map of NOEs corresponding to each computed conformation, (d) determination of few experimental restraints, able to select the real adduct among those calculated, and (e) validation and cyclic in silico refinement of the ligand position in the protein scaffold. The identification of the protein binding site can be conveniently performed from the analysis of the chemical shifts acquired in the presence and in the absence of the ligand. NOEs between ligand protons and protein protons are obtained from ^{15}N NOESY−HSQC

Scheme 1



spectra. The protocol requires that the protein structure and the assignment of its ^{15}N ^{1}H HSQC spectrum is known.

HSQC spectra of the protein in the presence and in the absence of the ligand must be acquired. Most of the protein peaks will coincide in the two spectra. Only peaks corresponding to amide protein protons close to the ligand will be in different positions, but their shift is usually small enough to be easily assigned. This information is used to identify the protein binding site, according to the value of the combined $^{1}H/^{15}N$ shift perturbation upon complexation, given by $\Delta = (\Delta\delta(^{1}H)^2 + (\Delta\delta(^{15}N)/6)^2)^{1/2}$.[19] The residues with a significantly large value of Δ, except those at sizably larger distance from all others, are used to identify the grid for docking calculations. The latter is centered on the protein surface atom closest to the center of the smallest sphere that comprises all the selected nitrogen atoms.

Due to the complexity of the energy landscape on the path to the global minimum region,[20] a specific ligand−protein docking program is invoked in order to accurately probe and select the conformations of the ligand according to appropriate scoring functions. We use the program Autodock because it has been amply validated and tested on the target proteins selected for this study. The docking program can be run to obtain clusters of the possible adducts. Such clusters are then used to predict NOEs between protein and ligand nuclei. In fact, a map of distances between ligand and protein nuclei can be obtained for each of the different clusters. The presence of cross-peaks can thus be predicted for the different possible adducts and compared with the cross-peaks actually present in the experimental spectra.

The following experiments must be performed: ^{15}N NOESY−HSQC spectra of the protein−ligand adduct and of the free protein, and the 1D ^{1}H spectrum of the free ligand in water. The latter experiment provides an estimate of where the chemical shifts of ligand signals in the adduct have to be looked for. The presence of intermolecular cross-peaks, i.e., peaks between frequencies close to those of the free ligand in one dimension, and those of the protein amide protons predicted to be in the vicinity of the ligand in the other dimension, is checked. Such cross-peaks, if absent in the free protein spectrum and not attributable to nuclei of other neighboring protein residues, are unambiguously assigned. A good correspondence between expected and observed cross-peaks is a clear indication of the goodness of the corresponding cluster. On the other hand, direct evidence of the unacceptability of some of the clusters generated by Autodock can be obtained. The experimental NOEs, translated into upper distance limits, can then be used to refine the remaining acceptable structures and possibly to further discriminate among them. The refinement procedure has been developed using Xplor-NIH. In such procedure, the protein side chains are left free to move, thus allowing a better docking to be obtained with respect to docking programs where the protein is completely rigid.

The refinement procedure consists of loading the calculated adduct and performing an in vacuo molecular dynamics simulation in internal coordinates, with backbone atoms grouped together to constitute a rigid structure. A simulated annealing is performed by heating the system to 1500 K and then cooling it to 50 K in steps of 50 K. At each temperature, 750 steps of molecular dynamics simulations are performed with time steps of 2 fs. The force constant of NOE restraints is fixed to 30 kcal mol^{-1} $Å^{-2}$, and van der Waals, electrostatic terms and the protein and ligand force field (angles, bonds, dihedrals and impropers) are also included. The resulting structures are then refined with a Powell minimization, and ordered according to the value of the target function. The latter is calculated considering the ligand-residue and residue−residue interactions only for residues up to 8 Å from the ligand. This helps reducing the energy "noise" originating from slight changes in residue−residue interactions far away from the ligand site. In all cases, the best 10 structures over 200 calculated through Xplor-NIH starting from each tentative docking structure are very similar to one another.

The structure of the adduct is thus calculated through the consecutive use of the program Autodock and the refinement procedure working in Xplor-NIH. Xplor-NIH calculations can significantly change the protein side chain positions after complexation. Therefore, cycling between Autodock and Xplor-NIH refinement is necessary until convergence to a fixed protein structure is achieved. We have tested that such approach can actually select the correct ligand−protein docking, among those proposed by Autodock. Furthermore, the introduction of experimental data and the allowed mobility of the protein side chains provide more confidence in the obtained adduct.

MMP systems, the receptors that we used in this work, have a catalytic zinc ion as active center, coordinated to three histidines. The three zinc-coordinated histidines were treated as the neutral form with the hydrogen on ND1, whereas other histidines used the default option with hydrogen on NE2. Glutamates were treated as charged form as default, except the catalytically essential glutamate 219,[21] at the second shell of the zinc binding site. The latter residue was protonated, with the hydrogen on the oxygen nearest to the catalytic zinc, or deprotonated depending on whether the zinc donor atom closest to it was deprotonated (hydroxamate ligands)[21] or protonated (thiol ligands). To take into account the electron density delocalization due to coordination of ligands, the charge of the zinc ion was distributed among the protein ligands.[21]

Results

Test with a Known Structure: MMP-12−NNGH. NNGH is a broad spectrum MMP inhibitor able to interact with both the catalytic zinc and the S1′ cavity.[6,22] In particular, it is able to bind MMP-12 with nanomolar affinity (K_d = 10 nM),[6] and for this reason it has been chosen as a model system to study protein-inhibitor interactions. Its molecular structure is reported in Chart 1.

The structure of MMP-12 complexed to NNGH is already known,[6] as both the X-ray (Figure 1A) and the NMR structures of the adduct have been solved. Therefore, we used such system as a test for our protocol. The HSQC spectra of the protein without and with the NNGH in solution were acquired. Inspection of residues showing significant chemical shift perturbation (see

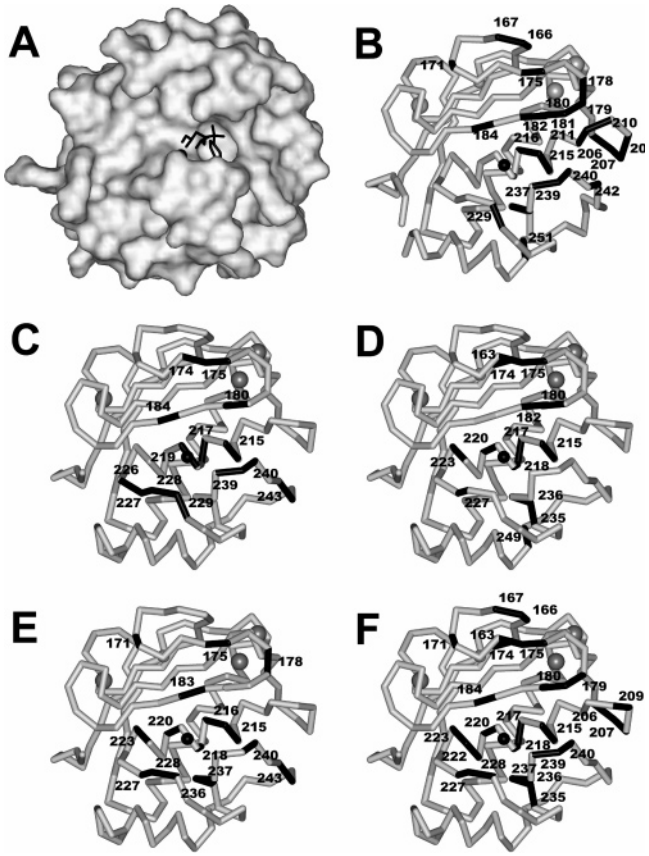


Figure 1. X-ray structure of the MMP-12−NNGH adduct (PDB 1RMZ) (A), residues of MMP-12 affected by chemical shift perturbation upon complexation with NNGH (B), and residues of MMP-1 affected by chemical shift perturbation upon complexation with NNGH (C), actinonin (D), galardin (E), and SIMP-1 (F) (see Table 2).

Table 1) permitted to define the ligand binding region as the protein catalytic site (defined here as constituted by the zinc binding region, the S1′ pocket and the substrate binding groove) with reasonable accuracy. As expected from the crystal structure of the MMP-12−NNGH adduct, among the affected resonances are residues 210, 211, 215, and 216 on the α helix at the bottom of zinc binding site, residues 237, 239−240 and 242 forming the hydrophobic S1′ cavity, and residues 179−182 and 184 on the strand facing both the catalytic metal and the S1′ pocket (Figure 1B).

Calculations were performed using the X-ray structure of the protein (PDB 1Y93) at 1.03 Å resolution.[6] Autodock was used to select the lower docking energy conformations. Docked conformations were clustered according to a maximal RMSD of 1 Å (Figure 2). The docking energies for the first, second, third and fourth clusters were −15.89, −15.84, −15.01, and −14.44 kcal mol^{-1}, respectively. The second cluster is in accordance with the X-ray structure of the adduct (PDB 1RMZ).[6] The plane containing the hydroxamic group in the first and third cluster is oriented perpendicularly to the plane containing the hydroxamic group in the second cluster. The *p*-methoxy-phenyl group enters more deeply in the S1′ pocket in the first than in the third cluster. In the fourth cluster the *p*-methoxy-phenyl group does not sit in the S1′ pocket.

These structures were separately refined with Xplor-NIH using the already available NOEs with protein backbone NH atoms (see Figure 3B).[6] The second cluster remains essentially unchanged, with total energy −1184 kcal mol^{-1} (see Figure 2). The structure calculated using the third cluster as starting conformation is similar to the previous one, with total energy −1179 kcal mol^{-1}. The structure calculated from the first cluster has total energy −1062 kcal mol^{-1}, and no coordination of the hydroxamic group to the metal ion; the one calculated from the fourth cluster has total energy −734 kcal mol^{-1}, and the *p*-methoxy-phenyl group outside the S1′ pocket.

