Human Rhinovirus 3C Protease: Generation of Pharmacophore Models for Peptidic and Nonpeptidic Inhibitors and Their Application in Virtual Screening

Theodora Steindl, Christian Laggner, and Thierry Langer*

Institute of Pharmacy, Computer Aided Molecular Design Group, University of Innsbruck, Innrain 52, A-6020 Innsbruck, Austria

Received December 3, 2004

Three-dimensional pharmacophore models for peptidic and small organic nonpeptidic inhibitors of the human rhinovirus 3C protease were generated in a structure-based as well as in a ligand-based approach, using the software package Catalyst. The inhibitors possess an electrophilic moiety, often a Michael acceptor function, which covalently binds to a cysteine in the active site of the enzyme. Since this process presents the key step for virus inactivation, the creation of a new function in Catalyst was required in order to include this decisive functionality into the pharmacophore models. In the present study we focus on this feature definition process because it presents an innovative strategy to expand the pharmacophore description ability of the Catalyst software to also include covalent bonds between ligand and binding site. The resulting hypotheses were then used for virtual screening of 3D databases in order to verify their quality and to search for structurally diverse, possible new lead substances.

INTRODUCTION

The human rhinovirus (HRV), a member of the picornavirus family, is the major cause of common cold. So far, only symptomatic treatment is available to medicate rhinovirus infections. In this context, the HRV 3C protease (3CP) presents an intensively studied target for the development of antirhinitis drugs.

Picornaviruses have a single-stranded, positive-sense RNA genome, which is translated into a polyprotein precursor. 3C and 2A proteases process the latter in order to generate mature viral proteins and functional enzymes and therefore ensure viral replication.[1] Especially the HRV 3CP is considered a very promising drug target, as can be concluded from intensive research on this topic throughout the last years.[2] Although more than 100 HRV serotypes have been identified, the highly conserved binding site[3] of the 3CPs indicates that a single compound could be active against a multitude of different serotypes. Besides, the lack of homology with known mammalian enzymes renders this an advantageous target.[4] Recently, HRV 3CP has also received increased interest because of its structural similarity to SARS-associated coronavirus main protease and the assumption that available rhinovirus 3CP inhibitors may be modified to make them useful for treating SARS.[5−7]

HRV 3CP is a cysteine protease with strong resemblance to trypsin-like serine proteases. The cleavage site in the polyprotein substrate lies between glutamine and glycine residues. The first described inhibitors resemble the natural substrates and therefore display peptidic structures (**1**−**4**). Due to their many disadvantages, e.g. the lack of oral availability and the poor synthetic accessibility, the search for peptidomimetic and small organic, nonpeptidic, highly potent substances began, which resulted in the development of the isatines (**9**, **10**), homophthalimides (**15**), azodicarboxamides (**17**), substituted benzamides (**7**, **8**), 2-pyridone containing peptidomimetics (**5**, **6**), and many more (Chart 1).

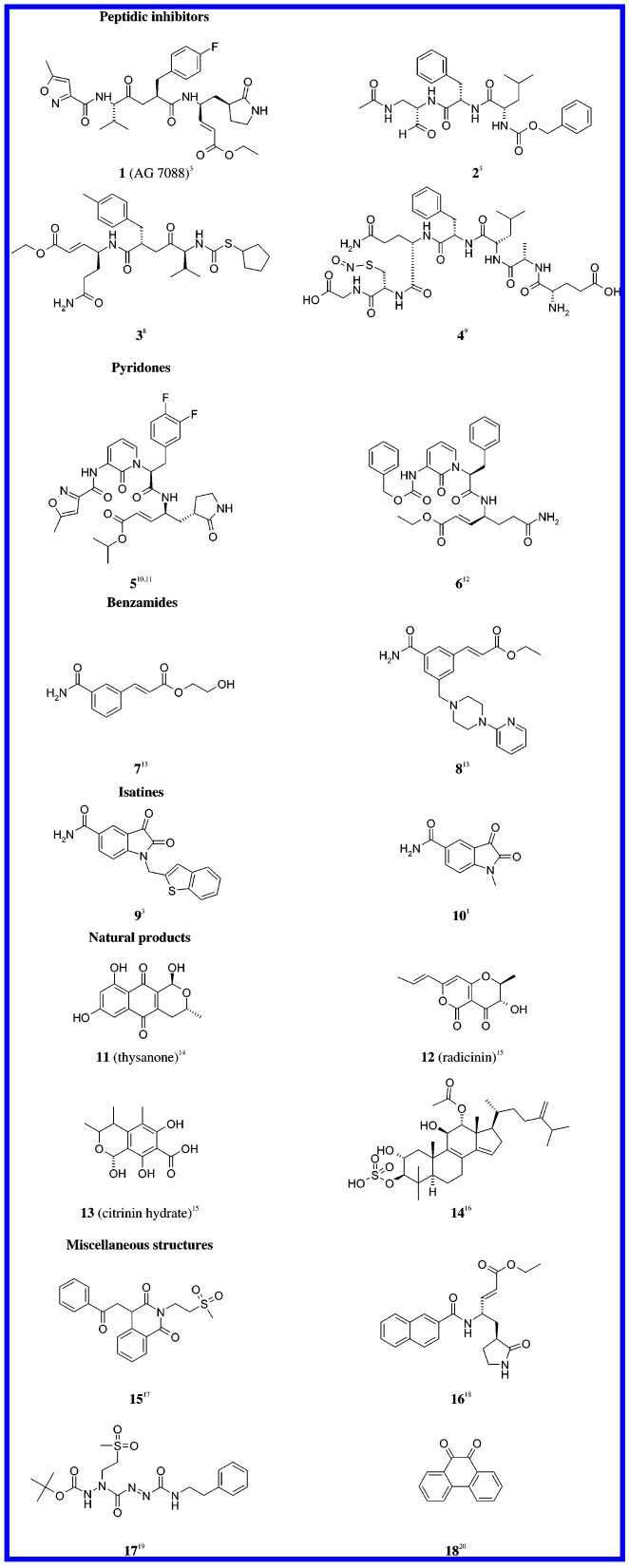
The aim of our study was the generation of selective pharmacophore models that describe the type, the size, and the position of chemical functions essential for a compound's antirhinoviral activity via 3CP inhibition, for both peptidic and nonpeptidic inhibitors. The pharmacophore models were generated within the program Catalyst.[21] This software package supports the modification of predefined features from the implemented feature dictionary as well as the creation of new ones, which was important for our application. A decisive structural requirement of HRV 3CP inhibitors is a group acting as a target for nucleophilic addition, because it enables the formation of a covalent bond between the ligand and the enzyme via a nucleophilic attack of the active site Cys 147. Literature sources provide much information on the design of reversible and irreversible inhibitors: Reversible inactivation is often achieved by introducing an aldehyde function,[22,23] whereas a prominent example of an irreversible inhibitor is the phase III compound AG 7088 (compound **1**, Chart 1),[3] which carries an α,β-unsaturated ethylester, a so-called Michael acceptor (MA). Scheme 1 shows the mechanism of enzyme inactivation.[18]

Accounting for this key interaction, the creation of a new function was the crucial step toward the development of pharmacophore hypotheses for HRV 3CP inhibitors. It includes a selection of electrophilic, chemical functionalities present in known HRV 3CP inhibitors, or defined in analogy thereto, that can act as targets in a nucleophilic addition reaction leading to covalent bond formation and enzyme inactivation. Including this function into the pharmacophore models guarantees that only molecules are retrieved as hits

* Corresponding author phone: +43 512 507 5252; fax: +43 512 507 5269; e-mail: thierry.langer@uibk.ac.at. Corresponding author address: Institute of Pharmacy, Department of Pharmaceutical Chemistry, Innrain 52, University of Innsbruck, A-6020 Innsbruck, Austria.

