

VIRAL PROTEASE INHIBITORS WITH A HYDROXYBENZYL ALCOHOL WARHEAD**Field of the invention**

The present invention relates to compounds that have new designs for viral protease inhibitors. The compounds can have structures that offer improved properties, particularly for warheads and P1 residues. The invention also relates to compositions and medical uses of such compounds, and to intermediate products that are useful for the provision of such compounds.

Background art

Coronaviruses are prominent sources of infections such as respiratory viral infections. They are a group of related RNA viruses known to cause diseases in mammals and birds. Particularly prominent are respiratory tract infections that can range from mild to lethal. Mild illnesses in humans include some cases of the common cold (which is also caused by other viruses, predominantly rhinoviruses), while more lethal varieties of coronavirus can cause SARS, MERS, and COVID-19. In cows and pigs they are known to cause diarrhea, while in mice they can cause hepatitis and encephalomyelitis.

Severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1; or Severe acute respiratory syndrome coronavirus, SARS-CoV) is a strain of coronavirus that causes severe acute respiratory syndrome (SARS), the respiratory illness responsible for the 2002–2004 SARS outbreak. The species is a member of the genus Betacoronavirus and subgenus Sarbecovirus. It is an enveloped, positive-sense, single-stranded RNA virus that infects the epithelial cells within the lungs. The virus enters the host cell by binding to angiotensin-converting enzyme 2. It infects humans, bats, and palm civets. The SARS-CoV-1 outbreak was largely brought under control by simple public health measures. Testing people with symptoms (fever and respiratory problems), isolating and quarantining suspected cases, and restricting travel all had an effect. SARS-CoV-1 was most transmissible when patients were sick, and so by isolating those with symptoms, onward spread was prevented. As SARS is a viral disease, antibiotics do not have direct effect but may be used against bacterial secondary infection. Treatment of SARS is mainly supportive with antipyretics, supplemental oxygen and mechanical ventilation as needed. While ribavirin is commonly used to treat SARS, there seems to have little to no effect on SARS-CoV, and no impact on patient's outcomes. There is no proven antiviral therapy known in the art. Tested substances, include ribavirin, lopinavir, ritonavir, type I interferon, which have thus far shown no conclusive contribution to the disease's course.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a strain of coronavirus that causes COVID-19 (coronavirus disease 2019), a respiratory illness responsible for a pandemic. First identified in the city of Wuhan, Hubei, China, the World Health Organization declared the outbreak a public health emergency of international concern on January 30, 2020, and a pandemic on March 11, 2020. The species is a member of the genus Betacoronavirus and subgenus Sarbecovirus. It is a positive-sense single-stranded RNA virus that is contagious in humans. The symptoms of COVID-19 are variable but often include fever, cough, headache,

fatigue, breathing difficulties, loss of smell, and loss of taste. Symptoms may begin one to fourteen days after exposure to the virus. At least a third of people who are infected do not develop noticeable symptoms. Of those who develop symptoms noticeable enough to be classified as patients, most (81%) develop mild to moderate symptoms (up to mild pneumonia), while 14% develop severe symptoms (dyspnea, hypoxia, or more than 50% lung involvement on imaging), and 5% develop critical symptoms (respiratory failure, shock, or multiorgan dysfunction). Older people are at a higher risk of developing severe symptoms. Some people continue to experience a range of effects (long COVID) for months after recovery, and damage to organs has been observed. Multi-year studies are underway to further investigate the long-term effects of the disease. The primary treatment is symptomatic. Management involves the treatment of symptoms through supportive care, isolation, and experimental measures. One of the few known treatment options is based on nirmatrelvir/ritonavir, which requires administration of many pharmaceutical doses, which can strain patient compliance.

Middle East respiratory syndrome-related coronavirus (MERS-CoV), or EMC/2012 (HCoV-EMC/2012), is the virus that causes Middle East respiratory syndrome (MERS). It is a species of coronavirus which infects humans, bats, and camels. The infecting virus is an enveloped, positive-sense, single-stranded RNA virus which enters its host cell by binding to the DPP4 receptor. The species is a member of the genus Betacoronavirus and subgenus Merbecovirus. MERS is a viral respiratory infection and symptoms may range from none, to mild, to severe depending on age and risk level. Typical symptoms include fever, cough, diarrhea, and shortness of breath. The disease is typically more severe in those with other health problems. There is no specific vaccine or treatment for the disease. Using extra-corporeal membrane oxygenation (ECMO) seems to improve outcomes significantly. Neither the combination of antivirals and interferons (ribavirin + interferon alfa-2a or interferon alfa-2b) nor corticosteroids improved outcomes (Zumla A, et al., 2016, Nature Reviews Drug Discovery. 15 (5): 327–47).

At present, all treatments given for this infection are off-label because no standard care or effective drug is yet available; out of necessity, only the symptoms can be treated. COVID-19 patients for instance regularly end up on ventilation at the ICU. However, ventilation in COVID-19 patients has been shown to be associated with many complications and high mortality. Patients enrolled in the ICU show severe edema of the lung, a decrease in serum albumin, and also appear to have a relatively low heart rate (consistent with SARS and ARDS image). Some patients worsen after 2 weeks of illness due to sterile inflammation. In addition, they have poor gas exchange, following breathlessness symptoms. Not uncommon with and after a COVID-19 infection are bacteria and fungal co-infections as well as cardiological and haematological (thromboembolic) complications.

Flavivirus is a genus of the family *Flaviviridae*. This genus includes the West Nile virus (WNV), dengue virus (DEN), Tick-borne Encephalitis Virus, Japanese Encephalitis virus, Yellow Fever Virus, and several other viruses that may cause encephalitis.

Flaviviruses are viruses that cause millions of infections each year. Dengue viruses cause a self-limiting disease in humans called dengue fever (DF), which is often resolved in 7-10 days.

However, more severe forms of the disease, known as Dengue hemorrhagic fever (DHF) and Dengue shock syndrome (DSS), are common in areas endemic to DEN and lead to considerable morbidity and mortality. According to World Health Organization estimates, 50-100 million cases of DEN infections in tropical and subtropical countries occur each year.

5 Complicating matters further is the fact that DEN exists as at least four separate serotypes (DEN-1, DEN-2, DEN-3, and DEN-4) with DEN-2 being the most prevalent in many recent epidemics. Unfortunately infection by one serotype does not provide protection from infections by the other serotypes. Furthermore, evidence suggests that subsequent infections by different serotypes may increase the probability of developing the more serious forms of the
10 disease like DHF and DSS.