Slight changes in the side-chain protein structure were obtained, and new Autodock calculations were thus performed using the three lowest energy Xplor-NIH protein structures. Remarkably, the lowest docking energy clusters calculated by Autodock now converge to similar conformations using the second and third Xplor-NIH protein structure (see Figure 2). These conformations are in agreement with the X-ray structure, with docking energy from −15.85 kcal mol^{-1} to −15.63 kcal mol^{-1}. Xplor-NIH refinements provided structures (see Figures 2 and 3B) with lowest total energy from −1197 to −1180 kcal mol^{-1}, in agreement with the X-ray structure (see Figure 3A).

Analogous calculations were performed also using the X-ray structure PDB 1OS9, with 1.85 Å resolution.[23] In this structure the active site of one molecule is not hosting an external ligand but the N-terminal part of the neighboring protein molecule. The calculations converged to the same adduct obtained starting from the 1Y93 structure.

Determination of Structural Models for Ligand Adducts of MMP-1. NNGH itself and three other known strong inhibitors of MMPs were selected as representatives of different classes of ligands and tested against MMP-1. The test consists in following the protocol described above and checking whether (i) unambiguous NOEs could be obtained and (ii) the cycling between Autodock and Xplor-NIH calculations permits the selection of one ligand conformation. Calculations were performed using the X-ray structure of the inhibitor-free protein (PDB 1CGE) with 1.90 Å resolution.[24]

MMP-1−NNGH. The first ligand examined is the same ligand used to validate the protocol with MMP-12. The structure of the NNGH adduct with MMP-1 is not known, although it is reasonable to believe that it will adopt a similar conformation. We measured an IC_{50} value for the adduct of 174 nM.

Chemical shift perturbation affects the zinc binding histidine 228 and the neighboring residues 226, 227, and 229, residues 239, 240, and 243 forming the S1′ hydrophobic pocket, residues 215, 217, and 219 on the α-helix where the metal binding site is inserted, and residues 180 and 184 on the parallel strand (see Figure 1C). This is an expected feature, but it is a new independent experimental information based on which an Autodock grid was generated. The grid resulted nicely centered around the known catalytic site. Autodock calculations using this grid were thus performed.

The four lower docking energy clusters were analyzed. The docking energies were −14.68, −13.99, −13.94, and −13.81 kcal mol^{-1}, respectively. The structures in the first and third clusters show similar hydroxamate

Table 1. MMP Residues Subjected to Significant Chemical Shift Perturbations (in bold)[a]

seq	163									171	..	174											186	..	204					209
MMP12+NNGH	F	A	R	**G**	**A**	*H*	*G*	D	**D**	..	F	**D**	G	*K*	**G**	**G**	**I**	**L**	**A**	H	**A**	F	G	..	T	T	**H**	**S**	**G**	G
MMP1+NNGH	F	V	R	G	D	H	R	D	N	..	**F**	**D**	G	*P*	G	G	**N**	L	A	H	**A**	F	**Q**	..	T	N	N	F	R	E
MMP1+act	**F**	V	R	G	D	H	R	D	N	..	**F**	**D**	G	*P*	G	G	**N**	L	**A**	H	A	F	**Q**	..	T	N	N	F	R	E
MMP1+SIMP1	**F**	V	*R*	**G**	**D**	H	R	D	**N**	..	**F**	**D**	G	*P*	G	**G**	**N**	*L*	*A*	*H*	**A**	*F*	*Q*	..	T	*N*	**N**	**F**	R	**E**
MMP1+gal	F	V	R	G	D	H	R	D	**N**	..	F	**D**	G	*P*	**G**	G	*N*	*L*	*A*	**H**	A	F	Q	..	T	N	N	F	R	E

seq	210																				230		235							243	249
MMP12+NNGH	**T**	**N**	L	F	L	**T**	**A**	V	H	E	I	G	H	S	L	G	L	G	H	**S**	S	..	V	M	**F**	*P*	**T**	**Y**	K	**Y**	V	**S**
MMP1+NNGH	Y	N	L	H	R	**V**	A	**A**	H	**E**	L	G	H	S	L	G	**L**	**S**	**H**	**S**	T	..	L	M	Y	P	**S**	**Y**	T	F	**S**	A
MMP1+act	Y	N	L	H	R	**V**	A	**A**	**H**	E	**L**	G	H	**S**	L	G	L	**S**	H	S	T	..	**L**	**M**	Y	*P*	S	Y	T	F	S	**A**
MMP1+SIMP1	Y	N	L	H	R	**V**	A	**A**	**H**	E	**L**	G	**H**	**S**	L	G	L	**S**	**H**	S	T	..	**L**	**M**	**Y**	*P*	**S**	**Y**	*T*	*F*	S	A
MMP1+gal	Y	N	L	H	R	**V**	**A**	A	**H**	E	**L**	G	H	**S**	L	G	L	**S**	**H**	S	T	..	*L*	**M**	**Y**	*P*	S	**Y**	T	*F*	**S**	A

[a] Residue numbers refer to the MMP-1 sequence. Chemical shifts for residues in italics are not available.

coordination to the catalytic zinc. The ligands in the second cluster are oriented similarly to those in the first cluster, but the hydroxamic acid is coordinated to zinc only though the carboxylic oxygen. The structures in the fourth cluster show coordination of the sulfonate oxygen (SO) atoms to zinc. In all cases the *p*-methoxy-phenyl group sits in the S1′ hydrophobic pocket. The position of the *i*-butyl group changes in the four adducts. In the first and second clusters it prevents the formation of hydrogen bonding between the hydroxamic HN and alanine 182 oxygen, whereas the latter hydrogen bond is present in the third cluster.

NOE restraints were obtained in the following way. In the ^{15}N NOESY−HSQC spectrum, a cross-peak at chemical shift of 11.7 ppm is present in the N leucine181 plane (Figure 4). Such a shift is too high to be assigned to a protein signal, as no tryptophan residue is close to the active site. Therefore, it was assigned to the unique amide proton of NNGH. Aromatic protons of NNGH are close to N of residues glycine 221, histidine 222, alanine 216, and arginine 214 according to the structures calculated by Autodock. We have searched in the spectrum all the long range NOEs between aromatic protons and amide groups of these residues. New peaks in the spectrum of the adduct that cannot be due to intraprotein interactions actually appear in the aromatic region (Figure 4) and were assigned as reported in Table 2.

The structural families obtained with Xplor-NIH starting from the first three lowest Autodock docking energy structures converged to the same conformation (see Figure 5). This conformation was similar to the conformation of the third Autodock cluster, with the exception that the sulfur oxygen H-bonded to alanine 182 was the most external oxygen atom rather than the internal one. The lowest total energies were −622, −617, and −615 kcal mol^{-1}, respectively. The lowest total energy of the structural family obtained with Xplor-NIH starting from the fourth Autodock structure was −592 kcal mol^{-1}. This adduct, slightly different from the other three for the fact that zinc coordination by hydroxamate was loose, can be excluded due to its larger energy.

No appreciable changes in the protein side chain positions are observed and thus further Autodock/Xplor-NIH cycles were not needed. Therefore, the structural family shown in Figure 3C represents an experimentally validated and unique structural model for the MMP-1−NNGH adduct.

MMP-1−Actinonin. Actinonin, whose molecular structure is reported in Chart 1, is a well-known inhibitor of aminopeptidases and peptide deformylase.[25] It is also a strong inhibitor for some MMPs, with a K_i of 300 nM for its adduct with MMP-1.[26]

Chemical shift perturbations again allow us to map the region of interest on the protein surface. Residues 215, 217, 218, 220, 223, and 227 forming the metal binding site, residues 235, 236, and 249 on the loop that covers the S1′ pocket, and residues 180 and 182 on the spatially close strand (see Figure 1D) define the ligand binding region and were used for the definition of the Autodock grid. Despite the incomplete correspondence of the affected residues with those found for the NNGH adduct, the resulting grid was quite similar. Four clusters were then calculated (docked conformations were again clustered according to a maximal RMSD of 1.0 Å, see Figure 6). In all structures the hydroxamate is bound to the catalytic zinc. However, whereas in the first two structures the pentyl group is located inside the S1′ hydrophobic pocket and the external propyl group is differently oriented, in the third and fourth structures the two groups are interchanged. The lowest docking energy for the structures in the four clusters were −19.91, −19.11, −18.82, and −18.62 kcal mol^{-1}, respectively.

Cross-peaks of all protons belonging to the ligand with the HN protein protons expected at distances shorter than 5 Å for one or another cluster were looked for in the ^{15}N NOESY−HSQC spectrum. Since the NH of tyrosine 240 has two unassigned cross-peaks at frequencies typical of methyls, they must be related to two methyl groups that are close in the structure of the adduct. From the clusters generated by Autodock, they can only be H′ and H‴. The following peaks were thus assigned: (a) methyl protons H‴ with tyrosine 240 and with the aligned threonine 241 and (b) H′ with tyrosine 240. Among the clusters generated by Autodock, the third and fourth clusters can be readily excluded, because in such structures the above cross-peaks could not be observed. Therefore, by looking at the other two clusters, we also assigned the following cross-peaks, which cannot be assigned to other intraresidue protons or to side chain protons of close residues: (c) alanine 184 with H1 and tyrosine 240 with H6, as such protons are the closest to the coupled HN protons; (d) leucine 181 and tyrosine 240 with H′, as they are aligned and close to one another.