10.1021/ci049638a CCC: $30.25
Published on Web 03/29/2005

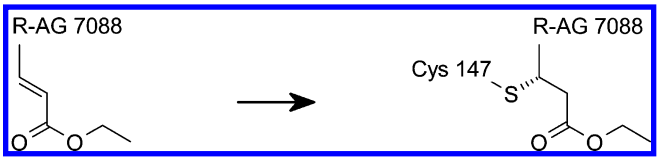
Chart 1. Examples of HRV 3CP Inhibitors Belonging to Different Structural Classes



that contain groups known to participate in such a reaction.

For the pharmacophore model generation, we took advantage of information from the Brookhaven Protein Databank (PDB)[24] on the only published cocrystallized complex of HRV 3CP with its inhibitor AG 7088 (PDB code 1CQQ). First, a structure-based approach was applied to achieve

Scheme 1. Inactivation Reaction of the HRV 3CP by Covalent Bond Formation of Active Site Cys 147 with the MA Moiety of Inhibitor AG 7088



pharmacophore models for the class of peptidic inhibitors. To refine and improve these structure-based models, they were compared and combined with automatically generated hypotheses from a training set of several highly active peptidic HRV 3CP inhibitors.

Fewer possibilities for interaction with the enzyme exist for the class of smaller, nonpeptidic inhibitors. They only occupy the innermost part of the binding site, where they also cause inactivation, by functioning as targets for nucleophilic attack. Identification of the essential pharmacophore features was achieved via comparison of the structural similarities of known nonpeptidic HRV 3CP inhibitors as well as from crystallographic data of the PDB complex 1CQQ.

Catalyst hypotheses can serve as query features in 3D database (DB) search. These virtual screening processes were focused on evaluating our models by searching an internal HRV 3CP inhibitor DB consisting of 30 peptidic and nonpeptidic compounds and the Derwent World Drug Index (WDI),[25] which contains drugs and other bioactive molecules. On the other hand, our aim was to retrieve structurally diverse hits, which provide a pool of possible new lead substances and candidates for biological testing, by screening other DBs.

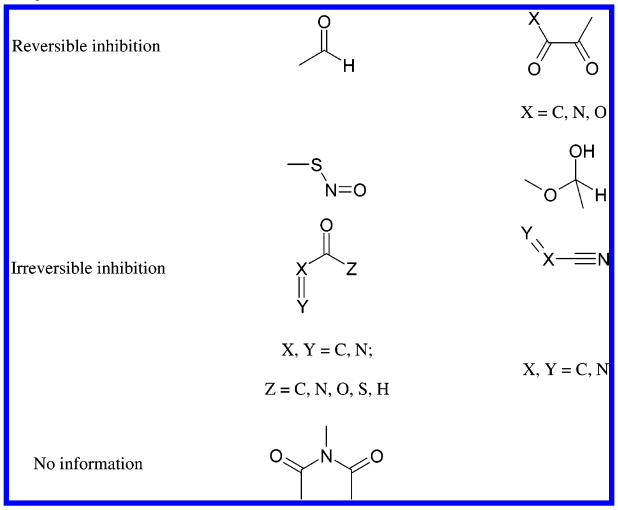
RESULTS AND DISCUSSION

1. Creation of a New Function. The ability of HRV 3CP inhibitors to inactivate the enzyme results from the existence of a chemical group acting as a target for nucleophilic addition by the sulfhydryl moiety of Cys 147 in the active site. This provokes the formation of a covalent bond between the inhibitor and the protein. Depending on the properties of this electrophilic key group in the ligand, the addition is either reversible or irreversible. Information on different kinds of electrophilic groups in HRV 3CP inhibitors and their modes of action was collected from literature.

Reversible inhibition results from reactive carbonyl groups such as aldehydes,[22,23] sometimes masked as hemiacetals,[15] and α-oxocarbonyl[1] compounds via addition to the carbon atom of the carbonyl group. Furthermore, S-nitrosothiols,[9] which act by transnitrosination of the Cys 147 thiol, have been suggested as reversible 3CP inhibitors. As distinct from this first class, all types of MAs (α,*β*-unsaturated carbonyl compounds)[3,8,10,12,13,18,26] form irreversible, covalent bonds to the enzyme. No data concerning the reversibility of bond formation is available for the category of diacylamines.[17] Occurrence of these groups in HRV 3CP inhibitors can be observed from the example structures in Chart 1.

To be able to include these reactive functions of the inhibitors in our pharmacophore models, a new Catalyst function, called TNA (target for nucleophilic addition), was created. Chart 2 shows the groups included in its definition. The displayed fragments are supposed to enable nucleophilic addition resulting in reversible or irreversible enzyme inhibi-

Chart 2. Fragment Set for the Definition of the New TNA Function in Catalyst



tion and were either directly taken from known HRV 3CP inhibitor structures and the appertaining literature sources or were defined in analogy thereto.

Within the Catalyst *Quicktool* manager, several fragments whose atom and bond properties are user-definable can be selected and combined to a new feature. Catalyst allows the inclusion of such newly created features into the feature dictionary and supports their application for automatic and manual hypothesis generation as well as for DB screening. We employed the fragments and fragment definitions displayed in Chart 2 for the assembly of a new Catalyst feature. Centroids for the fragments were defined, which function as origin for later feature positioning. Figure 1 shows the associations that were set in Catalyst for the selected fragment set composing the TNA function.

The occurrence of the newly created TNA function could be confirmed in all compounds in our internal HRV 3CP inhibitor DB using the *show function mapping* command in Catalyst. The TNA feature was therefore considered valuable for usage in the following pharmacophore model generation and DB screening. These models can be expected to retrieve structures containing the TNA function, i.e., the basic requirement for HRV 3CP inhibition. However, careful inspection of these hits is indispensable, since various influences, like sterical hindrance or participation in an aromatic π-electron system, might considerably decrease the reactivity of the groups that are mapped by our TNA feature.

To ensure that the TNA function itself does not already impose such high restrictions that a simple substructure search for its included fragments would reduce a DB to a sufficiently small number of compounds, various DBs were screened with this feature. Thereby obtained hitlists were found to be much larger than those resulting from later pharmacophore model screening approaches. E.g., when searching the Derwent WDI, 12 923 structures were retrieved, versus only a few hundred with the original pharmacophore hypothesis, stressing the need for a reasonable combination of substructure search represented by the TNA feature and other important functions for interaction in correct sterical arrangement. Besides, this prefiltering of DBs for TNA moiety containing compounds helped to decrease time consumption in the pharmacophore model screening processes, since only subsets of the DBs had to be searched.

