Mechanism of the Serine Protease Reaction. Stereoelectronic, Structural, and Kinetic Considerations as Guidelines To Deduce Reaction Paths

HANS DUTLER* and SPARTACO A. BIZZOZERO

Laboratorium für Organische Chemie, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

Received October 4, 1988 (Revised Manuscript Received May 30, 1989)

Introduction

Understanding the mechanisms by which enzymes catalyze chemical reactions is a fundamental problem in theoretical chemistry as well as in biochemistry. Valuable mechanistic information may, in appropriate cases, be derived from detailed knowledge of the structure of reacting species along the reaction path. By extrapolating from these ground-state structures it should eventually be possible to obtain estimates of transition-state structures for the individual bond-breaking and bond-making processes. Since the reacting species represent transients that are not directly accessible by experiment, their structures need to be deduced by starting from appropriate structural and kinetic information. This is usually done by molecular modeling using molecular mechanics or molecular dynamics procedures. The starting information for these procedures is of two kinds: (1) Structure of enzyme–inhibitor complexes and of enzyme derivatives. Obviously, these stable enzyme species are not themselves catalytically competent, but have the capacity to serve as reference species from which the deduction procedure can start. (2) Kinetics describing structure–function relationships. To be applicable in the molecular modeling procedures, this kinetic information must be expressed in terms of interactions occurring between protein and substrate. However, deducing the structure of reacting species on the basis of exploiting these two kinds of information turned out to be a very difficult endeavor even when highly sophisticated computer-aided methods were used. On the basis of our work concerned with understanding enzyme catalysis from the stereochemical point of view, we developed a new design tool that is expected to greatly facilitate this deduction procedure and that is particularly well suited to combine with the experimentally available structural information.

The new design tool is developed from generalization about the preferred angles of attack by nucleophiles on carbonyl carbon and about preferred conformations of the reacting species. Such stereoelectronic concepts have been long known in organic chemistry but have found only reluctant acceptance in biochemistry. We now want to make use of these concepts by way of exploiting their properties to express intrinsic reaction requirements and to design a specific intrinsic reaction module (IRM) for each of the considered reacting species. These IRMs are incorporated into the modeled structures as individual structural entities belonging partly to the enzyme and partly to the substrate.

With trypsin as model enzyme, this Account describes how IRMs are used in combination with information on protein–substrate interactions to deduce the structure of the four reacting species on the reaction path to the acyl enzyme. With the aim of obtaining insight into the molecular events taking place during reaction, the resulting changes of the IRMs are related to the conformational adaptability of the protein.

The Reaction in Summary

The structure of peptide substrates is formulated as shown in Figure 1.[1] The amino acid residues are labeled L_{x1}, L_{x2},... on the N-terminal side and L_{y1}, L_{y2},... on the C-terminal side. The substrate is cleaved between L_{x1} and L_{y1}. A tetrapeptide, for example, is composed of the four amino acid residues L_{x1}, L_{x2}, L_{y1}, and L_{y2}, and the corresponding side chains are labeled R_{x1}, R_{x2}, R_{y1}, and R_{y2}, respectively. Individual atoms obtain the index of the amino acid residue they belong to: for example, C and N of the peptide bond to be cleaved and C^{α} of L_{x1} are formulated as C_{x1}, N_{y1}, and C^{α}_{x1}, respectively. In shortened form and for all peptide substrates equally, X–CO– is used for the N-terminal acyl part (X part) and –NH–Y for the C-terminal leaving-group part (Y part).

The structure of trypsin is known to high precision,[2-5] there exists a wealth of kinetic information,[6-9] and it has a well-defined mechanism.[2,10-15] The trypsin-cat-

Hans Dutler was born in Switzerland in 1929. He received his Ph.D. from ETH Zürich in 1957, and during 1957–1960 he was a postdoctoral fellow in synthetic organic chemistry with Robert B. Woodward at Harvard University. In 1961 he joined the Laboratorium für Organische Chemie at the ETH, where he received the title of Professor in 1981. His research interests are in the chemistry of enzymatic reactions.

Spartaco A. Bizzozero was born in 1942 in Switzerland. He received his Ph.D. from the ETH Zürich in 1971 and is now Oberassistent at the Laboratorium für Organische Chemie at the ETH. His research interests are in the mechanism of serine protease reactions.