Xplor-NIH calculations were thus performed to refine the selected Autodock structures. Actually, we per-

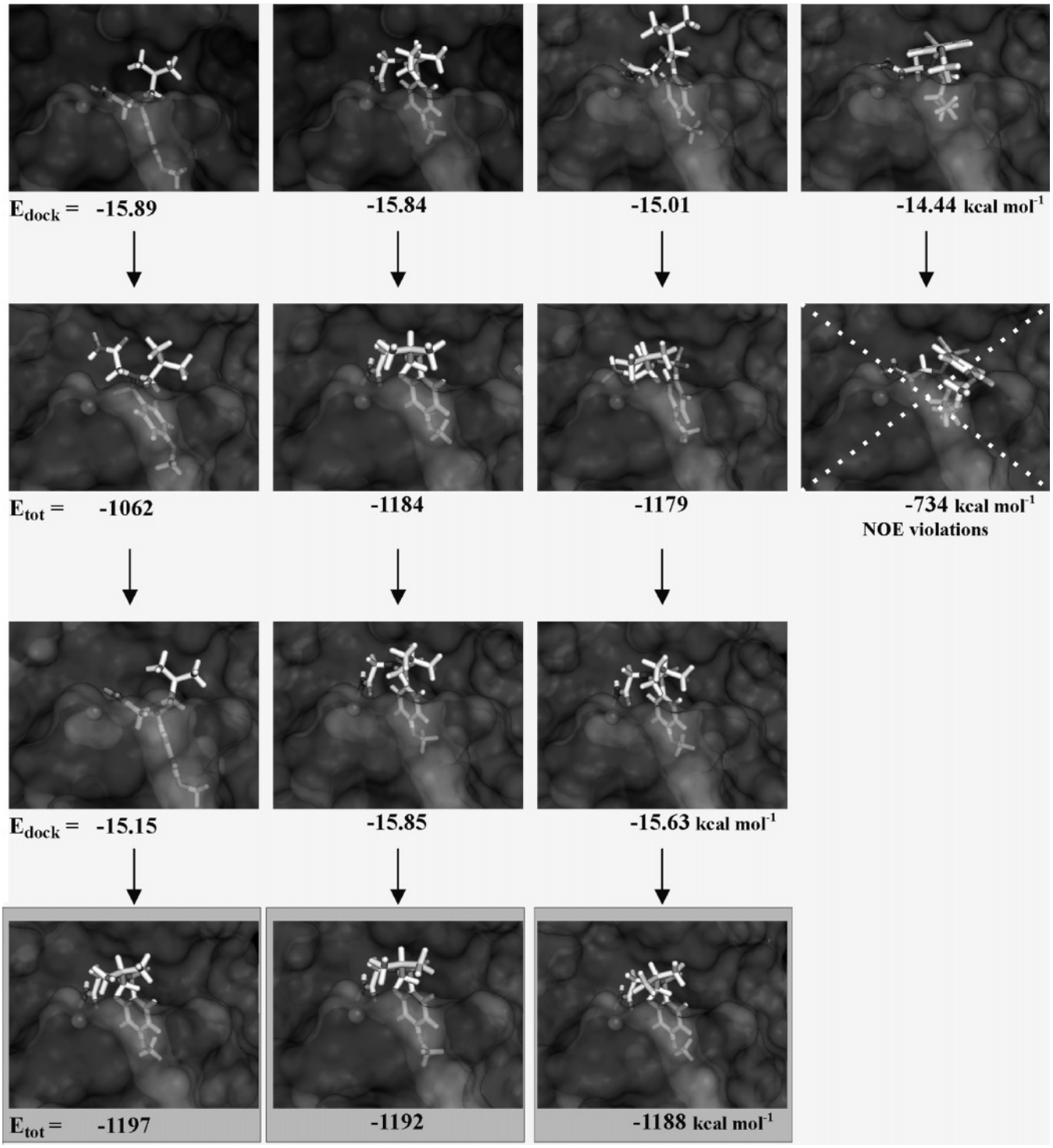


Figure 2. Representative structures of the MMP-12−NNGH adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted.

formed the calculations not only starting from the first two structures, but also starting from the structures excluded according to the observation of the ^{15}N NOESY−HSQC spectrum. The first and second family of structures calculated with Xplor-NIH are very similar to the corresponding Autodock structures; the third Xplor-NIH structural family shows significant rearrangements in the position of the ligand branches, but the pentyl group remains located outside the hydrophobic pocket; in the fourth Xplor-NIH structural family the pentyl group lies in the hydrophobic pocket, thus resulting similar to the first and second families. Xplor-NIH energies for the four families are −870, −862, −811, and −840 kcal mol^{-1}, respectively. This indicates that the third structure, quite different from the other three, is not acceptable. The calculations show that the method is indeed robust. In fact, the first two Autodock structures that were selected from the observation of the NMR spectra actually have the lowest energy, whereas the third has a sizably larger energy even after

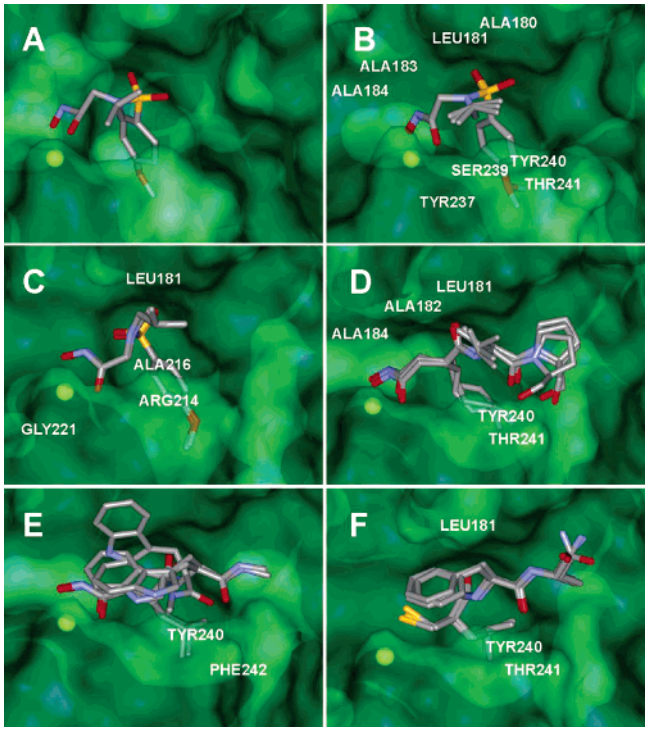


Figure 3. X-ray structure of the MMP-12−NNGH adduct (A), structures calculated with the proposed protocol of the MMP-12−NNGH adduct (B), and structures calculated with the proposed protocol for the adduct of MMP-1 with NNGH (C), actinonin (D), galardin (E), SIMP-1 (F). Labels in panels B−F indicate the residue numbers of amino acids exhibiting NOE contacts to the ligands.

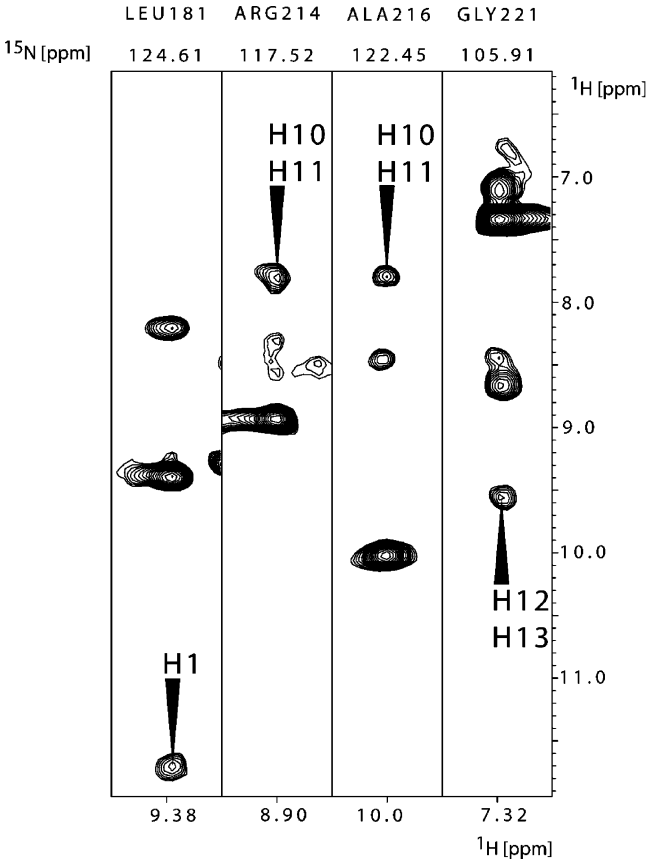


Figure 4. Protein−ligand cross-peaks observed in the ^{15}N NOESY−HSQC spectrum of the MMP-1−NNGH sample.

Xplor-NIH refinement. Interestingly, the fourth Autodock structure, initially completely different from the first two, was brought by Xplor-NIH calculations to converge with the first two.

A second Autodock and Xplor-NIH cycle was performed starting from the lowest energy protein structure. The calculated Xplor-NIH structures, in fact, showed slightly different positions of protein side chains, in particular of residues leucine 181, proline 238, and tyrosine 240. Such new protein conformation was provided to Autodock for a new docking calculation. The best four Autodock clusters (docking energy −20.61, −19.07, −18.99, and −18.88 kcal mol^{-1}) were then provided to Xplor-NIH. The first, third and fourth clusters display both the hydroxamate and the pentyl group in similar positions; the second one is completely different (the hydroxamate does not bind the zinc ion). All lowest energy Xplor-NIH structures (see Figure 3D), with the exception of those calculated starting from the second Autodock structures, converged to the third Autodock conformation, and are equivalent to the lowest energy Xplor-NIH family calculated in the first cycle. The total energies for these structures are −881, −880, and −870 kcal mol^{-1}. The Xplor-NIH structure calculated starting from the second Autodock structure has a total energy of −810 kcal mol^{-1} and can thus be excluded. Therefore, the structure family of Figure 3D is a unique structural model for the MMP-1−actinonin adduct.

MMP-1−Galardin. Galardin (see Chart 1) is a broad spectrum peptidomimetic inhibitor of MMPs[16] with an IC_{50} of 1.5 nM for MMP-1.[27] Chemical shift perturbation involved residues 215, 216, 218, 220, 223, 227, and 228 at the metal binding site, 236, 237, 240, and 243 at the large loop covering the S1′ cavity, and 179 and 183 at the strand facing the S1′ cavity and the metal binding site (Figure 1E). These residues were used to define the Autodock grid, which again was found very similar to the previous ones. The four lowest docking energy clusters calculated by Autodock (−19.79, −19.51, −18.59, and −17.48 kcal mol^{-1}, respectively) showed the following features (see Figure 7). In the first, second and third cluster the *i*-butyl group enters the S1′ pocket, whereas in the fourth cluster it is outside. The structures in the first and second clusters are very similar, as they differ only for the orientation of the indole group, positioned outside the S1′ pocket. The structures in the third and fourth cluster are quite different from those in the first and second cluster, including the position of the indole group, which in any case remains outside the S1′ pocket.