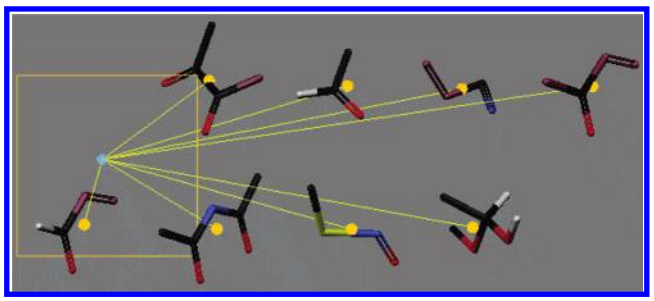


Figure 1. Association setting for fragments forming the newly created TNA function in the Catalyst Quicktool manager.

2. The HRV 3CP Binding Site. The most common junction sites for picornaviral 3CP cleavage are Gln-Gly peptide bonds. In rhinovirus 3CP, the catalytically important residues Cys 147, His 40, and Glu 71 form a linked cluster of amino acids. The highly conserved sequence Gly-X-Cys-Gly-Gly in viral 3CPs serves to position Cys 147 for nucleophilic attack on the carbonyl carbon of the substrate and to orient backbone −NH groups of Gly 145 and Cys 147 to form an "oxyanion hole" for the stabilization of a tetrahedral transition state. Efforts to study geometric and electronic factors for protein−ligand recognition identified specific, conserved residues that form the binding site and are arranged similarly in 3CPs from different rhinovirus serotypes.[3,23]

We retrieved information on the HRV 3CP binding site, essential interactions, and the bioactive ligand conformation from a crystallized enzyme−inhibitor complex in the PDB (entry code 1CQQ determined at a resolution of 1.85 Å). The enzyme-bound inhibitor AG 7088 is an irreversible, peptidic inactivator of HRV 3CP, which has reached phase III of clinical development. The following important possibilities for interactions are reported in the literature.[3]

From the C_β atom of the α,β-unsaturated carbonyl system, the MA moiety, formation of a covalent bond to Cys 147—the step of enzyme inactivation—can be observed. Aside from this essential covalent bond, several hydrophobic contacts and a dense network of hydrogen bonds between protein and ligand are accountable for the affinity of the ligand toward the binding site. Especially hydrogen bonds in the proximity of the MA are important to orientate the electrophilic β carbon into a favorable position for nucleophilic attack by Cys 147. A possibility for hydrophobic interaction emerges from the ethyl of the MA ester, which is situated near Phe 25. The MA carbonyl oxygen forms a hydrogen bond to the backbone amide of Cys 147. A Gln or Gln-like side chain, the so-called P1 region of the ligand, is decisive because it satisfies hydrogen bonds within the complementary S1 pocket as does the Gln of the natural substrate. The P1 side chain of AG 7088, a methylpyrrolidinone moiety, fulfills perfectly this requirement: The cyclic amide nitrogen donates a hydrogen bond to the backbone carbonyl oxygen of Thr 142, and the oxygen atom functions as an acceptor for hydrogen bonds from the side chains of His 161 and Thr 142. Val 162 accepts a hydrogen bond from the P1 backbone amide nitrogen of the ligand. Pocket S2 formed by His 40 and residues 127−130 accommodates the bulky P2 fluorophenyl group of AG 7088, which allows favorable hydrophobic contacts. Two main chain hydrogen bonds emerge from the P3 backbone region of the ligand to Gly 164. Furthermore,

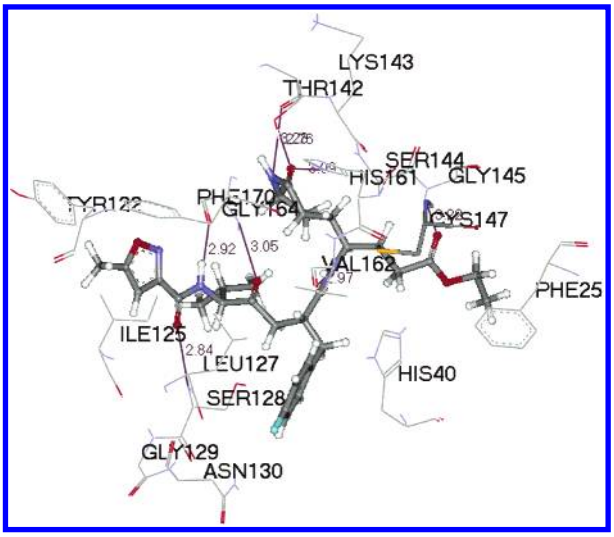


Figure 2. Inhibitor AG 7088 in HRV 3CP binding site. Interacting amino acid residues are labeled and distances of atoms participating in protein−ligand hydrogen bonding are displayed.

the backbone nitrogen of Ser 128 enables a hydrogen bond interaction to the terminal carbonyl oxygen of AG 7088. The methyl residue of its isoxazole ring is accommodated within a shallow hydrophobic pocket (Tyr 122, Phe 170, Ile 125).

Important residues forming the different pockets of the binding site and the interacting parts of the ligand as well as possibilities for hydrogen bond formation are displayed in Figure 2.

3. Hypothesis Generation. Hypotheses for peptidic inhibitors: Using the binding information and the 3D coordinates from PDB entry 1CQQ we manually generated a structure-based hypothesis, designed after the bioactive conformation of AG 7088. The two possibilities for hydrogen bonds from the P1 side chain oxygen to His 161 and Thr 142 were united in our model to one hydrogen bond accepting feature, whose projection point was given a sphere sufficiently large to allow interactions in both directions. This simplification hardly altered the quality of the hypotheses but made virtual screening considerably easier and less time-consuming. Other hydrogen bonds and hydrophobic interactions as well as the new TNA feature were arranged according to the interactions in the binding site described in the previous section. Owing to the numerous features, this first model turned out to be too restrictive since it left no hits from DB screening and was not even able to recognize AG 7088. Therefore, we concluded that special features, which do not strongly contribute to binding, had to be eliminated from the pharmacophore.