(1) Baumann, W. K.; Bizzozero, S. A.; Dutler, H. *Eur. J. Biochem.* **1973**, *39*, 381–391.
(2) Huber, R.; Kukla, D.; Bode, W.; Schwager, P.; Bartels, K.; Dreisenhofer, J.; Steigemann, W. *J. Mol. Biol.* **1974**, *89*, 73–101.
(3) Bode, W.; Schwager, P. *J. Mol. Biol.* **1975**, *98*, 693–717.
(4) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Crystallogr.* **1983**, *B39*, 480–490.
(5) Chambers, J. L.; Stroud, R. M. *Acta Crystallogr.* **1979**, *B35*, 1861–1874.
(6) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331–358.
(7) Pozsgay, M.; Szabo, G.; Bajusz, S.; Simonsson, R.; Caspar, R.; Elodi, P. *Eur. J. Biochem.* **1981**, *115*, 497–502.
(8) Bizzozero, S. A.; Dutler, H. *Arch. Biochem. Biophys.* **1987**, *256*, 662–676.
(9) Fiedler, F. *Eur. J. Biochem.* **1987**, *163*, 303–312.
(10) Bode, W.; Huber, R. In *Molecular and Cellular Basis of Digestion*; Desnuelle, P., Ed.; Elsevier: Amsterdam, 1986; pp 213–234.
(11) Kossiakoff, A. A.; Spencer, S. A. *Nature (London)* **1980**, *288*, 414–416.
(12) Umeyama, H.; Hirono, S.; Nakagawa, S. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6266–6270.
(13) Weiner, S. J.; Seibel, G. L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 649–653.

0001-4842/89/0122-0322$01.50/0 © 1989 American Chemical Society

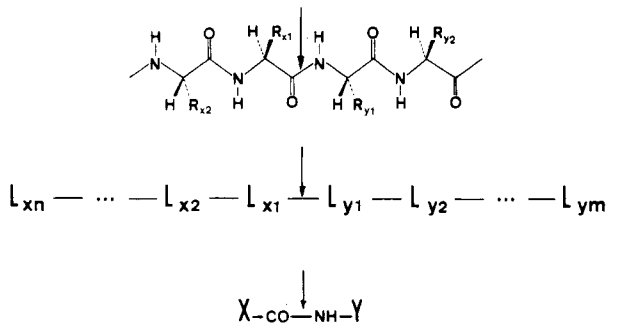

Figure 1. Peptide substrate nomenclature. The arrows indicate the bond to be cleaved.

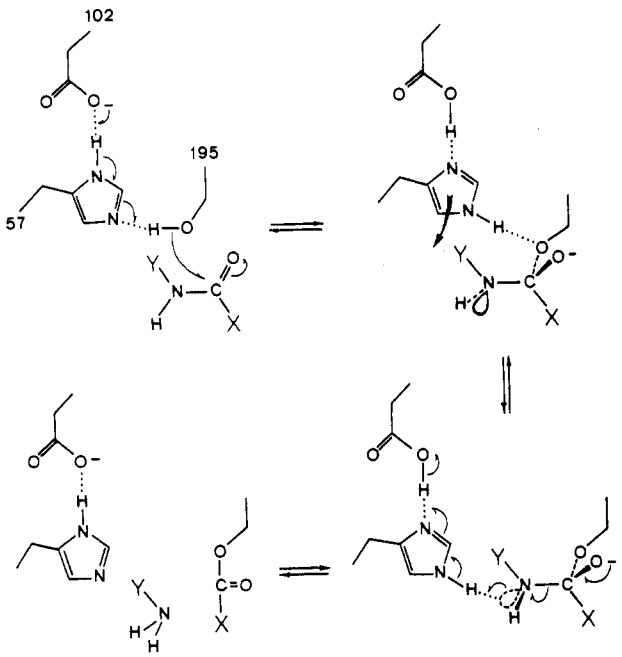

Figure 2. Schematic diagram showing the molecular events taking place during reaction. Proton transfer is assumed to involve the full catalytic triad. The two processes accompanying interconversion of the two tetrahedral intermediates are (1) inversion at N_{y1} and (2) in–out movement of imidazole (indicated by curved arrow).

alyzed reactions can be divided into two acyl transfer steps (Figure 2). In the first acyl transfer, X–CO– is transferred from the peptide to O^{γ}_{195}, and H–NH–Y of the cleaved peptide substrate is released. In the second acyl transfer, X–CO– is transferred from the acyl enzyme to water, forming free enzyme and X–COOH. Taking into account that the enzyme makes use of the catalytic triad Ser-195 (O^{γ}_{195}), His-57 ($N^{\epsilon 2}_{57}$ and $N^{\delta 1}_{57}$), and Asp-102 (one of the O^{δ}_{102}),[16-18] nucleophilic attack by O^{γ}_{195} on the carbonyl is accompanied by the transfer of the proton from O^{γ}_{195} to $N^{\epsilon 2}_{57}$. Whether the proton at $N^{\delta 1}_{57}$ is transferred to one of the O^{δ}_{102} (as assumed in Figure 2) or whether charge interaction between the two parts occurs remains an open question and is not of importance for the arguments presented in this Account. With the associated decomposition of the thus formed tetrahedral intermediate, proton displacement takes place in reverse, with the proton being transferred from $N^{\epsilon 2}_{57}$ to the N_{y1}.