In the ^{15}N NOESY−HSQC spectrum, two cross-peaks are present in the N phenylalanine 242 and tyrosine 240 planes. Such peaks are at chemical shifts typical of methyl groups and cannot be assigned to any intraresidue proton or proton of close residues. Since in galardin there are three methyl groups, two of them being close in the structure, the latter (H′ and H″) were assigned to these peaks. Another cross-peak is present that cannot be assigned to protein protons in the plane of tyrosine 240. This cross-peak falls into the aliphatic region, and therefore it could be provided by CH or CH_2 protons. Since such proton must be close to H′ and H″, which have also a cross-peak with tyrosine 240, it was assigned to H6 or H7.

Table 2. Observed NOEs between MMP-1 Amide Protons and Ligand Protons

	181	182	184	214	216	221	240	241	242
NNGH	H1			(H10,H11)	(H10,H11)	H12 H13			
actinonin	H′	(H3,H4) H6	H1				H‴ H′ H6	H‴ (H23,H24)	
galardin							H′ H″ (H6,H7)		H′ H″
SIMP-1	(H17,H18,H19,H20,H21)						H′ H″	H″	

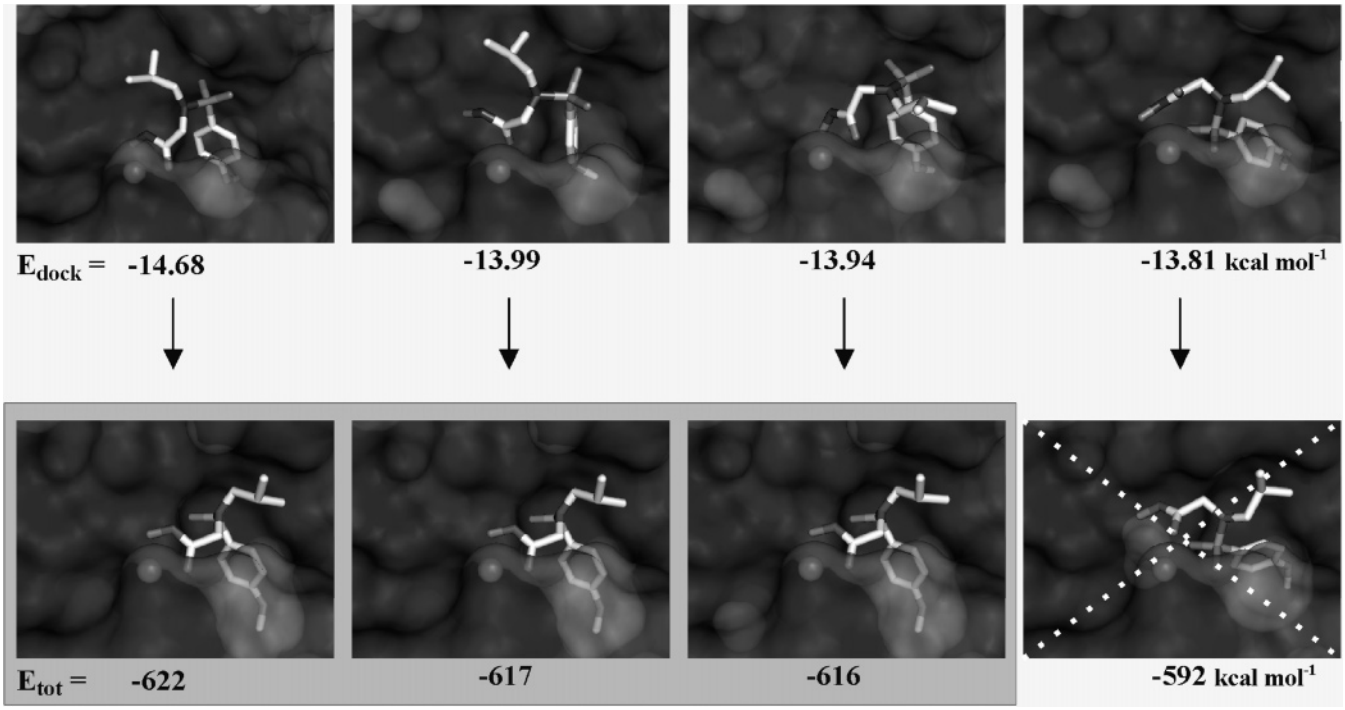


Figure 5. Representative structures of the MMP-1−NNGH adduct for the four lowest energy clusters obtained from Autodock (first row) and from Xplor-NIH (second row). At this point convergence was obtained. The final validated structures are highlighted.

Xplor-NIH calculations change only slightly the conformations obtained with Autodock relative to the first three clusters. The structure obtained starting from the fourth Autodock cluster is instead modified by the NOE restraints to have the *i*-butyl group inside the S1′ pocket as in the other three clusters. The total energy of the Xplor-NIH structures are −575, −578, −565, and −527 kcal mol^{-1}, respectively. Only small changes in the side chain positions have been observed, regarding in particular residues from 238 to 241.

The three protein structures with the smallest Xplor-NIH energy were used to repeat Autodock calculations. In the first case Autodock produced the two lowest docking energy clusters very similar to those obtained in the first run (−18.87 and −18.57 kcal mol^{-1}), whereas the third and fourth clusters (with docking energy −18.08 kcal mol^{-1}) have now the indole group inside the S1′ pocket. These conformations can be excluded by the observed NOEs. It is remarkable that such faulty Autodock behavior occurs in the second round, i.e., after adjustment of the structure by Xplor-NIH minimization. This observation underlines the need for experimental restraints to gain confidence in in silico models. In the second case, the three lowest docking energy clusters are again very similar to those obtained in the first run (−19.59, −19.56, and −18.14 kcal mol^{-1}), whereas in the fourth cluster (with docking energy −17.83 kcal mol^{-1}) the *i*-butyl group is outside the S1′ pocket. In the third case, the lowest docking energy cluster is again similar, with energy −18.42 kcal mol^{-1}. The Xplor-NIH calculations performed with the four lowest docking energy structures as starting conformations converged to a unique conformation (−586, −584, −576, and −576 kcal mol^{-1}), except for the indole group, which, being outside the S1′ pocket, is free to move (Figure 3E). Again, the family of Figure 3E can be confidently taken as a validated structural model for the galardin adduct of MMP-1.

MMP-1−SIMP-1. SIMP-1 is a polypeptide derivative able to inhibit collagenases.[16] Its molecular structure is reported in Chart 1. We measured an IC_{50} value for the adduct of 46 nM. In the MMP-1−SIMP-1 adduct, affected resonances include residues 215, 217−218, 220, 222, 223, 227, and 228 on the α helix of the zinc binding site, residues 235−237, 239, 240, and 242 forming the S1′ cavity, and residues 179, 180, and 184 on the strand facing both the catalytic metal and the S1′ pocket (Figure 1F). The four clusters with smallest docking energy calculated by Autodock (see Figure 8) have docking energy of −16.15, −15.92, −15.76, and −15.72

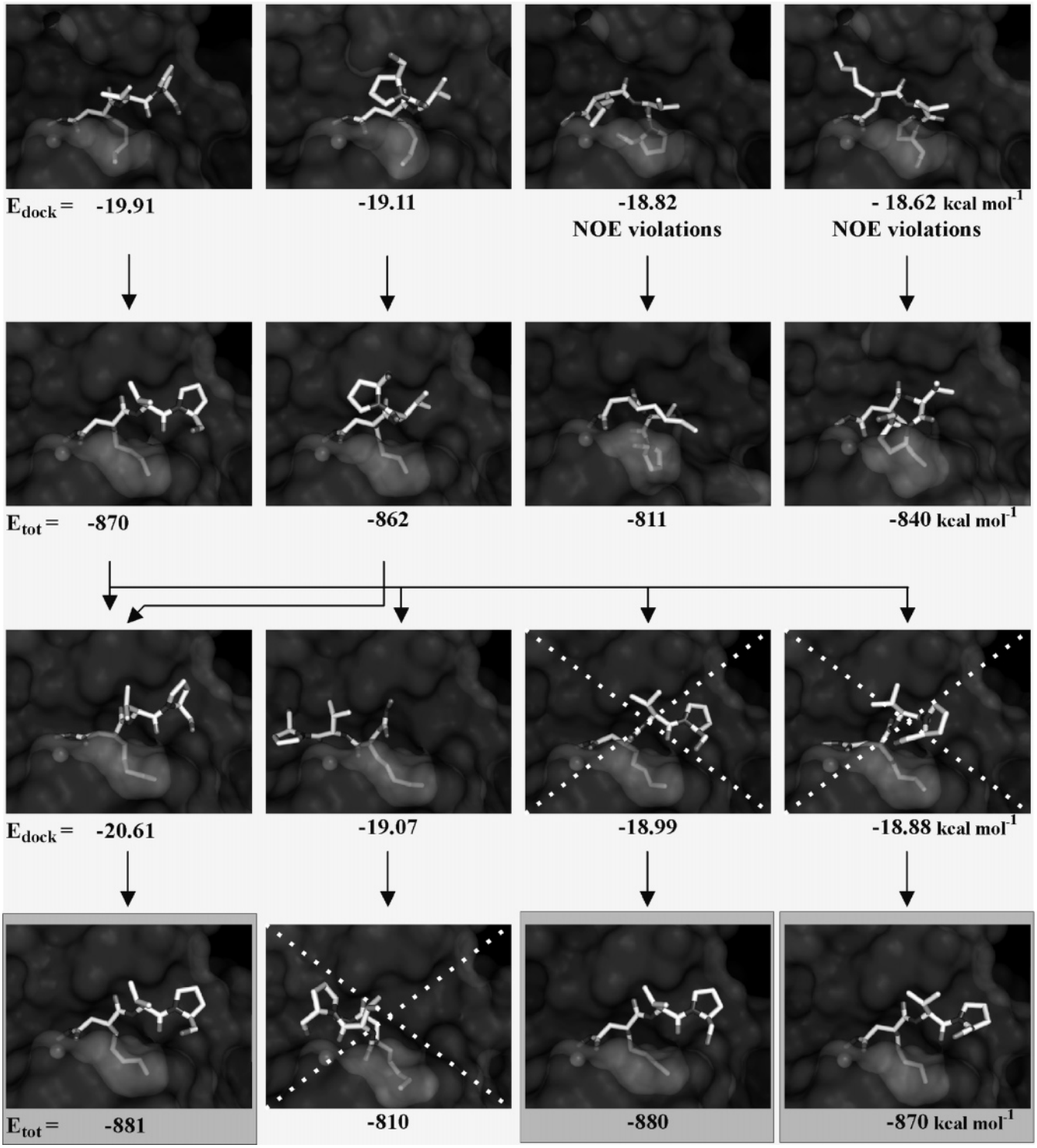


Figure 6. Representative structures of the MMP-1−actinonin adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH calculations (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted.