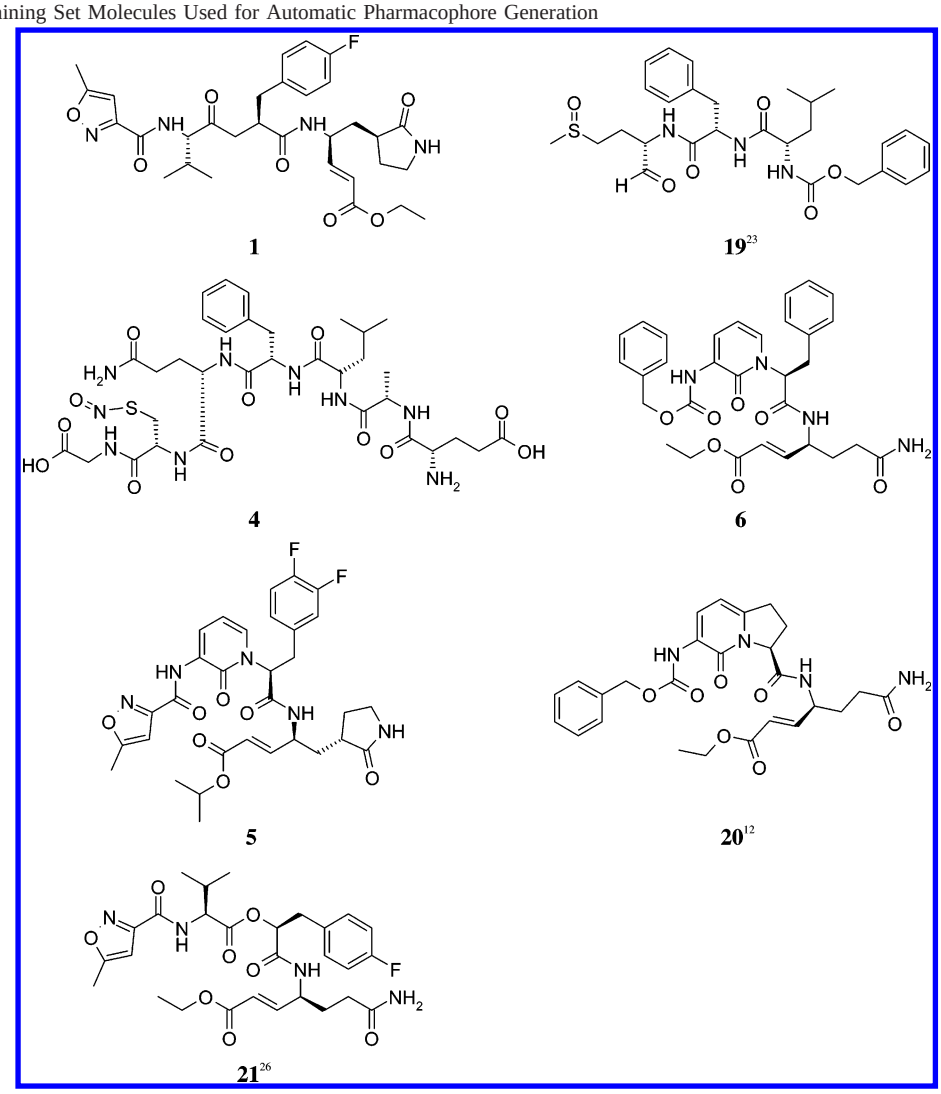
To find a combination of features, which is common to all peptidic inhibitors and therefore important for binding, automatic common feature hypotheses were generated with the Catalyst module HipHop. The training set, consisting of seven highly active peptidic HRV 3CP inhibitors, is shown in Chart 3.

For all compounds, 3D structures were built and conformational models were generated. In the pharmacophore generation process, AG 7088, whose bioactive conformation is known from crystallographic studies, was given a principal value of two, meaning its structure and conformation would have the strongest influence in the model building phase. A minimum of one hydrogen bond donor (HBD), one hydrogen bond acceptor (HBA), one hydrophobic feature (H), and one TNA feature was demanded. In this experiment, 10 five-featured hypotheses were obtained. The two best, judged by their resemblance to the structure-based model, were merged and complemented each other to a six-featured hypothesis (hypo1, see Figure 3). Its features cover the interacting chemical functions in the ligand AG 7088 as follows: TNA for the electrophilic MA moiety, HBA for the P1 side chain oxygen, HBD for the P1 backbone amide nitrogen, H for the P2 hydrophobic phenyl residue, and H and HBA representing the interaction of the P3 and P4 area. This hypothesis displays favorable selectivity for peptidic inhibitors—it detects all 14 peptidic structures from the internal HRV 3CP inhibitor DB—and maps AG 7088 with a fit value of 5.8. The fit value is a measurement of a compound's structure matching to the pharmacophore hypothesis and in this case ranges from 0 to 6 (the latter describing perfect fit). Studying the crystal structure proves that this hypothesis describes the situation in the enzyme quite precisely. However, because of the reduced number of features it is considerably less restrictive than the structure-based, manual model described above. Screening the Derwent WDI with this hypothesis as query feature results in 367 hits (0.77% of the entire DB). Among them, 14 substances with known virucide activity could be found. The 1D property field "mode of action" showed that 18 of these compounds had previously been reported as protease inhibitors. These include all three known HRV 3CP inhibitors included in the Derwent WDI. To obtain new ideas for lead structure development, additionally other DBs were screened: From the commercially available Maybridge DB 2004 (nearly 60 000 compounds)[27] and the Interbioscreen DBs (approximately 30 000 natural and 290 000 synthetic compounds)[28] seven (0.01%) and 260 hits (0.08%) could be found, respectively.

A merging operation of hypo1 with the original structure-based model led to a nine-featured hypothesis. One HBA and one HBD interacting with the S3 region of the binding site as well as one HBD interacting with the S1 pocket had to be deprived of their projection points in order to prevent too high restrictiveness of the model leading to hypo2 (Figure 3). From our internal HRV 3CP inhibitor DB 80% (11 out of 14) of the peptidic compounds are mapped, among them AG 7088 with a fit value of 8.6 (maximum 9.0). The hitlist from the WDI amounts to only 77 (0.16%) structures. Many of those hits are indicated in 1D data searching as antibiotic, cytostatic, fungicide, or hydrolase inhibiting agents. Among them eight protease inhibitors can be found including the three known active compounds.

A less restrictive hypothesis resulted from the elimination of the P1 side chain HBD (hypo3), see Figure 3, since it also allows the detection of HRV 3CP inhibitors such as the peptidic compound **19**[23] (Chart 3), lacking this interaction possibility. From the WDI DB, 247 compounds (0.51%) including the three antirhinoviral protease inhibitors fulfill the requirements of this model. Figure 5 shows one hit from the Interbiosceen natural compound DB aligned with hypo3. However, these three models strongly focus on the peptidic character of the training set molecules, so that other structures, e.g. compounds from the Maybridge DB, whose chemical structures were not designed to influence biological systems or to exhibit drug-like properties, will hardly ever be retrieved as hits in 3D DB screening, as can be seen from

Chart 3. Peptidic Training Set Molecules Used for Automatic Pharmacophore Generation



the very low number of compounds resulting in the hitlists. Details about the performance of these peptidic models (hypo1−3) can be seen in Table 1. Chart 4 displays some examples of screening hits that were provided with high fit values.

Hypotheses for nonpeptidic inhibitors: As distinct from large molecules, like the peptidic inhibitors, where it is not necessary, sometimes even not advisable to include all interacting functions in the pharmacophore model, a different challenge occurs for the small, nonpeptidic compounds. Because of their smaller size, they incorporate fewer functional groups and consequently possess only a limited number of possibilities for interaction with the protein. Therefore, all interacting functions should be included into hypothesis generation to achieve enough features for sufficient selectivity in virtual DB screening. Although crystallographic data remain unpublished, literature sources describe proteins cocrystallized with nonpeptidic inhibitors.[3,13,18] They give evidence that these small ligands occupy the S1 and S2 area of the rhinovirus 3CP binding site and that the binding process is guided by interactions in this region.[1,18] The HBA and HBD in the P1 side chain as well as the TNA and a hydrophobic feature for the P2 region were considered as essential pharmacophore features. These interactions ensure that the activated, electrophilic TNA moiety acquires a favorable position for nucleophilic attack.