As discussed in detail previously,[19,20] Delongchamps's stereoelectronic theory[21-23] requires that the lone-pair orbital at N_{y1} assumes an antiperiplanar position with respect to the newly forming O–C bond, thereby directing the proton located at N_{y1} toward $N^{\epsilon 2}_{57}$. Since $N^{\epsilon 2}_{57}$ is occupied by the proton it had accepted from O^{γ}_{195}, this requirement results in two-proton interference. In the next step, the proton at $N^{\epsilon 2}_{57}$ should be transferred to N_{y1}, a process that cannot directly take place due to said orientation of the proton at N_{y1}. Both these structural relationships, therefore, require that N_{y1} undergoes inversion. As we had also pointed out previously,[20] since $N^{\epsilon 2}_{57}$ is too far from N_{y1} for the proton to be transferred, the imidazole must move from a position in which it is hydrogen bonded to O^{γ}_{195} to one in which it is hydrogen bonded to the N_{y1}. Thus acyl enzyme formation involves two tetrahedral intermediates whose interconversion is accompanied by two complementary processes, (1) inversion of N_{y1} and (2) movement of imidazole from an in-position into an out-position.

Our experiments with chymotrypsin using *N*-acetylphenylalanylsarcosine amide as substrate-analogue inhibitor[19] provided a first confirmation of the applicability of this mechanistic scheme: the methyl group replacing the proton at the N_{y1} comes too close to the imidazole ring, thereby preventing formation of the tetrahedral intermediate. Crystallographic and NMR experiments with porcine pancreatic elastase using the hexapeptide Thr-Pro-nVal-NMeLeu-Tyr-Thr[24] show in convincing structural detail how such an *N*-methyl group interferes with the imidazole. In both cases, the reaction is blocked at a stage close to the starting enzyme–substrate complex.

Intrinsic Reaction Modules

The stereoelectronic concepts used for designing the IRMs are based on three kinds of effects: (1) Primary electronic effects as a result of n–π^* interactions. These effects are due to the delocalization of electron pairs between amide nitrogen or ester oxygen and the carbonyl group forming the conjugated system of the amide or ester function, respectively. (2) Secondary electronic effects as a result of n–σ^* interactions. According to Deslongchamps,[21-23] these effects are due to the preferred orientation of the specific σ C–N or σ C–O bond of the tetrahedral intermediate antiperiplanar to the lone-pair orbitals on the other two heteroatoms. (3) Directionality in the approach of a nucleophile, Nu, on carbonyl and in the ejection of a leaving group from the tetrahedral intermediate. According to Bürgi and Dunitz,[25-27] preferred approach occurs at a Nu···C=O

(14) Fink, A. L. In *Enzyme Mechanisms*; Page, M. I., Williams, A., Eds.; The Royal Society of Chemistry: London, 1987; pp 159–177.
(15) Kossiakoff, A. A. In *Biological Macromolecules and Assemblies*; Jurnak, F. A., McPherson, A., Eds.; Wiley: New York, 1987; Vol. 3, pp 370–412.
(16) Elrod, J. P.; Hogg, J. L.; Quinn, D. M.; Venkatasubban, K. S.; Schowen, R. L. *J. Am. Chem. Soc.* **1980**, *102*, 3917–3922.
(17) Lesyng, B.; Meyer, E. F. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 211–217.
(18) Meyer, E.; Cole, G.; Radhakrishan, R. *Acta Crystallogr.* **1988**, *B44*, 26–38.
(19) Bizzozero, S. A.; Zweifel, O. *FEBS Lett.* **1975**, *59*, 105–108.
(20) Bizzozero, S. A.; Dutler, H. *Bioorg. Chem.* **1981**, *10*, 46–62.
(21) Deslongchamps, P. *Tetrahedron* **1975**, *31*, 2463–2490.
(22) Deslongchamps, P. *Heterocycles* **1977**, *7*, 1271–1317.
(23) Deslongchamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon Press: Oxford, 1983.
(24) Meyer, E. F.; Clore, G. M.; Gronenborn, A. M.; Hansen, H. A. *Biochemistry* **1988**, *27*, 725–730.
(25) Bürgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc.* **1973**, *95*, 5065–5067.

angle of 100–110° and in the π-orbital plane of the carbonyl group. In the reverse molecular event, i.e., ejection of a leaving group, the corresponding directionality is thought to be applicable.

The way we want to make use of these concepts consists in producing an IRM for each of the considered species along the reaction path. These structural entities include two types of atoms, which may be part of substrate or protein. The atoms that either are directly involved in bond breaking and making and/or undergo hybridization changes are called *reacting atoms*. They constitute the actual "transposers" of the electronic effects, and their function along the entire reaction path is independent of the species considered. The reacting-atom structure of an IRM is based on the σ-bond skeleton and made in accord with the above-mentioned electronic effects. Deviations in bond and torsion angles due to the presence of heteroatoms are not considered at the present stage.[28] The thus-resulting topographical features are propagated to the residual part of the considered species via atoms which are "slaved" by the reacting atoms and are called *satellite atoms*. Such atoms are fixed in their position within the IRM whenever their neighboring reacting atoms (reacting atoms to which they are bound) are constrained due to primary and/or secondary electronic effects. Since these constraints vary from species to species, the IRMs may contain different satellite atoms.