kcal mol^{-1}. In the first and second clusters the sulfur atom coordinates the catalytic zinc; in the first cluster the S1′ pocket interacts with the ligand benzyl group, in the second with the *i*-butyl group. In the third cluster the sulfur atom is hydrogen bonded to the oxygen of glycine 179, on the other site of the catalytic pocket with respect to the zinc ion, and the ligand benzyl group sits in the S1′ pocket. In the fourth cluster, the ligand is oriented similarly as in the first cluster, but the ligand sulfur atom is loosely coordinated to the zinc ion, and hydrogen bonded to glutamate 219.

In the ^{15}N NOESY−HSQC spectrum, in the N plane of residue leucine 181, there are two signals in the aromatic region that cannot be assigned to protein side chains. Therefore, they must be assigned to protons of the aromatic ring of the SIMP-1. Two cross-peaks, one of low and one of high intensity, are present in the N plane of tyrosine 240 at frequencies typical of methyl

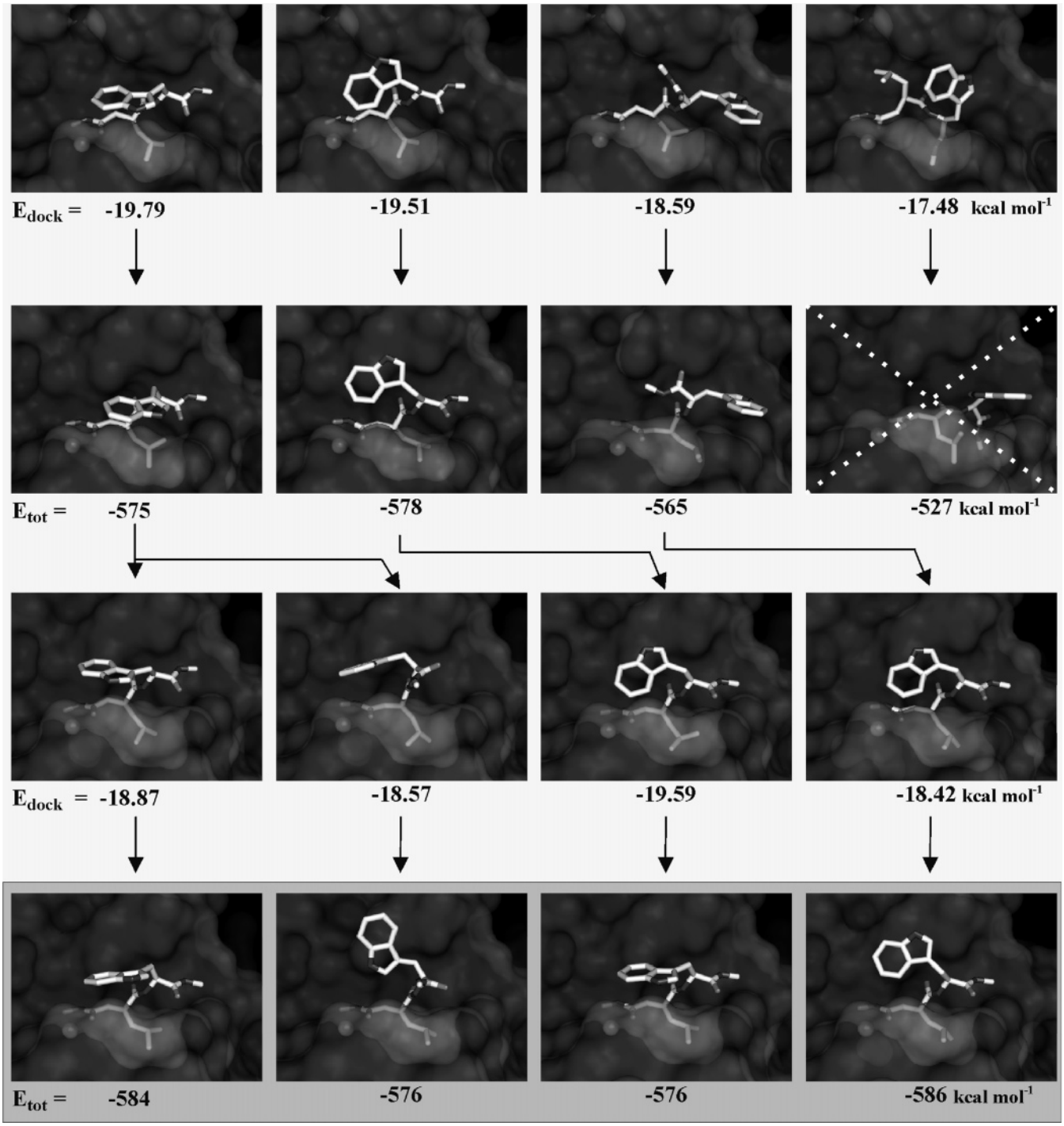


Figure 7. Representative structures of the MMP-1−galardin adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH calculations (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted. The configuration of the indole ring is not defined because of lack of experimental restraints and strong energetic preference in Xplor-NIH calculations.

groups that cannot be assigned to intraresidue or sequential contacts. In one of the clusters calculated by Autodock, N of tyrosine 240 is close to two of the three methyls of SIMP-1, H′ being closest than H″, and thus the cross-peaks were correspondingly assigned (see Table 2). A further distance restraint is determined from another cross-peak in the N plane of threonine 241, aligned with the signal assigned to H″.

Xplor-NIH calculations select the second Autodock cluster as the correct one. In fact it remains almost unchanged after refinement, with total energy −709 kcal mol^{-1}. Calculations performed starting from the other clusters provide structures very different from the starting ligand conformation, and with the ligand not coordinated to the zinc ion. Their total energies are larger than −634 kcal mol^{-1} and such structures are thus excluded.

Slight changes in the protein side chain positions are observed, in particular on residues 180, 214 and 219. A second Autodock calculation was thus performed. The first three clusters (with docking energy of −16.15, −16.12 and −15.88 kcal mol^{-1}, respectively) show a

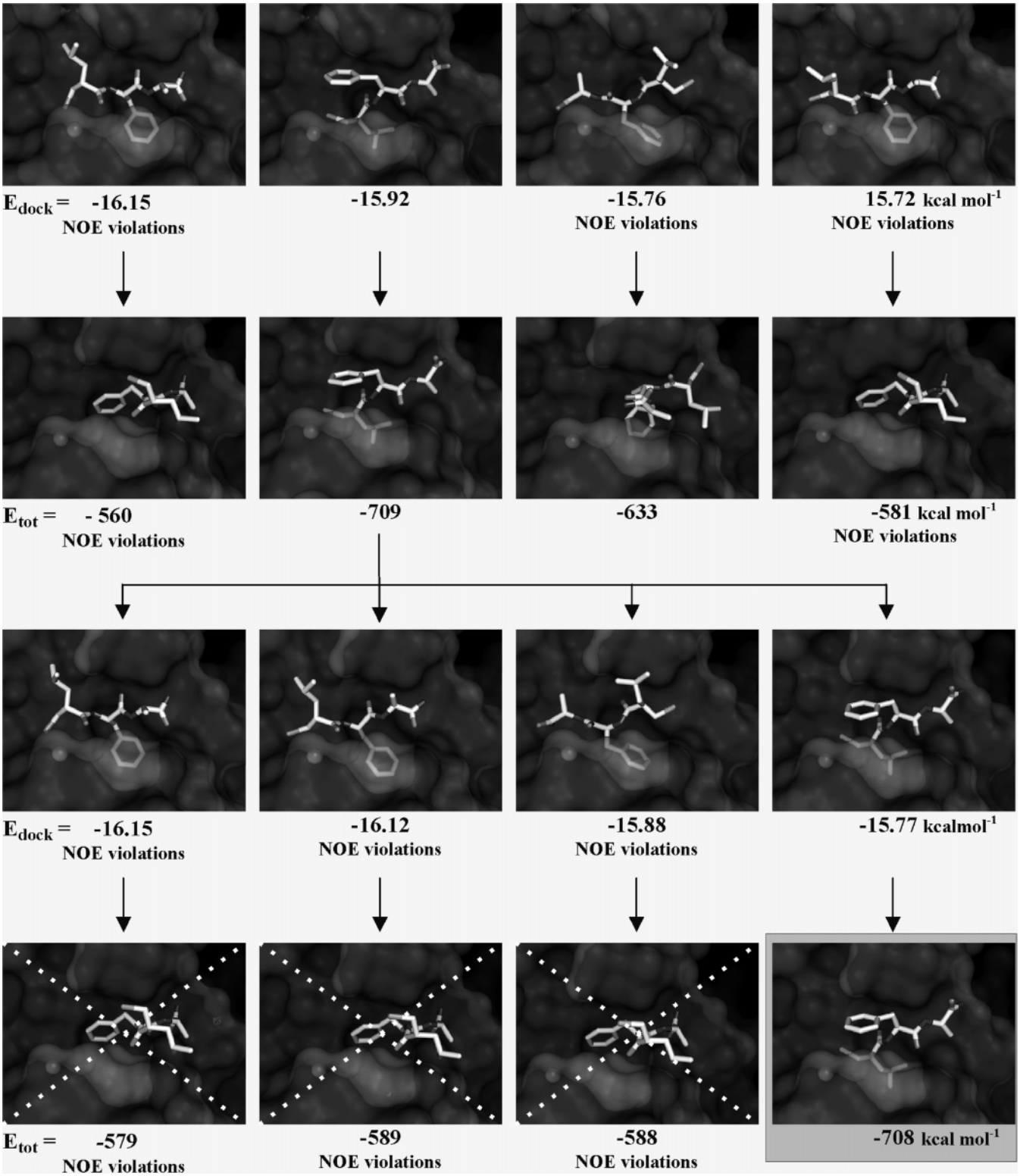


Figure 8. Representative structures of the MMP-1−SIMP-1 adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH calculations (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted.

ligand pose similar to that calculated in the first and fourth clusters of the first Autodock run. The fourth cluster, with docking energy −15.77 kcal mol^{-1}, is instead similar to the pose already identified as correct. Xplor-NIH calculations again confirmed such structure as the correct one, with total energy −708 kcal mol^{-1}. The corresponding family is shown in Figure 3F. This family represents the validated structural model of the MMP-1−SIMP-1 adduct.