Information on the design, activity, and mode of action of several classes of nonpeptidic HRV 3CP inhibitors was gained from various literature sources. Members of the structural classes (Chart 1) discussed in the following paragraph were used for pharmacophore model generation and as test compounds—as they are part of our internal inhibitor DB—to verify the quality of the obtained models.

Among the most prominent classes of HRV 3CP inhibitors rank the 2,3-dioxindoles (isatines).[1] Despite the fact that they proved unsuitable for clinical use because of their disadvantageous pharmacological properties,[18] some of them possess excellent selectivity and affinity for the target enzyme and were therefore considered valuable for pharmacophore model generation. Furthermore, studies on nonpeptidic ligands, containing a lactame moiety and an MA[18] as well as studies on substituted benzamides have been published.[13] Inclusion of the latter in model generation was avoided because of their missing occupancy of the S2 pocket. Nevertheless, they were integrated in the internal inhibitor DB. Also for some natural products, like thysanone[14] (compound **11**) from *Thysanophora penicilloides* or radicinin[15] (compound **12**) from *Curvularia*, inhibitory activity on HRV 3CP could be demonstrated. 2-Pyridone containing peptidomimetics[10,12] were also included in the test set. The group of homophthalimides[17] does not conform to these other nonpeptidic inhibitors because of the lack of a P1 HBD. So they could

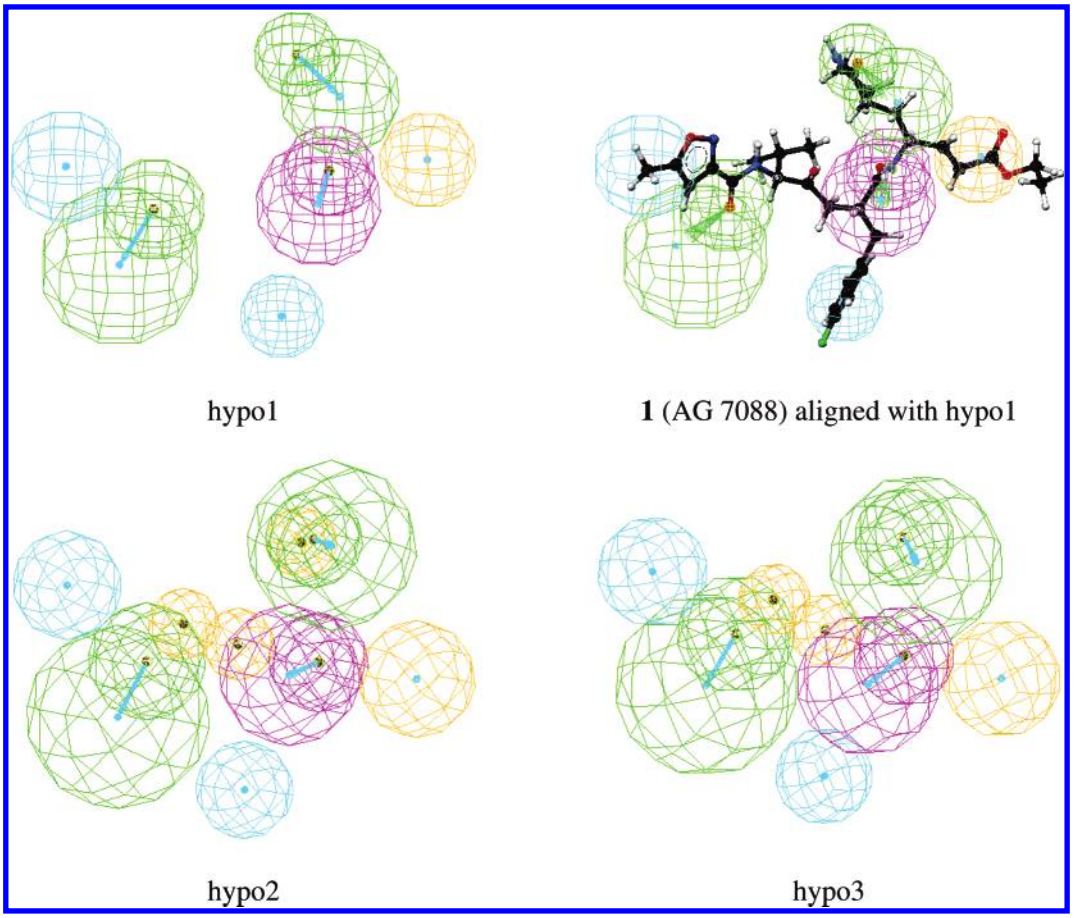


Figure 3. Pharmacophore hypotheses hypo1, hypo2, and hypo3 for peptidic class of HRV 3CP inhibitors. Pharmacophore features are color coded (green: HBA; violet: HBD; orange (larger sphere): TNA (newly created function); orange (three smaller spheres): HBAs and HBDs deprived of their projection points; light-blue: H).

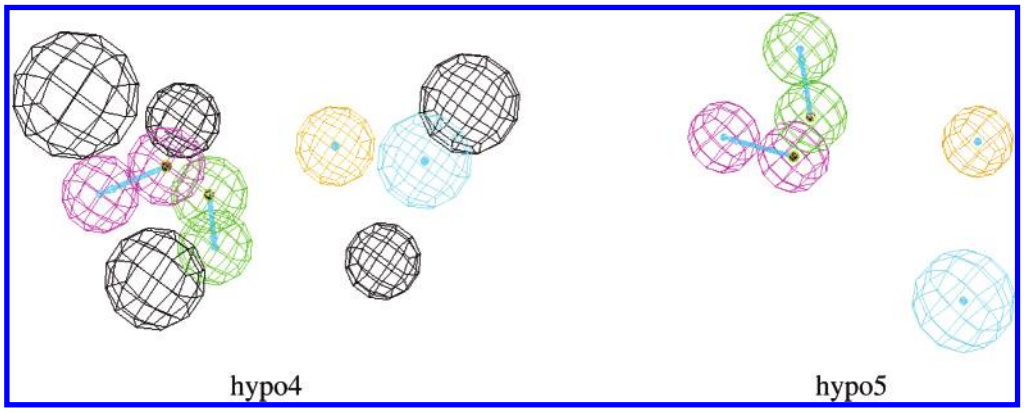


Figure 4. Pharmacophore hypotheses hypo4 and hypo5 for nonpeptidic class of HRV 3CP inhibitors. Pharmacophore features are color coded (green: HBA; violet: HBD; orange: TNA (newly created function); light-blue: H; black: excluded volume spheres).

neither be used for building the models nor were they detected in the following virtual screening process of our internal inhibitor DB.