The structural linkage between the IRM and the residual part of the considered species is primarily affected by the spatial arrangement of the atoms located at the periphery of the IRM (reacting or satellite atoms). This positional determinant combines with the directional determinant defined by the bonds leading from the IRM to the connecting atoms on the side of the residual part of the species. The bonds to the connecting atoms are conformationally restricted as dictated by the IRM and are contained in a conic surface with its peak at the corresponding atom located at the periphery of the IRM. In each IRM, such conic surfaces are visualized by circles approximating the geometric loci of the connecting atoms. Depending on the constraints effective at the neighboring reacting atom (as stated above), an atom functioning as a satellite atom in one IRM may appear as a connecting atom in another. The IRMs concerning the species of the first acyl transfer are depicted in Figure 3 (the species of the second acyl transfer are structurally related). They are all projected in the O_{x1}–C_{x1} direction. This unifying orientation takes into account that the C_{x1}–O_{x1} bond is common to both tetrahedral intermediates and defines the direction of attack of the nucleophile and ejection of the leaving group.

In the IRM for the starting species of the process (IRM 1), O^{γ}_{195} is acting as nucleophile to attack the carbonyl from the *Re* side. The reacting atoms C_{x1}, O_{x1}, and N_{y1} together with the satellite atoms C^{α}_{x1} and C^{α}_{y1} are located in one plane, the plane of the peptide bond. O^{γ}_{195} is at a distance of 2.8 Å from C_{x1}, and the O^{γ}_{195}···C=O angle is set at 107°. The chosen O^{γ}_{195}–C_{x1} distance is supposed to relate to a displacement of C_{x1} from the plane of its three-bonded atoms N_{y1}, C^{α}_{x1}, and O_{x1} of about 10% of the distance observed in the tetrahedral intermediates (distance at which directionality of bond making starts to be operative). C^{β}_{195} is not intrinsically fixed and therefore appears as a connecting atom. The geometric loci of the connecting atoms are visualized by circles around C^{α}_{x1}, C^{α}_{y1}, and O^{γ}_{195}.

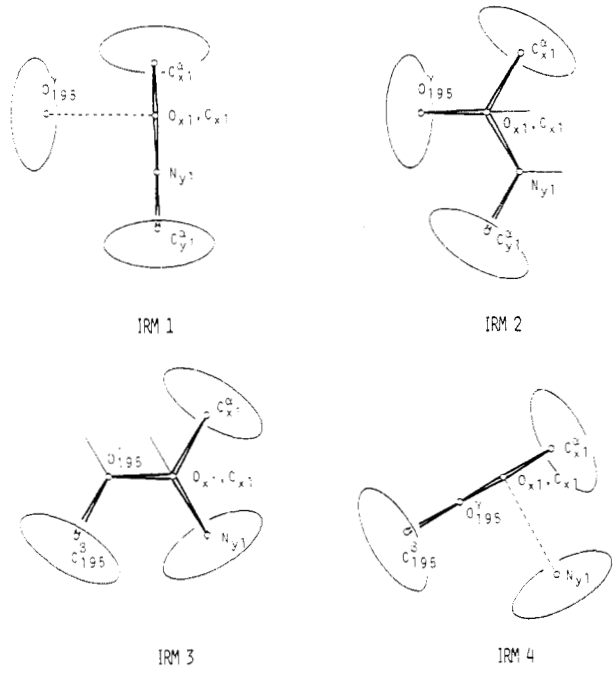

Figure 3. Intrinsic reaction modules (IRMs) projected along the O_{x1}–C_{x1} axis. Open-ended lines signify lone pair orbital axes, and dotted lines signify directions of nucleophile attack or leaving-atom ejection. The circular geometric loci of the connecting atoms appear as ellipses at the corresponding neighboring atoms.

In the IRM for the first tetrahedral intermediate (IRM 2), the reacting atoms show tetrahedral geometry. This geometry in principle is dictated by the antiperiplanar arrangement of a lone-pair orbital at O_{x1} and N_{y1} relative to the C_{x1}–O^{γ}_{195} bond and at O_{x1} and O^{γ}_{195} relative to the C_{x1}–N_{y1} bond (secondary electronic effects). As a result of these requirements, all reacting and satellite atoms have a tendency to adopt the geometry of a diamond lattice with slightly altered bond lengths due to the presence of the heteroatoms. Such geometry, however, would lead to severe steric congestion at C^{β}_{195} and C^{α}_{y1}. Since the antiperiplanar arrangement of the critical lone-pair orbital at N_{y1} and O_{x1} relative to the newly formed C_{x1}–O^{γ}_{195} bond also plays a role in lowering the energy of the transition state of nucleophilic attack, this requirement is the more important one, and C^{β}_{195} is thought to release the strain by rotation about the C_{x1}–O^{γ}_{195} bond. Thus, as in IRM 1, C^{β}_{195} is not intrinsically fixed and appears as a connecting atom. The connecting atoms are the same as in IRM 1.