Backbone Mobility. To test the protocol for possible protein backbone rearrangements upon complexation, Xplor-NIH calculations were also performed with allowing the protein backbone to move in the protein region affected by chemical shift perturbation. In all

cases we found no appreciable differences in the results. In fact, for all adducts the lowest energy structures corresponded to those identified as correct in the calculations performed with rigid backbones.

Discussion

A protocol has been developed to merge the "pure" docking capability of Autodock (or other docking programs) with the exploitation of available experimental restraints. For the relatively strong ligands (K_{diss} approximately micromolar or less) elected here, the protocol has been shown to be efficient, robust and reliable. As shown in Scheme 1, the protein binding site is identified from chemical shift perturbation in the HSQC spectrum of the protein upon complexation. The observation of shift perturbations on passing from the assigned spectrum of the free protein to the spectrum of the adduct permits the definition of the protein grid to be used in Autodock calculations. Autodock usually provides several clusters of structures for the adduct, which often have similar docking energy. These structures are used to calculate maps of NOEs, to be compared with NOEs actually observed in the ^{15}N NOESY−HSQC spectrum of the adduct. A few ligand−protein NOEs can always be assigned, and the latter can be used as restraints in Xplor-NIH calculations for selection, validation and refinement of the Autodock structures. One-two cycles at most may be needed in case Xplor-NIH calculations modify some protein side chain positions with respect to the structure provided to Autodock. All these steps could be performed semiautomatically, if required.

The protocol relies on the following information to be available: the protein structure; the assigned HSQC spectrum of the free protein; the ^{15}N NOESY−HSQC spectrum of the free protein; the HSQC and ^{15}N NOESY−HSQC spectra of the protein−ligand adduct; and the 1D ^{1}H spectrum of the ligand. The protocol has been developed in order to avoid preparation of doubly labelled samples and assignment of protein side chains, thus resulting in a much faster throughput.

We have shown that such approach is actually efficient in finding the protein−ligand structure for four adducts of MMP-1 with different ligands. The peculiarity that makes this approach successful in the cases here examined is the combination of a docking program, able to quickly and efficiently sample the possible binding poses, with a molecular dynamics program, which selects the proposed poses using few unambiguous experimental data. In this way the efficiency of the former program is coupled to the complexity of the latter, which also allows for protein side chain movements. The program has been deliberately tested using only unambiguous NOEs obtainable from the assignment of HN, but it is obviously open to the use of additional or different restraints. We decided to use the chemical shift perturbations only for the determination of the grid to be used for the docking program calculations, without including them as restraints in the molecular dynamics program due to their ambiguous nature, although ambiguous restrains could be in principle used, either as such, as recently proposed,[13] or through calculation of *j*-surfaces.[10] The use of chemical shift perturbations for the determination of the grid is much less stringent than their use as constraints, as a few "second sphere" shifts erroneously mistaken for first sphere shifts may drive the ligand in wrong positions, while the resulting grids are expected to be only somewhat broadened. As a matter of fact, differences in perturbed residues from one ligand to another do not result in grossly different grids, and the latter, in all cases, encompassed the whole catalytic site.

Several predicting programs for protein−ligand adducts have been proposed in the literature. Inclusion of biochemical and biophysical data in docking protocols, called guided docking,[28,29] is a common approach to reduce the conformational variety of the proposed solutions. Some other programs[7−9,30−35] work totally in silico, without experimental information on the investigated adduct, and perform docking calculations with an improved level of sophistication. They can be successful, but the level of confidence for the proposed adduct is difficult to establish. Furthermore, a strong bias toward known solutions or preconceived requirements is introduced if the docking is restrained according to chemical information derived from databases of protein−ligand complexes. Other programs[36−39] use the experimental NMR information more systematically, thus being similar to structural determination programs and therefore more time-consuming. NMR-derived restraints were also used in docking programs to identify the location of the ligand binding[10] and to restrict the conformational space for molecular modeling routines.[11] NMR experiments on selectively labeled proteins were also used to obtain structural information on protein−ligand complexes.[12] This approach, although more expensive than the one here proposed, is probably the only viable in case of large proteins. To our knowledge this is the first time that an approach is proposed where few experimental data are used to select and refine poses proposed by fast docking programs.

Autodock has been selected among the docking programs because in the case of MMPs it was demonstrated to be a robust program with good docking accuracy and reliability, including the correct geometry of the zinc binding groups.[21,40] It employs a genetic algorithm searching function, able to efficiently sample large search spaces. Different docking programs could however be used if considered more reliable in other cases. In the same way, other molecular dynamic programs could be used instead of Xplor-NIH. We used Xplor-NIH as an NMR-oriented widespread general program for structural calculations using simulated annealing. Ligand growing procedures[30] may also be implemented in Xplor-NIH, resulting probably useful especially in case of large ligands.

Although the presence of the metal ion in MMPs tends to restrict the number of Autodock clusters by favoring poses where the hydroxamic moiety is coordinated to the metal, the protocol is expected to be useful also in case of proteins not containing catalytic ions. Actually, docking programs are developed to work mainly in their absence, and, in case they propose several different conformations, the detection of NOEs may result decisive for the selection of the correct one. Indeed, as we have seen, Autodock does not always succeed in correctly binding the metal to the hydroxamic moiety. Furthermore, in the absence of the metal, further

H-bonds or van der Waals contacts should occur for strong ligands, which would likely provide additional intermolecular NOEs.

We have shown that it is possible to obtain few intermolecular experimental NOEs through fast NMR experiments without the necessity to assign all protein NOESY cross-peaks. Only unambiguous NOEs between protein and ligand protons have been considered; therefore, cross-peaks were assigned to ligand protons only if they could not be reasonably assigned to any protein side chain proton, taking into account the structural adducts proposed by Autodock. In all cases here addressed, experimental restraints have been shown to be necessary and sufficient to extract the adduct conformation among the several proposed by Autodock with similar docking energy, and thus are used to validate them. Furthermore, the approach proposed can also be useful to refine the structure of the ligand−protein adduct, especially because local small modifications in the protein structure (of side chains, if sufficient as in the present case, but also in the protein backbone, if needed − see below) can be accommodated by cycling between Autodock/Xplor-NIH runs. This makes the present approach preferable to the direct introduction of distance restraints in docking programs with a fixed protein matrix.

The solution structure of the inhibitor-free MMP-1, obtained from a series of 3D triple-resonance NMR experiments, shows nearly identical both backbone and secondary structures than the crystallographic structures.[41] Furthermore, the backbones of the solution structures of the inhibitor-free MMP-1 and of the MMP-1 complexed with a sulfonamide derivative of the hydroxamic acid compound have been shown to be essentially identical,[42] although mobility measurements indicate that the region near the active site is highly mobile.[41,42] It is thus reasonable, at least in our case, to assume that the protein backbone remains rigid during complexation in solution, and with structure identical to the crystallographic structure.

Although not necessary for the present calculations, also the protein backbone could be allowed to (partially) move in Xplor-NIH calculations (see results). This could be important if modest backbone rearrangements are expected upon ligand binding, as could be indicated by chemical shift perturbations spread out over a wider region.

It is known that effective MMP inhibitors achieve tight binding via extensive van der Waals contacts with the hydrophobic interior of S1′ and by strong electrostatic interactions with zinc and nearby charged or polar side chains.[43] All calculated adducts indeed show ligand coordination to the catalytic zinc and the formation of a net of hydrogen bonds between ligand and protein residues. This result is not trivial as it may seem, as several of the initially obtained Autodock structures had severely distorted − or were even lacking − hydroxamate coordination to the zinc ion.

The distance between zinc and hydroxamate oxygens is in all calculated structures between 1.95 and 2.25 Å. The O−Zn−O angle is always between 86 and 93°. The coordination geometry is distorted square-pyramidal in MMP-12−NNGH and MMP-1−actinonin, and distorted trigonal bipyramidal, with hydroxamic O2 and N histidine 222 in axial positions, in MMP-1−NNGH and galardin. All hydrogen bonding interactions between MMPs and ligands are reported in Table 3. In particular, H-bonds are present in all adducts with NNGH, actinonin, and galardin between oxygen of alanine 182 and the amide proton of the hydroxamic group, as well as between the protonated glutamate 219 and the oxygen of the hydroxamic group. H bonds are also present between ligands and HN of Leu 181, as previously seen in the MMP-1-CGS[42] and in the MMP-12−NNGH adducts. In the MMP-1−SIMP-1 adduct, with a distorted tetrahedral coordination geometry around the zinc ion, constituted by the three histidine nitrogen atoms and the sulfur SIMP-1 atom, a net of hydrogen bonds is formed, connecting the ligand to the protein atoms (see Table 3). Both the coordination geometry and the H-bonding network can be used to assess the reliability of the obtained adducts. In all the adducts, the inhibitors establish enough interactions to reach nanomolar affinity. In particular, all the ligands bind the metal, place a lipophilic moiety into the S1′ cavity and establish two or more hydrogen bonds with atoms of the protein groove. This binding mode is reasonable and is indeed adopted by many strong ligands of MMPs.