At first, approaches for pharmacophore identification were focused on the automatic generation process in Catalyst using the module HipHop. Various combinations of nonpeptidic HRV 3CP inhibitors belonging to different structural classes were employed as training set molecules. However, no appropriate generation conditions, i.e., a representative training set, a valid and sufficiently high number of minimally demanded features, and suitable feature selection, could be found for these automatic approaches. The corresponding processes either left no data for collection or returned very unspecific results.

Consequently, manual pharmacophore hypothesis construction was the strategy of choice. The *show function mapping* command in Catalyst allows confirming or denying the occurrence of a specific feature in a molecule. Intensive studies on different functions and their relative positions to each other among the nonpeptidic HRV 3CP inhibitors were carried out. Thereby, we focused on important ligand features interacting with the S1 and S2 region of the protein. Most ligands contain the required P1 HBA and HBD as well as the TNA function. Noticeable differences lie in the position

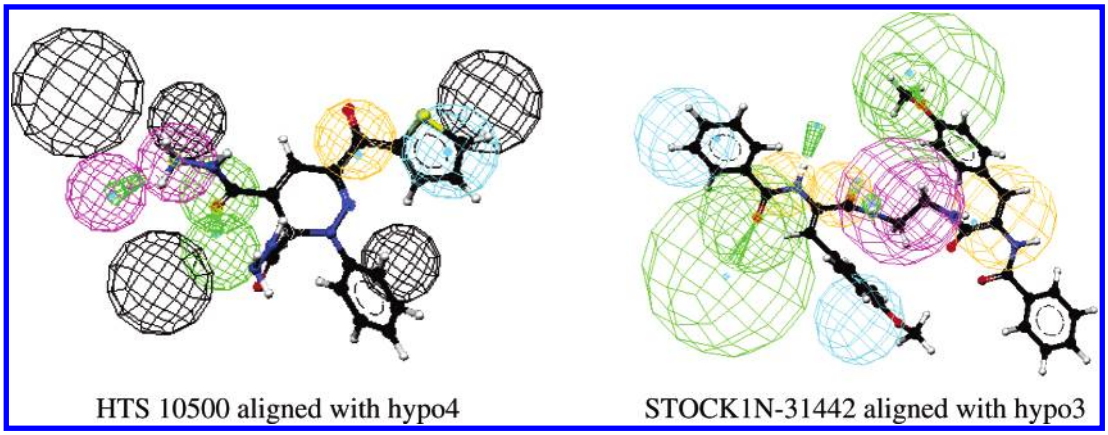
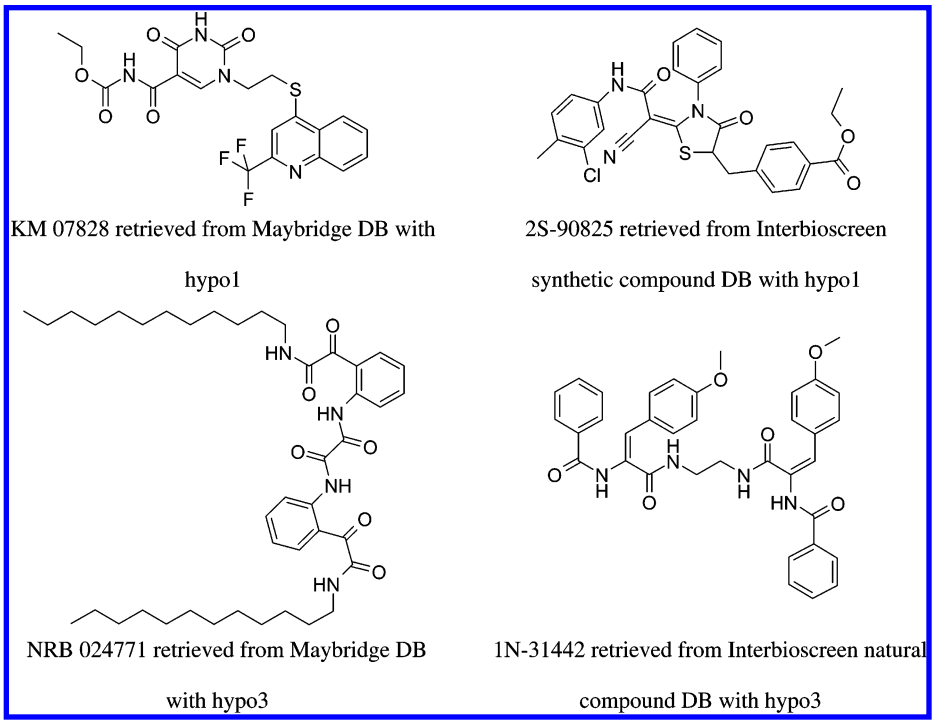


Figure 5. Examples of hits from virtual DB screening aligned with hypotheses used as query features. Color codes see Figures 3 and 4.

Table 1. Results from Virtual DB Screening Experiment of the Derwent WDI, the Maybridge DB, the Interbioscreen Synthetic, and Natural DBs Using Pharmacophore Models for Inhibition of HRV 3CP by Peptidic as Well as Nonpeptidic Inhibitors

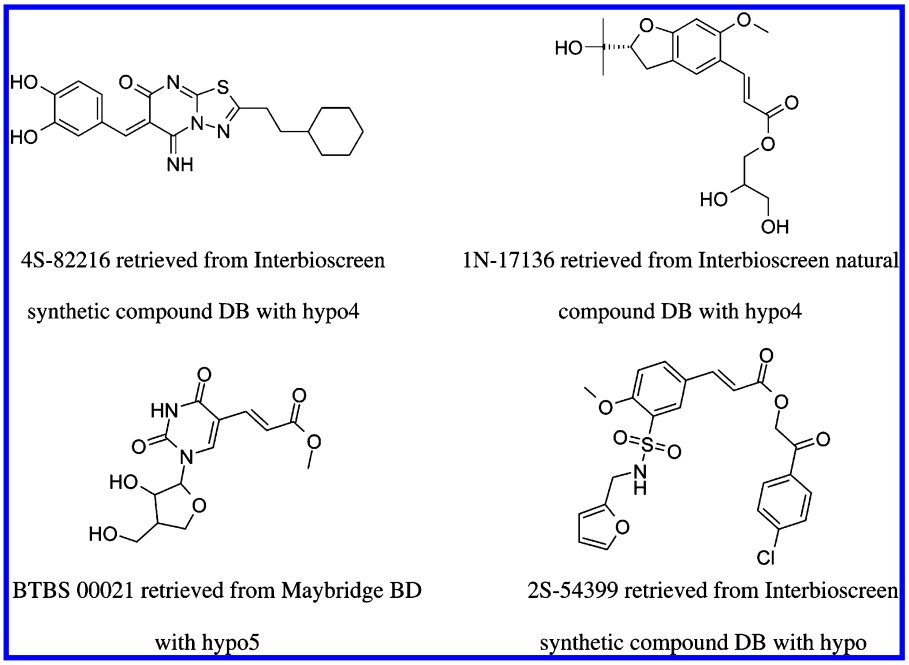
search query	hypo1	hypo2	hypo3	hypo4	hypo5
application area	good selectivity, for peptidic inhibitors	highly selective, for peptidic inhibitors	general model for peptidic inhibitors	general model for peptidic and nonpeptidic inhibitors	general model for peptidic and nonpeptidic inhibitors
WDI	376	77	247	2777	2467
% of DB	0.8	0.2	0.5	5.7	5.1
Maybridge	7	0	1	481	129
% of DB	<0.1	0	<0.1	0.8	0.2
Interbioscreen (synthetic)	180	1	16	12202	4910
% of DB	0.1	<0.1	<0.1	4.2	1.7
Interbioscreen (natural)	80	25	78	3069	2007
% of DB	0.3	0.1	0.3	10.2	6.7