In the IRM for the second tetrahedral intermediate (IRM 3), the reacting and satellite atoms tend to adopt the strained geometry of the diamond lattice for the same reasons as in IRM 2. But now C^{α}_{y1} is more likely to release the strain since the antiperiplanar arrangement of the critical lone-pair orbitals at O^{γ}_{195} and O_{x1} relative to the breaking C_{x1}–N_{y1} bond also helps to lower the energy of the transition state of bond breaking. In analogy to IRM 2, therefore, C^{α}_{y1} is not included in the IRM and appears as a connecting atom. The geometric

(26) Bürgi, H. B.; Dunitz, J. B.; Wipff, G. *Tetrahedron* **1974**, *30*, 1563–1572.
(27) Schweizer, W. B.; Procter, G.; Kaftory, M.; Dunitz, J. D. *Helv. Chim. Acta* **1978**, *61*, 2783–2808.
(28) Lehn, J.-M.; Wipff, G. *J. Am. Chem. Soc.* **1980**, *102*, 1347–1354.

loci of the connecting atoms are visualized by circles around C^{α}_{x1}, C^{β}_{195}, and N_{y1}.

In the IRM for the final species, the product of the process during which the C_{x1}–N_{y1} bond is broken (IRM 4), the reacting atoms C_{x1}, O_{x1}, and O^{γ}_{195} together with satellite atoms C^{α}_{x1} and C^{β}_{195} are located in one plane, the plane of the ester bond. N_{y1} is at a distance of 2.8 Å from C_{x1}, and the N_{y1}···C=O angle is set at 107°. This geometry is chosen in analogy to IRM 1. C^{α}_{y1} is not intrinsically fixed and therefore, as in IRM 3, appears as a connecting atom. The connecting atoms are the same as in IRM 3.

General Modeling Strategy

For deriving the structure of the considered reacting species, the information contained in the IRMs is to be combined with the structural information available from appropriate reference species. It is evident that such a combination of information can only be achieved when the system is subjected to a conformational adaptation procedure. Direct mutual adaptation of both informational components, however, is very difficult to achieve, particularly as dependable information on conformational energy changes concerning the protein backbone and the IRM is scarce. We circumvent this problem by application of a stepwise procedure in which geometric constraints are given a leading function.

A first step aims at structures that are characterized by a bias in favor of either the IRM or the protein. In both these biased structures, the protein backbone and the IRM bond angles remain unchanged. *IRM-biased structures* are obtained by adapting the protein to fully rigid IRMs. These structures have the valuable property of disclosing where and in which direction during adaptation the protein has to change its conformation as dictated by the IRM. *Protein-biased structures* are the result of adapting the protein to fully flexible IRMs. They disclose where and how far IRM torsion angles need to be changed as dictated by the protein with fixed backbone. The two types of biased structures mark the topographic limits within which protein conformational changes can take place when the structure-forming capacity of the IRMs is being balanced with that of the protein. In a second step, mutual adaptation of IRM and protein leads to *balanced structures*. This final modeling step can thus be performed within a clearly defined topographic framework. In these balanced structures, the protein backbone and IRM bond angles are both allowed to undergo controlled changes.

As a critical measure for the functional competence of the modeled structures, we use kinetically derived information on protein–substrate interactions. This information is topographic in nature and consists in the detailed specification of the kinetics-based nonbonded interaction pairs. Modeling then is performed on the basis of the criterion that kinetics-based and model-created interactions should conform as far as possible. This interaction criterion becomes particularly important when the protein backbone and the IRM bond angles are allowed to change. Methods and data concerning the kinetic investigations with substrates of systematically varying structure and the derivation of the information on the protein–substrate interactions are published elsewhere.[1,20,29-32]

(29) Bizzozero, S. A.; Kaiser, A. W.; Dutler, H. *Eur. J. Biochem.* **1973**, *33*, 292–300.

Table I
Relevant Torsion Angles and Distances Pertaining to the Two Tetrahedral Intermediates as Obtained by Reaction of Ac-Arg-Ala-NH_2 with Trypsin

geometric parameters	bias	1st tetrahedral intermediate (substructure 2)	2nd tetrahedral intermediate (substructure 3)
C_{x1}–O^{γ}_{195}–C^{β}_{195}–C^{α}_{195}	IRM	+110°	+110°
	protein	+110°	+107°
N_{y1}–C_{x1}–O^{γ}_{195}–C^{β}_{195}	IRM	+90°	+60°
	protein	+77°	+68° [a]
O^{γ}_{195}–C_{x1}–N_{y1}–C^{α}_{y1}	IRM	−60°	−90°
	protein	−70° [a]	−89°
C^{β}_{195}–C^{α}_{y1} distance	IRM	2.8 Å	2.8 Å
	protein	2.8 Å	2.9 Å

[a] The critical IRM torsion angles of the protein-biased structures.