Table 3. Predicted H-bonds between Protein and Ligand Nuclei

MMP-12−NNGH	181 Leu HN	NNGH O4/O3
	182 Ala HN	NNGH O3
	219 Glu HO1	NNGH O1
	NNGH H1	182 ALA O
MMP-1−NNGH	182 Ala HN	NNGH O4
	219 Glu HO1	NNGH O1
	NNGH H1	182 Ala O
MMP-1−actinonin	181 Leu HN	act O3
	219 Glu HO1	act O1
	240 Tyr HN	act O4
	act H1	182 Ala O
	act N2	238 Pro O
MMP-1−galardin	181 Leu HN	gal O3
	182 Ala HN	gal O3
	219 Glu HO1	gal O1
	240 Tyr HN	gal O4
	gal H5	238 Pro O
	gal H1	182 Ala O
MMP-1−SIMP-1	181 Leu HN	SIMP1 O1
	182 Ala HN	SIMP1 O1
	240 Tyr HN	SIMP1 O2
	SIMP1 H8	238 Pro O
	SIMP1 H1	219 Glu OE1
	SIMP1 H23	179 Gly O
	SIMP1 H26	210 Tyr OH

Cycling between fast docking programs and Xplor-NIH calculations can be used to assess ligand−protein structures also in the presence of restraints different from NOEs. Diamagnetic residual dipolar couplings have already been demonstrated to be extremely useful to predict the structure of protein−protein adducts.[20,44−46] Also pseudocontact shifts have been used for the study of protein−protein docking.[47] Paramagnetism-based restraints, and in particular paramagnetic relaxation rates, pseudocontact shifts and residual dipolar couplings, arising when a paramagnetic metal ion is coordinated to the protein, could be employed as restraints in the proposed protocol for protein−ligand docking. Xplor-NIH has the advantage that it already contains the tools needed to deal with such restraints.[48]

Conclusions

A novel protocol to obtain validated structural models of protein−ligand complexes has been developed and applied for the determination of the structure of the adducts of the protein MMP-1 with four different ligands. The method was shown to be reliable, as tested for the known structure of the adduct of one of these ligands with MMP-12. It uses NMR derived restraints obtained using singly (^{15}N) labeled proteins. The strategy that we propose promises to be generally useful also for the structural determination of different protein−ligand adducts, whenever the structure of the free protein is known and the structural changes upon complexation are not expected to be dramatic.

Experimental Section

Sample Preparation. The fragment of human fibroblast collagenase corresponding to proMMP-1 (Pro21-Pro269) and bearing an additional methionine at the N-terminal, was expressed in *Escherichia coli*. The cDNA was cloned into the pET21 vector (Novagen) using *Nde*I and *Xho*I as restriction enzymes. The *E. coli* strain BL21 Codon Plus cells, transfected with the above vector, were grown in 2 × YT media at 37 °C. The protein expression was induced during the exponential growth phase with 0.5 mM of IPTG. Cells were harvested for 4 h after induction. Uniform^{15}N-labeled protein was obtained by growing the transfected BL21 Codon Plus cells in minimal media at 37 °C. The cells were lysed by sonication and the inclusion bodies, containing the proMMP-1, were solubilized in 2 M urea, 20 mM Tris (pH 8.0). The protein was purified on the Hitrap Q column (Pharmacia) with a buffer containing 2 M urea and 20 mM Tris (pH 8.0). The elution was performed using a linear gradient of NaCl up to 0.35 M. The purified protein was then refolded by using a multistep dialysis against solutions containing 50 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl. The refolded protein was exchanged, by dialysis, against a buffer with 10 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl. The protein was activated by 1 mM APMA (4-aminophenylmercuric acetate) at 4 °C overnight and dialyzed with a buffer containing 10 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl, 0.2 M acetohydroxamic acid (AHA). The activated protein (Val 101−Pro 269) was concentrated using an Amicon stirrer and Centriprep concentrators, fitted with a YM10 membrane in nitrogen atmosphere at 4 °C. Catalytic domain of MMP-1 was purified using size-exclusion chromatography with the final dialysis buffer and concentrated up to 0.5 mM using an Centriprep concentrators in nitrogen atmosphere at 4 °C. The final protein sample was dialyzed against a solution containing 50 mM sodium acetate, 100 mM NaCl, 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, with 10% of D_2O (pH 6.5).

Inhibited proteins were prepared by titration of the free-MMP-1 with equimolar amounts of NNGH, SIMP-1, galardin, and actinonin.

NNGH, galardin, and actinonin were purchased by BIOMOL International; SIMP-1 was purchased by Peptide International, Inc.

In Vitro Assay. The compounds were evaluated for their ability to inhibit the hydrolysis of fluorescence-quenched peptide substrate Mca-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH_2 (Biomol, Inc.). The assays were performed in 50 mM HEPES buffer, containing 10 mM $CaCl_2$, 0.05% Brij-35, at pH 7, using 1 nM of MMP-1 catalytic domain and 1 μM of peptide. The enzyme was incubated at 25 °C with increasing concentration of inhibitor and the florescence ($excitation_{max}$ 328 nm; $emission_{max}$ 393 nm) was measured for 3 min after the addition of the substrate using a Varian Eclipse fluorimeter. Fitting of rates as a function of inhibitor concentration provided the IC_{50} values. In our experimental conditions with low enzyme concentration and peptide concentration much lower than K_M (the concentration of the substrate that leads to half-maximal velocity of the enzymatic hydrolysis reaction); the IC_{50} values provide a good estimate of the dissociation constant of the adduct. The inhibitor *N*-isobutyl-*N*-[4-methoxyphenylsulfonyl]-glycyl hydroxamic acid (Biomol, Inc.) was used as control.

NMR Measurements. ^{1}H ^{15}N HSQC experiments implemented with the sensitivity enhancement scheme[49] and ^{15}N NOESY−HSQC spectra[50] were performed on the free MMP-1 catalytic domain and on each protein−ligand adduct. ^{15}N NOESY−HSQC experiments were acquired with a mixing time of 110 ms and with data sets comprising 256(^{1}H) × 64(^{15}N) × 2048(^{1}H) data points. The NMR spectra were recorded on Avance 900 Bruker spectrometer, operating at proton nominal frequencies of 900.13 MHz and equipped with a triple resonance cryoprobe. All NMR experiments, recorded at 298 K, were processed using the standard Bruker software (XWIN-NMR), and analyzed through the XEASY program.[51]

Computer Programs. Autodock 3.0.5 was used to predict protein−ligand docking. It uses a Lamarckian genetic algorithm as global optimizer combined with energy minimization as a local search method.[52] Its scoring function is provided by the sum, with empirically determined scaling factors, of a Lennard-Jones 12−6 dispersion/repulsion term, a directional 12−10 hydrogen bond term, a Coulombic electrostatic potential, a term related to unfavorable entropy due to restrictions in conformational degree of freedom of the ligand, and a desolvation term. The PDB file was processed by Autodock Tool Kit. Reliable zinc parameters were provided as in ref 21. A box of 70 × 70 × 70 points with a grid spacing of 0.375 Å was defined as docking space. The ligands were generated and minimized using semiempirical calculations (AM1 type Gaussian 98),[53] and the pdbq files, comprising all protons, were provided to Autodock after all the Gasteiger−Marseli charges[54] were assigned by BABEL. For each run, a maximum number of 28 000 genetic algorithm operations were generated on a single population of 50 individuals. For each ligand, a total of 100 docking runs were performed, and the results were ranked according to the docking energy. Crossover, mutation, and elitism weights were set to 0.80, 0.02, and 1, respectively.

All minimization and dynamics calculations were carried out using the program Xplor-NIH.[55,56] The parameter and topology files for the ligands were generated using Xplo2D,[57] the improper angles being manually edited and the dihedral angles being set with force constant equal to zero. Protein electrostatic and van der Waal energy parameters have been evaluated using CHARMM nonbonded parameters.[58]

Acknowledgment. This work was supported by the European Union (Contract QLG2-CT-2002-00988 and Contract LSHG-CT-2004-512077), by MIUR−FIRB RBNE01TTJW, by MIUR−FISR "Modeling di strutture di metalloproteine e delle interazioni proteina-farmaco e proteina-proteina", and by Ente Cassa di Risparmio di Firenze.

References

(1) Anderson A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* **2003**, *10*, 787−797.
(2) Blundell, T. L.; Jhoti, H.; Abell, C. High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discovery* **2002**, *1*, 45−54.
(3) Lepre, C. A.; Moore, J. M.; Peng, J. W. Theory and Applications of NMR−Based Screening in Pharmaceutical Research. *Chem. Rev.* **2004**, *104*, 3641−3675.
(4) Pellecchia, M.; Becattini, B.; Crowell, K. J.; Fattorusso, R.; Forino, M.; Fragai, M.; Jung, D.; Mustelin, T.; Tautz L. NMR-based techniques in the hit identification and optimisation processes. *Exp. Opin. Ther. Targets* **2004**, *6*, 597−611.
(5) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531−1534.
(6) Bertini, I.; Calderone, V.; Cosenza, M.; Fragai, M.; Lee, Y.-M.; Luchinat, C.; Mangani, S.; Terni, B.; Turano, P. Conformational variability of MMPs: beyond a single 3D structure. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5334−5339.