Chart 4. Examples of Structures Obtained from Screening the Maybridge DB and the Interbioscreen DBs with Pharmacophore Models for Peptidic HRV 3CP Inhibitors



of the P2 hydrophobic substituent, which is rather small in some ligands, like the methyl group of the isatine compound **10**, or larger, for example in the case of the 2-methylbenzothiophene residue of another isatine (**9**), which reaches deeper into the S2 pocket. Several nonpeptidic HRV 3CP inhibitors were chosen and transformed into pharmacophore hypotheses by placing the desired P1 and P2 features from the Catalyst feature dictionary onto the complementary groups of the molecule. The best hypotheses, judged by their ability to detect the majority of the nonpeptidic compounds of our internal inhibitor DB, were derived from the isatine compound **10**, radicinin (**12**), and thysanone (**11**), the latter

Chart 5. Examples of Structures Obtained from Screening the Maybridge DB and the Interbioscreen DBs with Pharmacophore Models for Nonpeptidic HRV 3CP Inhibitors



representing the group of the quinone-like ligands. Since these inhibitors present rigid ring systems with only very limited rotational freedom, the usually vital question, which conformation to choose for manual pharmacophore model generation, did not pose any difficulties. A merging operation using these best-performing three hypotheses resulted in a model, which could also map ligand AG 7088 satisfactorily.

This hypothesis derived from compounds **10**, **11**, and **12** could be further refined: The positions of the four features (HBA, HBD, TNA, H) were slightly altered according to their locations in the bioactive conformation of AG 7088. The refinement process was again enabled by merging operations.

For further advancement and improved selectivity, the resulting hypothesis was provided with so-called excluded volume spheres. To be identified as a hit by the hypothesis, a ligand must not penetrate these forbidden, excluded areas, which are placed at the desired positions surrounding the pharmacophore features and characterize the steric limitations of the binding site. Size and coordinates of the excluded volumes spheres were determined according to crystallographic information on the active site amino acids forming the S1 and S2 pocket.

This hypothesis, hypo4, displayed in Figure 4, finds 15 out of 16 nonpeptidic components of the internal HRV 3CP inhibitor DB as well as a considerable share (70%) of the peptidic ones, including AG 7088. When applied for 3D DB screening, this model retrieved 2777 hits (5.68%) from the WDI, 481 hits (0.81%) from the Maybridge DB, 12 202 hits (4.24%) from the synthetic compounds of the Interbioscreen DB, and 3069 of the natural compounds (10.24%) (Table 1). The members of these hitlists were submitted to fit value calculations to determine their match with the hypothesis and allow a consequential ranking. Analysis of the WDI hitlist revealed 32 protease inhibitors. Aside from three HRV 3CP inhibitors these include many substances labeled as cysteine, serine, or HIV protease inhibitors.

To include as much knowledge as possible from the only available crystallographic structure, a four-featured model (HBA, HBD, TNA, H) was built from AG 7088 using the bioactive ligand conformation and the coordinates of the amino acid atoms involved in hydrogen bond interactions. Its four features represent those interactions identified as essential for the binding of nonpeptidic inhibitors. Thereby HBA and HBD represent the cyclic amide of the P2 region, TNA encodes for the MA, and H was put on the phenyl ring. The four-featured, structure-based pharmacophore model derived from AG 7088 was found to give a suitable characterization of nonpeptidic inhibitors with larger P2 substituents, which occupy the S2 pocket more thoroughly. Therefore, this hypothesis, called hypo5 (Figure 4), was used for 3D DB searching. Still nearly 80% of the compounds forming our internal inhibitor DB were mapped (23 out of 30). Among them, 13 nonpeptidic structures could be found, whereas 10 displayed peptidic characteristics. Of course now those compounds with the small-size P2 residues were missing. Searching the WDI resulted in 2467 hits (5.05%), whereas from the Maybridge DB only 129 structures (0.22%) were retrieved. The Interbioscreen synthetic and natural compound DBs could be reduced to 1.71% and 6.70%, respectively (Table 1). Among the hits from the Derwent WDI search we found three HRV 3CP as well as 26 other protease inhibitors.

A selection of well fitting, structurally interesting hits obtained from DB screening processes with hypotheses hypo4 and hypo5 can be viewed in Chart 5. These molecules, which acquired high fit values, may present candidates for biological testing and lead detection. Figure 5 shows one hit from the Maybridge DB aligned with hypo4.

EXPERIMENTAL SECTION

All molecular modeling studies were performed using Catalyst 4.7 installed on a Silicon Graphic Octane desktop workstation equipped with a 300 MHz MIPS R12000

processor (1 GB RAM) running the Irix 6.5 operating system. The images displayed in figures were derived from the Catalyst program and from WebLabViewerLite version 3.5. Molecular structures were drawn with the program Isis Draw 2.4. For molecules without defined stereochemistry no corresponding information is available. The flexibility of each molecule was represented by a set of energetically reasonable conformers which were generated with the Catalyst catConf module choosing a maximum number of 250 conformers, the *best* quality generation type, and an energy threshold of 20 kcal/mol beyond the calculated global energy minimum. Virtual DB screening was carried out within the *best flexible* search modus in the Catalyst DBServer module. The Interbioscreen DBs were generated with the catDB command (FAST method, maximum number of conformers = 100) using SD files downloaded from the Web dated July 2003. The Maybridge DB 2004 was provided by Accelrys.[21,27,28]

CONCLUSIONS

In this study we describe the development of pharmacophore models for peptidic and nonpeptidic inhibitors of the HRV 3CP using the software Catalyst. Both automatic and manual generation strategies were applied guided by crystallographic information on the binding mode of a peptidic inhibitor in complex with the enzyme. The main focus was laid on defining a new pharmacophore feature representing a target structure for nucleophilic addition in the ligands, which is vital for protease inactivation. Inclusion of this highly specific chemical function in the 3D arrangement of pharmacophoric features ensures that only structures of suitable chemical reactivity are identified by the models. Subsequently, our hypotheses were tested in virtual screening processes, returning promising hitlists from 3D DBs containing HRV 3CP inhibitors. Therefore, our pharmacophore models are considered as valuable tools for 3D virtual screening of DBs to detect novel lead structures for HRV 3CP inhibitors.

Abbreviations: HRV, human rhinovirus; 3CP, 3C protease; MA, Michael Acceptor; PDB, Brookhaven Protein Databank; DB, database; TNA, target for nucleophilic addition; WDI, Derwent World Drug Index; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; H, hydrophobic feature.