The BPTI (bovine pancreatic trypsin inhibitor) complex of trypsin[2] is chosen as the appropriate reference species for the starting species and the first tetrahedral intermediate. It fixes the side chain of His-57 with the imidazole in the in-position. Fixation of the side chain of His-57 with the imidazole in the out-position for the second tetrahedral intermediate and for the final species is achieved on the basis of the structure of diisopropyl fluorophosphate inhibited (DIP-inhibited) trypsin.[5] To start with, the IRMs and the residual parts of the substrates are built into the structures of the respective reference species from which all nonprotein atoms had previously been removed. Assembling these components is achieved with the aid of the connecting atoms. The models were generated on an Evans and Sutherland PS 300 instrument using the program FRODO.[33] The program YETI[34] served for adaptation of the biased structures and the program AMBER[35] for adaptation of the balanced structures.

For the purpose of demonstrating the essentials concerning the modeling procedure and the information that can be gained, the considered system is extensively simplified. The main arguments concerning the geometric constraints can be discussed with a focus on the near vicinity of the IRMs. The thus-considered parts of the modeled structures are composed of the side chains of Ser-195 and His-57 representing the protein and C^{α}_{x1} and C^{α}_{y1} representing the substrate. Two types of torsion angles are differentiated: IRM torsion angles with all four atoms belonging to the IRM (e.g., O^{γ}_{195}–C_{x1}–N_{y1}–C^{α}_{y1} in the first tetrahedral intermediate) and torsion angles with less than four atoms belonging to the IRM (e.g., N_{y1}–C_{x1}–O^{γ}_{195}–C^{β}_{195}in the first tetrahedral intermediate). The two types of torsion angles are used to describe the deviation from ideal diamond-lattice geometry with the aid of Newman projections. Among the protein–substrate interactions, the repulsive interaction consisting in steric congestion at C^{β}_{195} and C^{α}_{y1} is chosen to illustrate the modeling procedure. Substructures 2 and 3 representing, respectively, the first and the second tetrahedral intermediate with Ac-Arg-Ala-NH_2 as the substrate are discussed as examples.

(30) Bizzozero, S. A.; Baumann, W. K.; Dutler, H. *Eur. J. Biochem.* **1975**, *58*, 167–176.
(31) Bizzozero, S. A.; Rovagnati, B. A.; Dutler, H. *Helv. Chim. Acta* **1982**, *65*, 1707–1719.
(32) Bizzozero, S. A.; Baummann, W. K.; Dutler, H. *Eur. J. Biochem.* **1982**, *122*, 251–258.
(33) Jones, T. A. In *Computational Crystallography*; Sayre, D., Ed.; Oxford University Press: New York, 1982; pp 303–317.
(34) Vedani, A. *J. Comput. Chem.* **1988**, *9*, 269–280.
(35) Weiner, P.; Kollman, P. *J. Comput. Chem.* **1981**, *2*, 287–303.

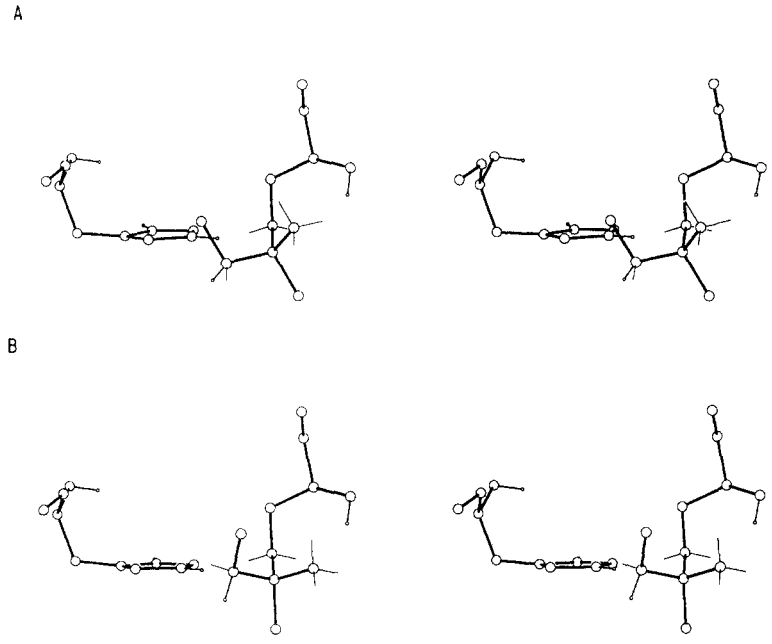

Figure 4. Stereodiagrams of substructures 2 (A) and 3 (B). Open-ended lines signify orbital axes, and small circles signify hydrogen.