(7) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(8) Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. Comparison of Automated Docking Programs as Virtual Screening Tools. *J. Med. Chem.* **2005**, *48*, 962−976.
(9) Trosset, J.-Y.; Scheraga, H. A. Reaching the global minimum in docking simulations: A Monte Carlo energy minimization approach using Bezier splines. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8011−8015.
(10) McCoy, M. A.; Wyss, D. F. Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J. Am. Chem. Soc.* **2002**, *124*, 11758−11763.
(11) Lugovskoy, A. A.; Degterev, A. I.; Fahmy, A. F.; Zhou, P.; Gross, J. D.; Yuan, J.; Wagner, G. A Novel Approach for Characterizing Protein Ligand Complexes: Molecular Basis for Specificity of Small-Molecule BcI-2 Inhibitors. *J. Am. Chem. Soc.* **2002**, *124*, 1234−1240.
(12) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D. S. NMR-based structural characterization of large protein−ligand interactions. *J. Biomol. NMR* **2002**, *22*, 165−173.
(13) Schieborr, U.; Vogtherr, M.; Elshorst, B.; Betz, M.; Grimme, S.; Pescatore, B.; Langer, T.; Saxena, K.; Schwalbe, H. How much NMR-Data are required for the Determination of a Protein−Ligand Complex? *ChemBioChem* **2005**, *6*, 1891−1898.
(14) Woessner, J. F., Jr. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J.* **1991**, *5*, 2145−2154.
(15) Yu, A. E.; Hewitt, R. E.; Connor, E. W.; Stetler-Stevenson, W. G. Matrix metalloproteinases: novel targets for directed cancer therapy. *Drugs Aging* **1997**, *11*, 229−244.
(16) Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. Design and Therapeutic Application of Matrix Metalloproteinase Inhibitors. *Chem. Rev.* **1999**, *99*, 2735−2776.
(17) Clendeninn, N. J.; Appelt, K. *Metalloproteinase Inhibitors in Cancer Therapy*; Humana Press: Totowa, NJ, 2001.
(18) Maeda, A.; Sobel, R. A. Matrix metalloproteinases in the normal human nervous system, microglial nodules, and multiple sclerosis lesions. *J. Neuropathol. Exp. Neurol.* **1996**, *55*, 300−309.
(19) Mitton-Fry, R. M.; Anderson, E. M.; Hughes, T. R.; Lundblad, V.; Wuttke, D. S. Conserved Structure for Single-Stranded Telomeric DNA Recognition. *Science* **2002**, *296*, 145−147.
(20) Clore, G. M.; Schwieters, C. D. Docking of Protein−Protein Complexes on the Basis of Highly Ambiguous Intermolecular Distance Restraints Derived from 1HN/15N Chemical Shift Mapping and Backbone 15N-1H Residual Dipolar Couplings Using Conjoined Rigid Body/Torsion Angle Dynamics. *J. Am. Chem. Soc.* **2003**, *125*, 2902−2912.
(21) Hu, X.; Shelver, W. H. Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction. *J. Mol. Graphics Modell.* **2003**, *22*, 115−126.
(22) MacPherson, L. J.; Bayburt, E. K.; Capparelli, M. P.; Carroll, B. J.; Goldstein, R.; Justice, M. R.; Zhu, L.; Hu, S.; Melton, R. A.; Fryer, L.; Goldberg, R. L.; Doughty, J. R.; Spirito, S.; Blancuzzi, V.; Wilson, D.; O'Byrne, E. M.; Ganu, V.; Parker, D. T. Discovery of CGS 27023A, a nonpeptidic, potent, and orally active stromelysin inhibitor that blocks cartilage degradation in rabbits. *J. Med. Chem*. **1997**, *40*, 2525−2532.
(23) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Mangani, S.; Terni, B. X-ray structures of ternary enzyme−product-inhibitor complexes of MMP. *Angew. Chem., Int. Ed.* **2003**, *42*, 2673−2676.
(24) Lovejoy, B.; Hassell, A. M.; Luther, M. A.; Weigl, D.; Jordan, S. R. Crystal Structures of Recombinant 19-kDa Human Fibroblast Collagenase Complexed to Itself. *Biochemistry* **1994**, *33*, 8207−8217.
(25) Chen, D. Z.; Patel, D. V.; Hackbarth, C. J.; Wang, W.; Dreyer, G.; Young, D. C.; Margolis, P. S.; Wu, C.; Ni, Z.-J.; Trias, J.; White, R. J.; Yuan, Z. Actiononin, a Naturally Occurring Antibacterial Agent, Is a Potent Deformylase Inhibitor. *Biochemistry* **2000**, *39*, 1256−1262.
(26) BIOMOL Research Laboratories data (www.biomol.com), 2005.
(27) Yamamoto, M.; Tsujishita, H.; Hori, N.; Ohishi, Y.; Inoue, S.; Ikeda, S.; Okada, Y. Inhibition of Membrane-Type 1 Matrix Matalloproteinase by Hydroxamate Inhibitors: An Examination of the subsite pocket. *J. Med. Chem.* **1998**, *41*, 1209−1217.
(28) van Dijk, A. D. J.; Boelens, R.; and Bonvin, A. M. J. J. Data-driven docking for the study of biomolecular complexes. *FEBS J.* **2005**, *272*, 293−312.
(29) Fradera, X.; Mestres, J. Guided Docking Approaches to Structure-Based Design and Screening. *Curr. Top. Med. Chem.* **2004**, *4*, 687−700.
(30) Ota, N.; Agard, D. A. Binding Mode Prediction for a Flexible Ligand in a Flexible Pocket using Multi-Conformation Simulated Annealing Pseudo Crystallographic Refinement. *J. Mol. Biol.* **2001**, *314*, 607−617.
(31) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.
(32) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235−249.
(33) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225−242.
(34) Carlson, H. A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447−452.
(35) Gervasio, F. L.; Laio, A.; Parrinello, M. Flexible docking in solution using metadynamics. *J. Am. Chem. Soc.* **2005**, *127*, 2600−2607.
(36) Maurer, M. C.; Trosset, J. Y.; Lester, C. C.; DiBella, E. E.; Scheraga, H. A. New general approach for determining the solution structure of a ligand bound weakly to a receptor: structure of a fibrinogen Aα-like peptide bound to thrombin-(S195A) obtained using NOE distance constraints and an ECEPP/3 flexible docking program. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 29−48.
(37) Vogtherr, M.; Fiebig, K. NMR-based screening methods for lead discovery. *EXS* **2003**, *93*, 183−202.
(38) Zabell, A. P. Z.; Post, C. B. Docking multiple conformations of a flexible ligand into a protein binding site using NMR restraints. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 295−307.
(39) Hajduk, P. J.; Mack, J. C.; Olejniczak, E. T.; Park, C.; Dandliker, P. J.; Beutel, B. A. SOS−NMR: A saturation transfer NMR-based method for determining the structures of protein−ligand complexes. *J. Am. Chem. Soc.* **2004**, *126*, 2390−2398.
(40) Hu, X.; Balaz, S.; Shelver, W. H. A practical approach to docking of zinc metalloproteinase inhibitors. *J. Mol. Graphics Modell.* **2004**, *22*, 293−307.
(41) Moy, F. J.; Pisano, M. R.; Chanda, P. K.; Urbano, C.; Killar, L. M.; Sung, M. L.; Powers, R. Assesignments, secondary structure and dynamics of the inhibitor-free catalytic fragment of human fibroblast collagenase. *J. Biomol. NMR* **1997**, *10*, 9−19.
(42) Moy, F. J.; Chanda, P. K.; Chen, J. M.; Cosmi, S.; Edris, W.; Skotnicki, J. S.; Wilhelm, J.; Powers, R. NMR Solution Structure of the Catalytic Fragment of Human Fibroblast Collagenase Complexed with a Sulfonamide Derivative of a Hydroxamic Acid Compound. *Biochemistry* **1999**, *38*, 7085−7096.
(43) Wasserman, Z. R. Making a New Turn in Matrix Metalloprotease Inhibition. *Chem. Biol.* **2005**, *12*, 143−148.
(44) Clore, G. M. Accurate and rapid docking of protein−protein complexes on the basis of intermolecular nuclear overhauser enhancement data and dipolar couplings by rigid body minimization. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9021−9025.
(45) Clore, G. M.; Bewley, C. A. Using Conjoined Rigid Body/Torsion Angle Simulated Annealing to Determine the Relative Orientation of Covalently Linked Protein Domains from Dipolar Couplings. *J. Magn. Reson.* **2002**, *154*, 329−335.
(46) Jain, N. U.; Wyckoff, T. J. O.; Raetz, C. R. H.; Prestegard, J. H. Rapid Analysis of Large Protein−Protein Complexes Using NMR-derived Orientational Constraints: The 95 kDa Complex of LpxA with Acyl Carrier Protein. *J. Mol. Biol.* **2004**, *343*, 1379−1389.
(47) Crowley, P. B.; Otting, G.; Schlarb-Ridley, B.; Canters, G. W.; Ubbink, M. Hydrophobic interactions in a cyanobacterial plastocyanin-cytochrome *f* complex. *J. Am. Chem. Soc.* **2001**, *123*, 10444−10453.
(48) Banci, L.; Bertini, I.; Cavallaro, G.; Giachetti, A.; Luchinat, C.; Parigi, G. Paramagnetism-based restraints for Xplor-NIH. *J. Biomol. NMR* **2004**, *28*, 249−261.
(49) Palmer, A. G., III; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Magn. Reson.* **1991**, *93*, 151−170.
(50) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S. J.; Sørensen, O. W.; Griesinger, C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR* **1994**, *4*, 301−306.
(51) Bartels, C.; Xia, T. H.; Billeter, M.; Güntert, P.; Wüthrich, K. The program XEASY for computer-supported NMR. *J. Biomol. NMR* **1995**, *5*, 1−10.
(52) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(53) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(54) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity-A Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219−3228.
(55) Schwieters, C. D.; Kuszewski, J.; Tjandra, N.; Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **2003**, *160*, 65−73.

(56) Schwieters, C. D.; Clore, G. M. Internal Coordinates for Molecular Dynamics and Minimization in Structure Determination and Refinement. *J. Magn. Reson.* **2001**, *152*, 288−302.
(57) Kleywegt, G. J.; Jones, T. A. *Methods Enzymol.* **1997**, *277*, 208−230.
(58) MacKerell, A. D., Jr.; Wiòrkiewicz-Kuczera, J.; Karplus, M. An All-Atom Empirical Energy Function for the simulation of Nucleic Acids. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

JM050574K