Every year, it is estimated that there are 50-100 million dengue virus infections with ~1.5 million documented cases of dengue fever, and ~500,000 cases of dengue hemorrhagic fever and shock syndrome. Reported cases increase annually. Approximately 40% of the world's population is at risk of dengue infection from living in regions endemic with the virus.

15 Symptoms of DENV infection include sudden onset of fever, body pain, headache, joint pain, rashes, and retro-orbital pain. Although usually mild or asymptomatic, the infection can progress into life-threatening serious conditions such as haemorrhagic fever (DHF). The symptoms of mild haemorrhage include petechiae, purpura, ecchymoses, and epistaxis. Up to 2% of the total number of DENV cases (mostly children below 15 years of age) experience
20 progression of the infection to severe DHF, characterized by thrombocytopenia, haemorrhagic manifestations that could affect the skin, nose, gum, and gastrointestinal tract. Dengue shock syndrome (DSS), the most severe form of DHF, is characterized by a weak pulse and sudden drop in blood pressure, which is the result of the collapse of the vascular system owing to hypovolemia caused by vascular leakage. Patients with co-morbidities (e.g. diabetes mellitus,
25 hypertension, cardiac or renal failure, sickle cell anaemia) or the at-risk populations (such as pregnant women, infants, the elderly) are at risk for developing severe disease. Patients with severe dengue have difficulty drinking, which hampers reversal of the effect of shock. Consequently, patients with the more severe forms of dengue may not be able to take oral medications.

30 West Nile virus (WNV) was introduced into the Western Hemisphere during an outbreak in the United States in 1999. Since this outbreak, WNV has spread throughout much of North America and has become a public health concern. Most WNV infections are asymptomatic; however, about 20% of cases are associated with mild flu-like symptoms. Some of these cases progress to more severe clinical manifestations, including encephalitis and/or flaccid paralysis.

35 In 1999, WNV emerged in the USA and has successfully spread across the entire country and into Canada, Mexico, and Central and South America. In 2007, the U.S. Centers for Disease Control reported 3630 clinical cases in the USA, with 2350 cases of West Nile fever, 1217 cases of meningitis or encephalitis, and 124 fatalities. Other regions at risk include Asia, Africa, Europe, and the Middle East.

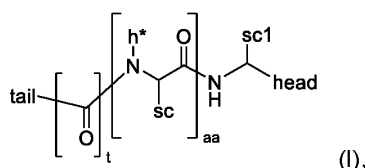
Some members of the genus Flavivirus are transmitted by a vector such as an insect, in many cases the insect is a mosquito. Flavivirus viruses are enveloped, positive-strand RNA viruses. The viral genome of the Flavivirus genus is translated as a single polyprotein and is subsequently cleaved into mature proteins. Both host signal peptidases and viral NS3 serine protease are involved in processing the polypeptide into viral proteins: structural proteins that form the virion particle, and non-structural proteins, that function in the virus life cycle. The viral NS3 protease has been shown to be required for viral replication, and provides a strategic target for inhibition in the development of flavivirus antivirals.

There is great demand for vaccines or antiviral therapeutics for Flaviviruses, such as Dengue or West Nile viruses. Currently, patients are treated with supportive care to relieve fever, pain, and dehydration. Attempts to treat West Nile disease with Ribavirin have been unsuccessful. Therefore, there exists a need for improved antiviral therapies, particularly to treat Flavivirus infections such as by dengue virus or West Nile virus. There is a need for compounds that reduce or hinder the progression of such infections. There is a need for compounds that interact with viral machinery. There is a need for compounds that reduce TNF- α or IL-6 levels associated with viral infections.

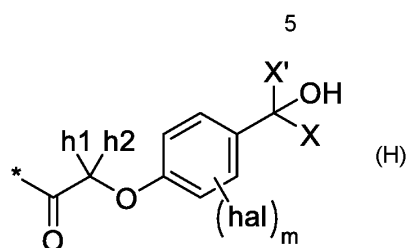
There is great demand for antiviral therapeutics for viruses, such as coronaviruses, particularly when these therapeutics are suitable for active infections. Currently, patients are treated prophylactically, or with supportive care to relieve symptoms. There exists a need for improved antiviral therapies, particularly to treat coronaviral infections such as by SARS-CoV, SARS-CoV-2, or MERS-CoV. There is a need for compounds that reduce or hinder the progression of such infections. There is a need for compounds that interact with coronaviral machinery. There is a need for compounds that interact with host machinery that is associated with progression of coronaviral infection. There is a need for improved administration of therapeutics for treatment of coronaviral infection.

Summary of the invention

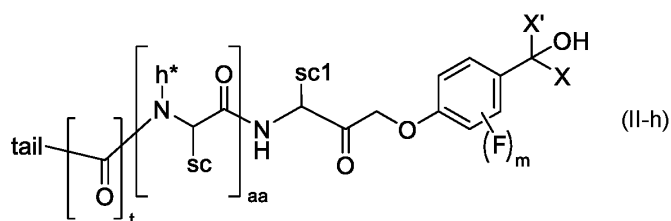
The invention provides a compound of general formula (I), or a salt thereof:



wherein **sc1** and each instance of **sc** are independently an amino acid side chain; **h*** is H or -CH₃ or -CH₂CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain; **aa** is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10; **t** is 0 or 1; **tail** is a linear, branched, or cyclic C1-22 alkyl, alkoxy, or alkylamine, wherein the alkyl, alkoxy, or alkylamine is optionally unsaturated and optionally substituted with halogen, alkylamine, alkylamide, aminoalkylamine, aminoalkylamide, alkoxy, or haloalkoxy, or is an independently chosen instance of **sc1**; **head** is a warhead of general formula (H)



Wherein **h1** and **h2** are each independently H, halogen, or C1-4 alkyl, or **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure; **hal** is a halogen; **m** is 0, 1, 2, 3, or 4; **X** and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure. In preferred embodiments **h1** and **h2** are H; **hal** is F; **m** is 3 or 4, preferably 4; and/or **X** and **X'** are each independently H or C1-2 alkyl, preferably H or -CH₃. Preferably the compound is of general formula (II-h)



wherein **m** is 1, 2, 3, or 4; **X** and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-4 membered cyclic structure, preferably **X** and **X'** are each independently H or -CH₃.