ACKNOWLEDGMENT

We thank Dr. Rémy D. Hoffmann (Accelrys SARL, Paris, France) for performing the searching experiments within the WDI database.

REFERENCES AND NOTES

(1) Webber, S. E.; Tikhe, J.; Worland, S. T.; Fuhrman, S. A.; Hendrickson, T. F. et al. Design, Synthesis, and Evaluation of Nonpeptidic Inhibitors of Human Rhinovirus 3C Protease. *J. Med Chem.* **1996**, *39*, 5072−5082.
(2) Hernandez, A. A.; Roush, W. R. Recent advances in the synthesis, design and selection of cysteine protease inhibitors. *Curr. Opin. Chem. Biol.* **2002**, *6*, 459−465.
(3) Matthews, D. A.; Dragovich, P. S.; Webber, S. E.; Fuhrman, S. A.; Patick, A. K. et al. Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11000−11007.
(4) Kaldor, S. W.; Hammond, M.; Dressman, B. A.; Labus, J. M.; Chadwell, F. W. et al. Glutamine-Derived Aldehydes for the Inhibtion of Human Rhinovirus 3C Protease. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2021−2026.
(5) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763−1767.
(6) Zhang, X. W.; Yap, Y. L. Exploring the binding mechanism of the main proteinase in SARS-associated coronavirus and its implication to anti-SARS drug design. *Bioorg. Med. Chem.* **2004**, *12*, 2219−2223.
(7) Jenwitheesuk, E.; Samudrala, R. Identifying inhibitors of the SARS coronavirus proteinase. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3989−3992.
(8) Dragovich, P. S.; Prins, T. J.; Zhou, R.; Fuhrman, S. A.; Patick, A. K. et al. Structure-Based Design, Synthesis, and Biological Evaluation of Irreversible Human Rhinovirus 3C Protease Inhibitors. 3. Structure−Activity Studies of Ketomethylene-Containing Peptidomimetics. *J. Med. Chem.* **1999**, *42*, 1203−1212.
(9) Xian, M.; Wang, Q. M.; Chen, X.; Wang, K.; Wang, P. G. S-Nitrosothiols as Novel, Reversible Inhibitors of Human Rhinovirus 3C Protease. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2097−2100.
(10) Dragovich, P. S.; Prins, T. J.; Zhou, R.; Brown, E. L.; Maldonado, F. C. et al. Structure-Based Design, Synthesis, and Biological Evaluation of Irreversible Human Rhinovirus 3C Protease Inhibitors. 6. Structure−Activity Studies of Orally Bioavailable, 2-Pyridone-Containing Peptidomimetics. *J. Med. Chem.* **2002**, *45*, 1607−1623.
(11) Dragovich, P. S.; Prins, T. J.; Zhou, R.; Johnson, O. T.; Hua, Y. et al. Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 8. Pharmacological optimization of orally bioavailable 2-pyridone-containing peptidomimetics. *J. Med. Chem.* **2003**, *46*, 4572−4585.
(12) Dragovich, P. S.; Prins, T. J.; Zhou, R.; Johnson, T. O.; Brown, E. L. et al. Structure-Based Design, Synthesis, and Biological Evaluation of Irreversible Human Rhinovirus 3C Protease Inhibitors. Part 7: Structure−Activity Studies of Bicyclic 2-Pyridone-Containing Peptidomimetics. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 733−738.
(13) Reich, S. H.; Johnson, T.; Wallace, M. B.; Kephart, S. E.; Fuhrman, S. A. et al. Substituted Benzamide Inhibitors of Human Rhinovirus 3C Protease: Structure-Based Design, Synthesis, and Biological Evaluation. *J. Med. Chem.* **2000**, *43*, 1670−1683.
(14) Singh, S. B.; Cordingley, M., G.; Ball, R. G.; Smith, J. L. M.; Dombrowski, A. W. et al. Structure and stereochemistry of thysanone: a novel human rhinovirus 3C-protease inhibitor from Thysanophora penicilloides. *Tetrahedron Lett.* **1991**, *32*, 5279−5282.
(15) Kadam, S.; Poddig, J.; Humphrey, P.; Karwowski, J.; Jackson, M. et al. Citrinin hydrate and radicinin: human rhinovirus 3C-protease inhibitors discovered in a target-directed microbial screen. *J. Antibiot.* **1994**, *47*, 836−839.
(16) Brill, G. M.; Kati, W. M.; Montgomery, D.; Karwowski, J. P.; Humphrey, P. E. et al. Novel triterpene sulfates from Fusarium compactum using a rhinovirus 3C protease inhibitor screen. *J. Antibiot.* **1996**, *49*, 541−546.
(17) Jungheim, L. N.; Cohen, J. D.; Johnson, R. B.; Villarreal, E. C.; Wakulchik, M. et al. Inhibition of Human Rhinovirus 3C Protease by Homophthalimides. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1589−1594.
(18) Johnson, T. O.; Hua, Y.; Luu, H. T.; Brown, E. L.; Chan, F. et al. Structure-Based Design of a Parallel Synthetic Array Directed Toward the Discovery of Irreversible Inhibitors of Human Rhinovirus 3C Protease. *J. Med. Chem.* **2002**, *45*, 2016−2023.
(19) Hill, R. D.; Vederas, J. C. Azodicarboxamides: A New Class of Cysteine Protease Inhibitor for Hepatitis A Virus and Human Rhinovirus 3C Enzymes. *J. Org. Chem.* **1999**, *64*, 9538−9546.
(20) Singh, S. B.; Graham, P. L.; Reamer, R. A.; Cordingley, M. G. Discovery, Total Synthesis, HRV 3C-Protease Inhibitory Activity, and Structure−Activity Relationships of 2-Methoxystypandrone and Its Analogues. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 3143−3146.
(21) Accelrys Inc., Catalyst, version 4.7, San Diego 2002. http://www.accelrys.com.
(22) Shepherd, T. A.; Cox, G. A.; McKinney, E.; Tang, J.; Wakulchik, M. et al. Small Peptidic Aldehyde Inhibitors of Human Rhinovirus 3C Protease. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2893−2896.
(23) Webber, S. E.; Okano, K.; Little, T. L.; Reich, S. H.; Xin, Y.; Fuhrman, S. A. et al. Tripeptide Aldehyde Inhibitors of Human Rhinovirus 3C Protease: Design, Synthesis, Biological Evaluation, and Cocrystal Structure Solution of P1 Glutamine Isosteric Replacements. *J. Med. Chem.* **1998**, *41*, 2786−2805.
(24) Brookhaven National Laboratory: Protein Data Bank, http://www.rcsb.org/pdb/.
(25) Derwent World Drug Index, http://www.derwent.com/.
(26) Webber, S. E.; Marakovits, J. T.; Dragovich, P. S.; Prins, T. J.; Zhou, R. et al. Design and Synthesis of Irreversible Depsipeptidyl Human Rhinovirus 3C Protease Inhibitors. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2683−2686.
(27) Maybridge Database, Tintagel, http://www.maybridge.com/.
(28) Interbioscreen Database, Moscow, http://www.ibscreen.com/.

CI049638A