The resulting geometric parameters are summarized in Table I.

Modeled Structures

The *IRM-biased structures* allow full expression of the structure-forming capacity of the IRMs and hence are particularly useful to demonstrate the main lines of thought. The two considered substructures, both viewed in the same direction, are shown in Figure 4. The relevant conformational changes accompanying the adaptation processes are visualized in the Newman projections of Figure 5. In substructure 2, the steric congestion at C^{β}_{195} and C^{α}_{y1} is released by rotation about the C_{x1}–O^{γ}_{195} axis of approximately +30° away from ideal diamond-lattice geometry (Figure 5A). The resulting conformation (Table I) has the following consequence: it leads to a reduction of steric congestion of the proton at N_{y1} (intrinsically forced to appear on the side facing the imidazole) with $N^{\epsilon 2}_{57}$ of imidazole in the in-position and with the proton being transferred from O^{γ}_{195} to $N^{\epsilon 2}_{57}$ (Figure 4A). In substructure 3, the steric congestion at C^{β}_{195} and C^{α}_{y1} is responsible for rotation about the C_{x1}–N_{y1} axis of approximately −30° away from ideal diamond-lattice geometry (Figure 5B). This conformational change (Table I) has the consequence that the lone-pair orbital at N_{y1} (after inversion) now points toward $N^{\epsilon 2}_{57}$ of imidazole in the out-position, as required for proton transfer from $N^{\epsilon 2}_{57}$ to N_{y1} (Figure 4B).

The *protein-biased structures* are characterized by fully flexible IRMs, and hence the conformational changes involve rotations about both the C_{x1}–O^{γ}_{195} and the C_{x1}–N_{y1} axes. In substructure 2, the steric congestion at C^{β}_{195} and C^{α}_{y1} results in rotations about the two axes away from ideal diamond lattice geometry of, respectively, +17° and −10°. Interestingly, the smaller rotation concerns the IRM torsion angle O^{γ}_{195}–C_{x1}–N_{y1}–C^{α}_{y1} (Table I). In substructure 3, the corresponding rotations are, respectively, +8° and −29°. Here the difference in the extent of rotation is somewhat more pronounced. It is again the rotation concerning the IRM torsion angle that is the smaller one; in this structure, the IRM torsion angle is N_{y1}–C_{x1}–O^{γ}_{195}–C^{β}_{195} (Table I).

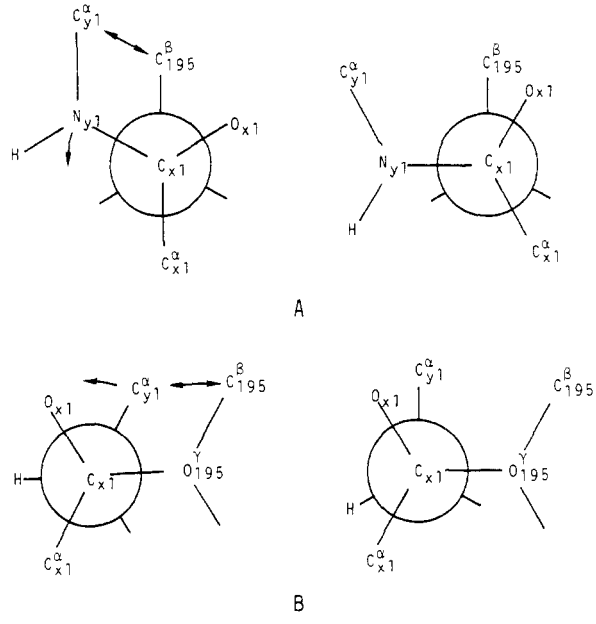

Figure 5. Newman projections of substructures 2 and 3. On the left are shown the conformations with ideal diamond-lattice geometry and on the right those with the geometry of the IRM-biased structure. (A) Substructure 2 (first tetrahedral intermediate) projected along the C_{x1}–O^{γ}_{195} axis. (B) Substructure 3 (second tetrahedral intermediate) projected along the C_{x1}–N_{y1} axis. The critical steric congestion is indicated by the double-headed arrows, and its decrease by rotation about the pertinent axis is indicated by the single-headed arrows. Open-ended lines signify lone pair orbital axes.