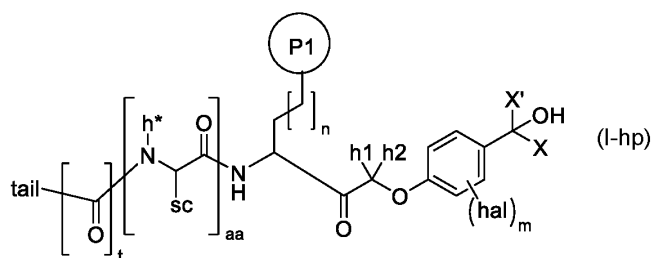
Preferably each instance of **sc** is independently a linear, branched, or cyclic C1-12 alkyl, wherein alkyl is optionally unsaturated and optionally substituted with halogen, or -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, or **sc** forms a cyclic amino acid side chain together with its adjacent **h***, wherein the cyclic amino acid side chain together with the carbon and nitrogen atoms to which it is attached forms a 4-7 membered ring.

Preferably **tail** is a linear or branched C1-20 alkyl or alkoxy, wherein the alkyl or alkoxy is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, alkyl, alkoxy, or -C(=O)O(CH₂)₀₋₄H.

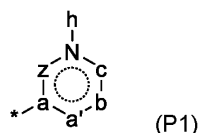
Preferably **aa** is 1 or 2, preferably 1; **t** is 0 or 1, preferably 1; **h*** is H or together with its adjacent **sc** forms a cyclic amino acid side chain; **sc** is butyl such as 2-methyl-n-propyl, is optionally substituted -CH₂-aryl such as -CH₂-fluorophenyl, or is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, wherein optional substitutions are preferably selected from halogen or alkoxy; and/or **tail** is linear C12-18 alkyl such as -C₁₅H₃₁, is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, more preferably methoxyl-1*H*-indol-3-yl, or is optionally substituted C4-11 alkoxy such as -O-CH₂-phenyl, wherein optional substitutions are preferably selected from halogen or alkoxy, preferably alkoxy, more preferably methoxyl.

In some embodiments the compound is of general formula (I-hp)

6

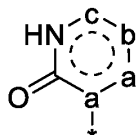


wherein **n** is 0 or 1; **P1** is a cyclic structure of general formula (P1):

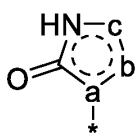


wherein **a** is C, CH, or N; **a'** is absent or is CH or -C(=O)- or N or NH; **b** is CH or N or NH; **c** is CH or -C(=O)- or N or NH; **h** is H or is absent; and **z** is -C(=O)- or N.

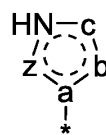
- 5 In some embodiments when **a** is CH, **n** is 0; when **a'** is CH, **n** is 0; when **a'** is absent and **z** is N, **n** is 0; when **a'** is absent, **n** is 1; or when **a'** is absent and **a** is N and **z** is -C(=O)-, **n** is 1; or wherein **a** is N or C. In some embodiments **P1** is selected from



(P1a)

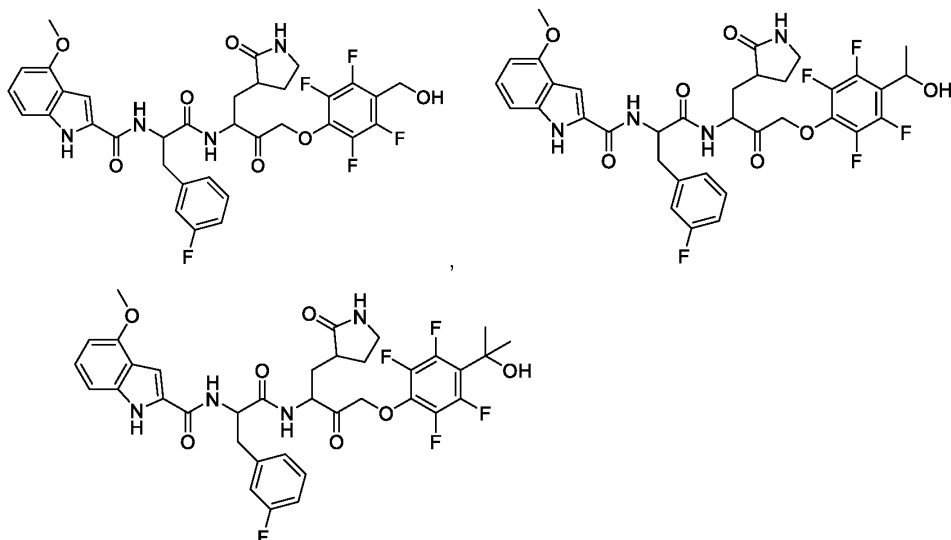


(P1b)



(P1c).

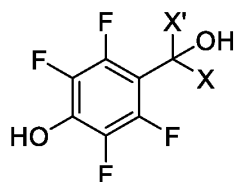
Preferably the compound is selected from



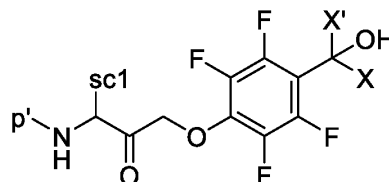
10

Also provided is a composition comprising a compound as defined above, and a pharmaceutically acceptable excipient, preferably wherein the composition is a pharmaceutical composition. Also provided is the compound or composition for use as a medicament.

Also provided is a compound according to general formula (intermediate-1) or (intermediate-2):



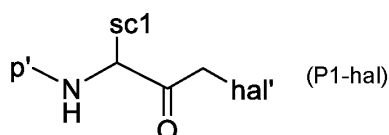
(intermediate-1)



(intermediate-2)

- wherein **p'** is H or a protecting group; **sc1** is an amino acid side chain; **X** is H, halogen, or C1-4 alkyl; **X'** is H, halogen, or C1-4 alkyl; or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure; wherein for the compound of general formula (intermediate-1) **X'** is halogen or C1-4 alkyl when **X** is H.

Also provided is a method for providing a compound as defined above, the method comprising the step of i) contacting a compound according to general formula (intermediate-1) as defined above with a compound of general formula (P1-hal):



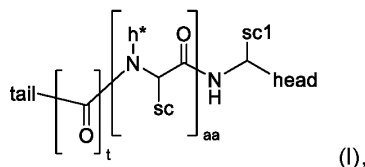
- wherein **p'** is a protecting group; **sc1** is an amino acid side chain; **hal'** a halogen, preferably chloride; to obtain a compound of general formula (intermediate-2) as defined above; and/or ii) removing the protecting group from a compound according to general formula (intermediate-2) as defined above, preferably by using a carboxylic acid such as trifluoroacetic acid.

- Also provided is a method for inhibiting a viral protease, the method comprising the steps of: i) providing a compound as defined herein, or a composition as herein; ii) contacting the viral protease with the provided compound or composition.

Description of embodiments

Compounds

- The inventors have identified new moieties that enable the design of new protease inhibitors, enabling new warheads and/or P1 residues. Accordingly the invention provides a compound of general formula (I), or a salt thereof:



wherein

- sc1** and each instance of **sc** are independently an amino acid side chain;
h* is H or -CH₃ or -CH₂CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain;

aa is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

t is 0 or 1;

tail is a linear, branched, or cyclic C1-22 alkyl, alkoxy, or alkylamine, wherein the alkyl, alkoxy, or alkylamine is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or is an independently chosen instance of **sc**¹;

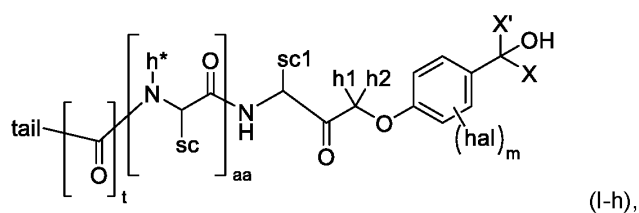
head is a warhead.

A salt is preferably a pharmaceutically acceptable salt. A salt is preferably a base addition salt wherein a cationic counterion is present. Examples of suitable salts are non-metallic salts such as ammonia salts, and metallic salts such as sodium salts and potassium salts. A skilled person can select suitable salt forms, and their means of production are well known (see e.g. "Occurrence of pharmaceutically acceptable anions and cations in the Cambridge Structural Database" Haynes et al., DOI: 10.1002/jps.20441). A salt can also be an acid addition salt. Acid addition salts are known in the art and examples are HCl salts and acetic acid salts. It should be recognized that the particular anion or cation forming a part of any salt of this invention is not critical, so long as the salt, as a whole, is pharmacologically acceptable. Additional examples of pharmaceutically acceptable salts and their methods of preparation and use are presented in Handbook of Pharmaceutical Salts: Properties, Selection and Use (2002). These salts can be prepared in situ during the isolation and purification of the compounds or by separately reacting the purified compound in its free base form with a suitable organic or inorganic acid and isolating the salt thus formed. Representative salts include the hydrobromide, hydrochloride, sulfate, bisulfate, nitrate, acetate, oxalate, valerate, oleate, palmitate, stearate, laurate, borate, benzoate, lactate, phosphate, tosylate, citrate, maleate, fumarate, succinate, tartrate, naphthylate mesylate, glucoheptonate, lactobionate, methane sulphonate, and laurylsulphonate salts, and the like. These may include cations based on the alkali and alkaline earth metals, such as sodium, lithium, potassium, calcium, magnesium, and the like, as well as non-toxic ammonium, quaternary ammonium, and amine cations including, but not limited to ammonium, tetramethylammonium, tetraethylammonium, methylamine, dimethylamine, trimethylamine, triethylamine, ethylamine, and the like. {See S.M. Barge et al., J. Pharm. Sci. (1977) 66, 1}.

The phrase "pharmaceutically-acceptable" or "pharmacologically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared.

In preferred embodiments a compound of general formula (I) is of general formula (I-h):

9



wherein

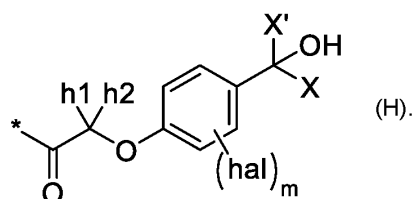
h1 and **h2** are each independently H, halogen, or C1-4 alkyl, or **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure;

hal is a halogen;

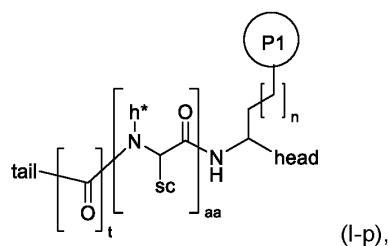
m is 0, 1, 2, 3, or 4;

X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure.

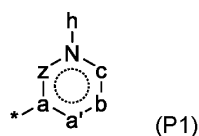
A compound of general formula (I-H) is a compound of general formula (I) wherein **head** is of general formula (H). Such compounds are preferred compounds. It should be understood that definitions for general formula (H) also apply to compounds that comprise general formula (H), when **head** is of that general formula.



In preferred embodiments a compound of general formula (I) is of general formula (I-p):



wherein **P1** is a cyclic structure of general formula (P1):



wherein **n** is 0 or 1; and

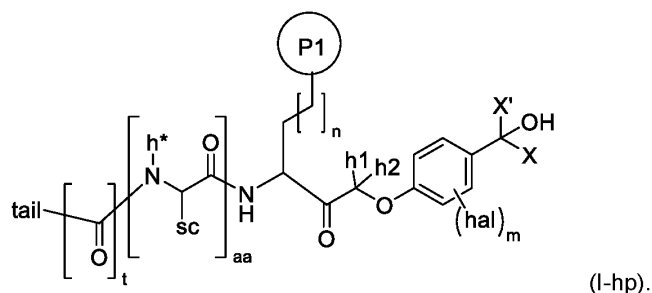
wherein

a is C, CH, or N;

a' is absent or is CH or -C(=O)- or N or NH;

b is CH or N or NH;
c is CH or -C(=O)- or N or NH;
h is H or is absent; and
z is -C(=O)- or N.

- 5 In preferred embodiments a compound of general formula (I) is of general formula (I-p) or of general formula (I-h). In more preferred embodiments a compound of general formula (I) is of general formula (I-hp):



In some embodiments **n** is 0. In some embodiments **n** is 1.