The final step leading to the *balanced structures* involves release of the conformational restriction caused by the fixed IRM bond angles and the rigid protein backbone. Controlled changes in these two structural features provide the conformational freedom necessary for further optimizing the protein–substrate interactions. Both considered reacting species show an astonishingly small difference in the IRM- and protein-biased structures with respect to the two considered

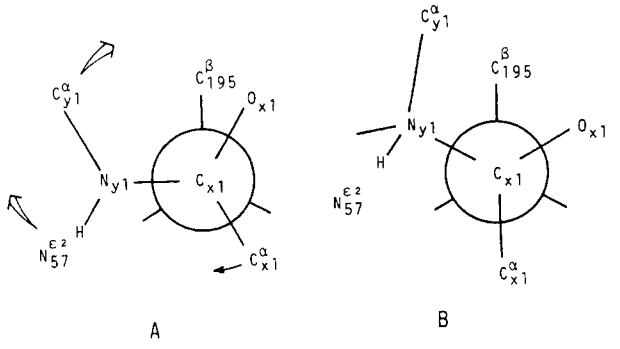

Figure 6. Analysis of the conformational changes occurring upon proceeding from substructure 2 (A) to substructure 3 (B). The Newman projections along the C_{x1}–O^{γ}_{195} axis in addition contain $N^{\epsilon 2}_{57}$ of imidazole in the respective in- and out-positions. The arrows indicate the movements occurring during transformation. Open-ended lines signify lone pair orbital axes.

torsion angles (Table I). These geometric parameters, therefore, are expected to undergo small changes only; their main function in the process is to mark the limits within which the conformational changes can take place. Modeling of these balanced structures is currently under investigation.

Conformational Changes Occurring during Reaction

The IRM-biased structures are used to obtain first insight into the molecular events taking place during reaction. The step proceeding from the first to the second tetrahedral intermediate is chosen as an example to demonstrate how to perform this comparative conformational analysis. The envisaged characterization of movements is best carried out by using Newman projections along the C_{x1}–O^{γ}_{195} axis showing rotation about the C_{x1}–N_{y1} axis in the near side view. Since the changes in the torsion angle C_{x1}–O^{γ}_{195}–C^{β}_{195}–C^{α}_{195} (Table I) are negligible, the in–out movement of imidazole can be taken care of by including in the Newman projections the corresponding $N^{\epsilon 2}_{57}$ positions. The analysis of the movements then comprises C^{α}_{x1}, C^{α}_{y1}, and $N^{\epsilon 2}_{57}$ and is expressed in terms of rotations and translations relative to the C_{x1}–O^{γ}_{195} axis with a fixed O^{γ}_{195}–C^{β}_{195} axis.

The two Newman projections in Figure 6 reveal the following movements upon proceeding from substructure 2 (Figure 6A) to substructure 3 (Figure 6B): (1) C^{α}_{x1} rotates clockwise in a plane perpendicular to the C_{x1}–O^{γ}_{195} axis; (2) C^{α}_{y1} rotates clockwise and translates antiparallel to the C_{x1}–O^{γ}_{195} axis and radially toward this axis; (3) $N^{\epsilon 2}_{57}$ rotates clockwise and translates antiparallel to the C_{x1}–O^{γ}_{195} axis (from the in- to the out-position) and radially away from this axis. With respect to the main concern of this conformational analysis, namely, to obtain information on how the protein framework conforms to the movements of C^{α}_{x1} and C^{α}_{y1}, this reaction step chosen as an example turns out to be a particularly favorable one. The derived movements suggest that there might exist a coupling between the movement of $N^{\epsilon 2}_{57}$ on the side of the protein and either C^{α}_{x1} or C^{α}_{y1} (or both) on the side of the substrate. Imidazole is incorporated in a hydrogen-bonding system reaching from the amido N of Ala-56 and His-57 on one end, across Asp-102, to the O^{γ} of Ser-214 on the other end. Interestingly, Ser-214 is among the amino acids essentially responsible for the interactions with the X part of the substrate, forming together with Trp-215 and Gly-216 a section of a β-pleated sheet.[8] This chain of bonds and interactions well supports the suggestion that the movement of imidazole is coupled with the movement of the satellite atom C^{α}_{x1}.

Concluding Remarks

The modeled structures obtained by adapting the protein to rigid and flexible IRMs (IRM-biased and protein-biased structures, respectively) show a small difference in the geometric parameters considered (primarily two types of torsion angles); this result applies to both tetrahedral intermediates. Apparently the protein "knows" about the stereoelectronic effects and orients the substrate so as to optimally obey the intrinsic reaction requirements inherent in the IRMs. Comparison of the two tetrahedral intermediates reveals that their interconversion involves significant specific changes in the geometric parameters considered. The source of this phenomenon appears to reside in the occurrence of strain (e.g., steric congestion at C^{β}_{195} and C^{α}_{y1}) as a consequence of obeying the intrinsic reaction requirements. In principle, the protein could compensate for this strain by better binding. Judged from our modeling results this is not the case; rather, the protein undergoes IRM-dictated conformational changes. We think that these marked structural and reaction-dynamic features will lead to new aspects in designing inhibitors. They should also contribute to the development of new viewpoints concerning the evolution of serine proteases.

We gratefully acknowledge Peter Zbinden, who actively participated in producing the modeled structures. It is a pleasure for both of us to thank the Swiss National Science Foundation for generous financial support over the years.

Registry No. Serine proteinase, 37259-58-8; trypsin, 9002-07-7.