10

Central peptide analogue

- The moieties enclosed by the brackets labelled **aa**, along with the moiety comprising **sc1**, can be said to together form a peptide or an analogue thereof, and together are referred to as “the AAx” herein. The AAx represent a sequence and as represented in general formula (I) they can be considered to depict the traditional representation of peptides, with the N-terminus at the left and the C-terminus at the right. At its N-terminus, when **t** is 1, there is an amide bond that connects the AAx to **tail**. For individual reference, the amino acid or analogue thereof bearing **sc1** can be referred to as P1, the adjacent residue in **aa** brackets is P2, etc.

- aa** is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, and thus determines the length of the central peptide analogue as being 1 to 11. In preferred embodiments **aa** is 1, 2, 3, 4, 5, 6, 7, or 8 preferably **aa** is 1, 2, 3, 4, 5, or 6, more preferably 1, 2, 3, or 4, most preferably 1 or 2. In some embodiments **aa** is 1. In some embodiments **aa** is 2. In some embodiments **aa** is 2, 3, 4, 5, 6, 7, or 8, preferably 2, 3, 4, 5, or 6, more preferably 2, 3, or 4, even more preferably 2 or 3. In some embodiments **aa** is 3, 4, 5, 6, 7, or 8, more preferably 3, 4, 5, or 6, even more preferably 3 or 4. In some embodiments **aa** is 0 or 1. In some embodiments **aa** is 0. In some embodiments **aa** is 3. In some embodiments **aa** is 4. In some embodiments **aa** is 5. In some embodiments **aa** is 6. In some embodiments **aa** is 7. In some embodiments **aa** is 8.

- sc1** and each instance of **sc** are independently an amino acid side chain. Many amino acid side chains are commonly known. As for canonical amino acids except glycine, backbone carbon atoms can be further substituted. For instance when a backbone methylene carbon atom is substituted with a further methyl group, the canonical amino acid residue alanine is formed. The methyl group in this case is generally referred to as the amino acid side chain of alanine.

sc and **sc1** are for each instance independently H or optionally substituted and optionally unsaturated C2-6(halo)alkyl-N(**scH**)₂, C1-6(halo)alkyl, 5-10-membered (hetero)aryl, C1-4(halo)alkyl-[5-10-membered (hetero)aryl], C2-6(halo)alkyl-N(**scH**)C(N(**scH**)₂)(=N**scH**), C1-6(halo)alkyl-C(O)-N(**scH**)₂, or C1-4(halo)alkyl-[3-10-membered (hetero)cycloalkyl].

5 **scH** is for each instance independently H or -C1-3(halo)alkyl or -CH₂-(halo)phenyl; preferably -CH₂-(halo)phenyl is -CH₂-phenyl; preferably -C1-3(halo)alkyl is -CH₃ or -CH₂-CH₃; preferably **scH** is H or -CH₃, more preferably H.

In some embodiments **sc** and **sc1** are not optionally substituted and not optionally unsaturated; in some embodiments **sc** and **sc1** are optionally substituted and not optionally unsaturated; in some embodiments **sc** and **sc1** are not optionally substituted and is optionally unsaturated; particularly preferred optional substitutions are halogen, C1-3(halo)alkyl, or C1-3(halo)alkoxyl, more preferably halogen, -CH₃, and -O-CH₃; wherein within **sc** optional substitutions are preferably -OH, -SH, -SeH, -S-CH₃, -O-CH₃, and -COOH, more preferably -OH, -SH, and -S-CH₃, most preferably -OH;

15 wherein in some embodiments **sc** and **sc1** are not optionally substituted; wherein in some embodiments **sc** and **sc1** are not optionally unsaturated; wherein in some embodiments **sc** and **sc1** are not optionally substituted and not optionally unsaturated; wherein **sc** and **sc1** are preferably H, C2-5(halo)alkyl-N(**scH**)₂, C1-4(halo)alkyl, 5-9-membered (hetero)aryl, C1-2(halo)alkyl-[5-9-membered (hetero)aryl], C2-4(halo)alkyl-N(**scH**)C(N(**scH**)₂)(=N**scH**), C1-4(halo)alkyl-C(O)-N(**scH**)₂, or C1-2(halo)alkyl-[3-9-membered (hetero)cycloalkyl];

20 wherein C2-6(halo)alkyl-N(**scH**)₂ is preferably -CH₂-CH₂-CH₂-NH₂ or -CH₂-CH₂-CH₂-CH₂-NH₂; wherein 5-10-membered (hetero)aryl is preferably phenyl; wherein optionally substituted C1-4(halo)alkyl-[5-10-membered (hetero)aryl] is preferably -CH₂-phenyl, -CH₂-CH₂-phenyl, -CH₂-imidazolyl, -CH₂-CH₂-imidazolyl, -CH₂-indolyl, -CH₂-CH₂-indolyl; -CH₂-hydroxyphenyl, -CH₂-CH₂-hydroxyphenyl; wherein optionally substituted C1-6(halo)alkyl is preferably -CH₃, -CH(CH₃)₂, -CH₂-CH(CH₃)₂, -CH(CH₃)-CH₂-CH₃, -CH₂-OH, -CH₂-SH, -CH₂-SeH, -CH₂-CH₂-CH₂-S-CH₃, -CH(CH₃)-CH₂-OH, -CH₂-CH₂-COOH, or -CH₂-COOH, more preferably -CH₃, -CH(CH₃)₂, -CH₂-CH(CH₃)₂, -CH(CH₃)-CH₂-CH₃, -CH₂-OH, -CH₂-SH, -CH₂-SeH, -CH₂-CH₂-CH₂-S-CH₃, or -CH₂(CH₃)-CH₂-OH; wherein C2-6(halo)alkyl-N(**scH**)C(N(**scH**)₂)(=N**scH**) is preferably -CH₂-CH₂-CH₂-N-C(=NH)-NH₂; wherein C1-6(halo)alkyl-C(O)-N(**scH**)₂ is preferably -CH₂-CH₂-C(O)NH₂ or -CH₂-C(O)NH₂;

30 preferably wherein AAx form optionally substituted and optionally unsaturated alanine, cysteine, aspartic acid, glutamic acid, phenylalanine, glycine, histidine, isoleucine, lysine, leucine, methionine, asparagine, ornithine, pyrrolysine, proline, pipecolic acid, glutamine, arginine, serine, 35 threonine, selenocysteine, valine, tryptophan, tyrosine, homophenylalanine, homohistidine, para-guanidiny-phenylalanine, meta-guanidiny-phenylalanine, (2-amino-1H-imidazol-4-yl)ethyl-alanine, (2-amino-1H-imidazol-4-yl)methyl-alanine, (2-amino-1H-imidazol-4-yl)-alanine, (pyridin-3-yl)homoalanine, (pyridin-3-yl)alanine, (pyridin-4-yl)alanine, (1H-imidazol-2-yl)aminomethyl-alanine, (1H-imidazol-2-yl)aminoethyl-alanine, (pyridin-4-yl)homoalanine, phenylglycine, (4-O-benzyl)phenylglycine, (4-O-4-pyridinyl)proline, pipecolic acid, or homo-analogues or nor-

analogues thereof; more preferably wherein amino acid residues are optionally substituted and optionally unsaturated alanine, cysteine, aspartic acid, glutamic acid, phenylalanine, glycine, histidine, isoleucine, lysine, leucine, methionine, asparagine, ornithine, pyrrolysine, proline, pipecolic acid, glutamine, arginine, serine, threonine, selenocysteine, valine, tryptophan, or tyrosine, or homo-analogues or nor-analogues thereof; in some embodiments these residues are not optionally substituted and not optionally unsaturated; in some embodiments these residues are optionally substituted and not optionally unsaturated; in some embodiments these residues are not optionally substituted and is optionally unsaturated; particularly preferred optional substitutions are halogen, C1-3(halo)alkyl, or C1-3(halo)alkoxyl, more preferably halogen, -CH₃, and -O-CH₃;

more preferably wherein AAx are optionally substituted and optionally unsaturated lysine, ornithine, arginine, histidine, phenylalanine, alanine, glutamine, tryptophan, proline, or valine, or homo-analogues or nor-analogues thereof such as homophenylalanine, homohistidine, pipecolic acid, and phenylglycine; as described above;

wherein optional substitutions are preferably halogen, C1-3(halo)alkyl, C1-3(halo)alkoxy, guanidinyl (which is -NH-C(=NH)-NH₂), or optionally unsaturated and optionally (halo)methylated -(O)₀₋₁-(CH₂)₀₋₁-5-6-membered (hetero)cycloalkyl such as phenyl or benzyl or imidazolyl or -O-pyridinyl or methyl-dihydropyrrolyl.

In preferred embodiments each instance of **sc** is independently a linear, branched, or cyclic C1-12 alkyl, wherein alkyl is optionally unsaturated and optionally substituted with halogen, or -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, or **sc** forms a cyclic amino acid side chain together with its adjacent **h***, wherein the cyclic amino acid side chain together with the carbon and nitrogen atoms to which it is attached forms a 4-7 membered ring. In other preferred embodiments each **sc** is independently a linear, branched, or cyclic C1-12 alkyl, wherein alkyl is optionally unsaturated and optionally substituted with halogen, or -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen.

In preferred embodiments **sc** is butyl such as 2-methyl-n-propyl, is optionally substituted -CH₂-aryl such as -CH₂-fluorophenyl, or is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, wherein optional substitutions are preferably selected from halogen of alkoxy.

h* is H or -CH₃ or -CH₂CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain. Preferably **h*** is H or -CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain. More preferably **h*** is H or -CH₃ or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain. Even more preferably **h*** is H or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain. Most preferably **h*** is H.

Cyclic amino acid side chains are commonly known, and examples of amino acids featuring cyclic side chains are proline and pipecolic acid. When **h*** together with its adjacent **sc** forms a cyclic amino acid side chain, it preferably forms a 3-9-membered cycle, more preferably a 4-7-membered cycle, still more preferably a 5-6-membered cycle. Preferably such a cycle is saturated. In some embodiments such a cycle is not further substituted. Such a cycle can be

fused to a further cycle, which is preferably 3-4-membered, most preferably 3-membered. When comprised in a cyclic amino acid side chain, **h*** and **sc** can together represent a bridging moiety. Suitable bridging moieties are $-(CH_2)_2-$, $-(CH_2)_3-$, $-(CH_2)_4-$, $-(CH_2)_5-$, and $-CH_2-CH(b^*)-CH(b^*)-$ wherein two instances of **b*** together form a bridging moiety $-C(CH_3)_2-$. Preferred bridging moieties are $-(CH_2)_2-$, $-(CH_2)_3-$, and $-(CH_2)_4-$, and $-CH_2-CH(b^*)-CH(b^*)-$, most preferably $-(CH_2)_3-$ and $-(CH_2)_4-$ and $-CH_2-CH(b^*)-CH(b^*)-$. With $-(CH_2)_3-$ the AAx can be described as proline and with $-(CH_2)_4-$ pipecolic acid. Compounds comprising such a and $-CH_2-CH(b^*)-CH(b^*)-$ can be of interest for showing attractive affinity for cathepsin.

AAx are for each independent instance preferably optionally substituted and optionally unsaturated alanine, cysteine, aspartic acid, glutamic acid, phenylalanine, glycine, histidine, isoleucine, lysine, leucine, methionine, asparagine, ornithine, pyrrollysine, proline, pipecolic acid, glutamine, arginine, serine, threonine, selenocysteine, valine, tryptophan, or tyrosine, or homo-analogues or nor-analogues thereof; in some embodiments these residues are not optionally substituted and not optionally unsaturated; in some embodiments these residues are not optionally substituted and not optionally unsaturated; in some embodiments these residues are not optionally substituted and is optionally unsaturated; particularly preferred optional substitutions are halogen, C1-3(halo)alkyl, or C1-3(halo)alkoxyl, more preferably halogen, $-CH_3$, and $-O-CH_3$. In some embodiments each individual instance of AAx is independently uncharged or positively charged under physiological conditions. More preferably each individual instance is independently lysine, arginine, phenylalanine, histidine, alanine, valine, leucine, or tryptophan, still more preferably it is lysine, arginine, or histidine, even more preferably it is lysine.

In preferred embodiments, AAx at P3 is lysine, arginine, glutamine, asparagine, alanine, phenylalanine, proline, histidine, alanine, valine, leucine, or tryptophan, still more preferably it is lysine, arginine, glutamine, asparagine, alanine, proline or histidine, even more preferably it is lysine, arginine, glutamine, alanine, asparagine, or, proline, even more preferably it is proline, arginine, alanine, or glutamine, even more preferably it is alanine, arginine, glutamine, most preferably it is alanine or glutamine.

In preferred embodiments, AAx at P2, when present, is lysine, arginine, phenylalanine, histidine, alanine, valine, leucine, or tryptophan, still more preferably it is lysine, alanine or phenylalanine, even more preferably it is lysine.

In some embodiments the AAx that comprises **sc1** (position P1) is uncharged or positively charged under physiological conditions, more preferably it is positively charged under physiological conditions. Preferably, none of the AAx are negatively charged under physiological conditions. Preferably, at most 4 of the AAx are positively charged under physiological conditions, more preferably at most 3 or the AAx positively charged under physiological conditions, most preferably 2 of the AAx are positively charged under physiological conditions. When one or more of the AAx are positively charged under physiological conditions, it is preferred that AAx at P1 is positively charged under physiological conditions. When two or more of the AAx are positively charged under physiological conditions, it is preferred that AA^x adjacent to **tail** and the AAx at P1

are positively charged under physiological conditions. In preferred embodiments **sc1** is not $-(CH_2)_4-NHBoc$ or $-(CH_2)_4-NH_2$ or $-(CH_2)_4-NH-Alloc$ (Alloc is allyloxycarbonyl, $-C(=O)O-CH_2-CH=CH_2$). In preferred embodiments **sc1** is not $-(CH_2)_4-NH_2$.

5 Amino acid residues as described herein can be L or D or a mixture thereof. Preferably residues are L, more preferably all residues within an individual embodiment are L. Analogues of residues that are known in the art can also be used, and additional such residues that can be represented by AAX are shown below, where residues can be shown as comprised in a peptide, or as comprised in H-AAX-OH for the free residue, or can be shown as the free residue.

10

In preferred embodiments **aa** is 2 and the AAX together form a tripeptide represented by KAK, KAH, KAF, KA-homoPhe, KAA, KPA, KPK, FAF, HAH, RQK, RQF, RAF, RQ-homoPhe, HQF, homoHis-QF, (para-guanidinyI-Phe)-QF, (meta-guanidinyI-Phe)-QF, ((2-amino-1H-imidazol-4-yl)ethyl-Ala)-QF, ((2-amino-1H-imidazol-4-yl)methyl-Ala)-QF, ((2-amino-1H-imidazol-4-yl)-Ala)-QF, ((pyridin-3-yl)homoAla)-QF, ((pyridin-3-yl) Ala)-QF, ((pyridin-4-yl) Ala)-QF, ((1H-imidazol-2-yl)aminomethyl-Ala)-QF, ((1H-imidazol-2-yl)aminoethyl-Ala)-QF, ((pyridin-4-yl)homoAla)-QF, or RNF. In some embodiments the tripeptide is represented by tripeptide represented by KAK, KAH, KAF, KA-homoPhe, KAA, KPA, KPK, FAF, HAH, RQK, RNF, or RQF. In some embodiments the tripeptide is represented by KAK, KAH, KAF, KA-homoPhe, KAA, KPA, KPK, FAF, HAH, RQK, RQF, RAF, RQ-homoPhe, HQF, homoHis-QF, or RNF. In some embodiments the tripeptide is represented by KAK, KAH, KAF, KA-homoPhe, KAA, KPA, or KPK. In some embodiments the tripeptide is represented by KAK, KAH, KAF, KA-homoPhe, KAA, FAF, or HAH. In some embodiments the tripeptide is represented by KAK, KAH, KPK, HAH, or RQK. In some embodiments the tripeptide is represented by RQK, RQF, RMF, RAF, KAH, or RQ-homoPhe. In some embodiments the tripeptide is represented by FAF, HAH, RQK, or RQF. In some embodiments the tripeptide is represented by RQK or RQF

In preferred embodiments **aa** is 3 and the AAX together form a tetrapeptide represented by KAAA, KAAF, KAAH, KAAK, KAAW, KAA-homoHis, KAA-homoPhe, KAA-phenylGly, KAA-(4-OBn-phenylGly),

30 KPAF, KPAH, KPAK, KPAW, KPA-homoHis, KPA-homoPhe, KPA-phenylGly, KPA-(4-OBn-phenylGly),

AAAK, AAKA, AKAA, AKAK, APAK, A-(4-O-(4-Py)-Pro)-AK, AVAK, AKKK, LKAK, LKKK, VKAK,

35 KPHK, KPH-homoPhe, K-(pipecolic acid)-AK, K-(pipecolic acid)-HK, KP-phenylglycine-K, K-(pipecolic acid)-phenylglycine-K,

ARQK, ARQF, ARRK, FKKK, RARK, FAAF, or WAAW.

In some embodiments the tetrapeptide is represented by KPAF, KPAH, KPAK, KPAW, KPA-homoHis, KPA-homoPhe, KPA-phenylGly, KPA-(4-OBn-phenylGly), AAAK, AAKA, AKAA, AKAK, APAK, A-(4-O-(4-Py)-Pro)-AK, AVAK, AKKK, LKAK, LKKK, VKAK, KPHK, KPH-

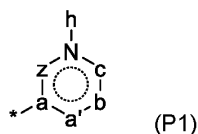
40

homoPhe, K-(pipecolic acid)-AK, K-(pipecolic acid)-HK, KP-phenylglycine-K, K-(pipecolic acid)-phenylglycine-K, ARQK, ARQF, ARRK, FKKK, RARK, FAAF, or WAAW. In some embodiments the tetrapeptide is represented by KAAA, KAAF, KAAH, KAAK, KAAW, KAA-homoHis, KAA-homoPhe, KAA-phenylGly, KAA-(4-OBn-phenylGly), AAAK, AAKA, AKAA, AKAK, APAK, A-(4-O-(4-Py)-Pro)-AK, AVAK, AKKK, LKAK, LKKK, VKAK, KPHK, KPH-homoPhe, K-(pipecolic acid)-AK, K-(pipecolic acid)-HK, KP-phenylglycine-K, K-(pipecolic acid)-phenylglycine-K, ARQK, ARQF, ARRK, FKKK, RARK, FAAF, or WAAW. In some embodiments the tetrapeptide is represented by KAAA, KAAF, KAAH, KAAK, KAAW, KAA-homoHis, KAA-homoPhe, KAA-phenylGly, KAA-(4-OBn-phenylGly), KPAF, KPAH, KPAK, KAPK, KPAW, KPA-homoHis, KPA-homoPhe, KPA-phenylGly, KPA-(4-OBn-phenylGly), KPHK, KPH-homoPhe, K-(pipecolic acid)-AK, K-(pipecolic acid)-HK, KP-phenylglycine-K, K-(pipecolic acid)-phenylglycine-K, ARQK, ARQF, ARRK, FKKK, RARK, FAAF, or WAAW. In some embodiments the tetrapeptide is represented by KAAA, KAAF, KAAH, KAAK, KAAW, KAA-homoHis, KAA-homoPhe, KAA-phenylGly, KAA-(4-OBn-phenylGly), KPAF, KPAH, KPAK, KAPK, KPAW, KPA-homoHis, KPA-homoPhe, KPA-phenylGly, KPA-(4-OBn-phenylGly), AAAK, AAKA, AKAA, AKAK, APAK, A-(4-O-(4-Py)-Pro)-AK, AVAK, AKKK, LKAK, LKKK, VKAK, ARQK, ARQF, ARRK, FKKK, RARK, FAAF, or WAAW. In some embodiments the tetrapeptide is represented by KAAA, KAAF, KAAH, KAAK, KAAW, KAA-homoHis, KAA-homoPhe, KAA-phenylGly, KAA-(4-OBn-phenylGly), KPAF, KPAH, KPAK, KAPK, KPAW, KPA-homoHis, KPA-homoPhe, KPA-phenylGly, KPA-(4-OBn-phenylGly), AAAK, AAKA, AKAA, AKAK, APAK, A-(4-O-(4-Py)-Pro)-AK, AVAK, AKKK, LKAK, LKKK, VKAK, KPHK, KPH-homoPhe, K-(pipecolic acid)-AK, K-(pipecolic acid)-HK, KP-phenylglycine-K, or K-(pipecolic acid)-phenylglycine-K.

In some embodiments the tetrapeptide is represented by KAAA, KAAF, KAAH, KAAK, KAAW, KAA-homoHis, KAA-homoPhe, KAA-phenylGly, KAA-(4-OBn-phenylGly), KPAF, KPAH, KPAK, KAPK, KPAW, KPA-homoHis, KPA-homoPhe, KPA-phenylGly, KPA-(4-OBn-phenylGly), KPHK, KPH-homoPhe, K-(pipecolic acid)-AK, K-(pipecolic acid)-HK, KP-phenylglycine-K, or K-(pipecolic acid)-phenylglycine-K. In some embodiments the tetrapeptide is represented by AAAK, AAKA, AKAA, AKAK, APAK, A-(4-O-(4-Py)-Pro)-AK, AVAK, AKKK, LKAK, LKKK, VKAK, ARQK, ARQF, ARRK, FKKK, RARK, FAAF, or WAAW.

The above tetrapeptides have the following SEQ ID NOs (in parentheses following the sequence): KAAA (1); KAAF (2); KAAH (3); KAAK (4); KAAW (5); KAA-homoHis (6); KAA-homoPhe (7); KAA-phenylGly (8); KAA-(4-OBn-phenylGly) (9); KPAF (10); KPAH (11); KPAK (12); KAPK (13); KPAW (14); KPA-homoHis (15); KPA-homoPhe (16); KPA-phenylGly (17); KPA-(4-OBn-phenylGly) (18); AAAK (19); AAKA (20); AKAA (21); AKAK (22); APAK (23); A-(4-O-(4-Py)-Pro)-AK (24); AVAK (25); AKKK (26); LKAK (27); LKKK (28); VKAK (29); KPHK (30); KPH-homoPhe (31); K-(pipecolic acid)-AK (32); K-(pipecolic acid)-HK (33); KP-phenylglycine-K (34); K-(pipecolic acid)-phenylglycine-K (35); ARQK (36); ARQF (37); ARRK (38); FKKK (39); RARK (40); FAAF (41); WAAW (42).

In highly preferred embodiments **sc1** comprises general formula (P1):



wherein

a is C, CH, or N;

5 **a'** is absent or is CH or -C(=O)- or N or NH;

b is CH or N or NH;

c is CH or -C(=O)- or N or NH;

h is H or is absent; and

10 **z** is -C(=O)- or N. Such a compound is preferably of general formula (I-p) as described above. More preferably it is of general formula (I-hp).

Within general formula (P1), in preferred embodiments

when **a** is CH, **n** is 0;

when **a'** is CH, **n** is 0;

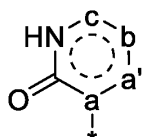
when **a'** is absent and **z** is N, **n** is 0;

15 when **a'** is absent, **n** is 1; or

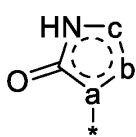
when **a'** is absent and **a** is N and **z** is -C(=O)-, **n** is 1;

or **a** is N or C.

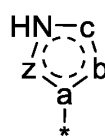
(P1) is preferably linked to the remainder of the compound via -(CH₂)₁₋₄-, preferably -(CH₂)₁₋₃-, more preferably -(CH₂)₁₋₂- such as -(CH₂)₁- or -(CH₂)₂-. In preferred embodiments of
20 compounds comprising general formula (P1), **a** is N or C. Further preferred embodiments of general formula (P1) are (P1a), (P1b), and (P1c). In some embodiments (P1) is of general formula (P1a) or (P1b). In some embodiments (P1) is of general formula (P1a) or (P1c). In some embodiments (P1) is of general formula (P1c) or (P1b).



(P1a)

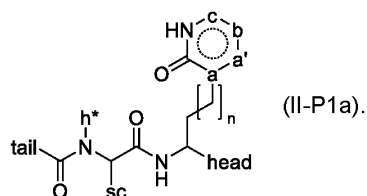


(P1b)



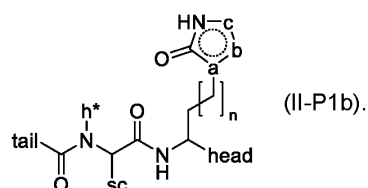
(P1c).

In preferred embodiments the compound is of general formula (II-P1a):

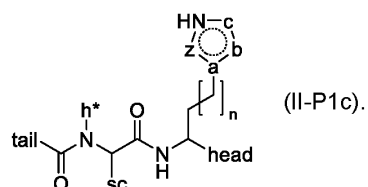


25 In preferred embodiments the compound is of general formula (II-P1b):

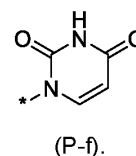
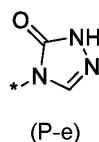
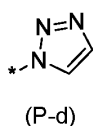
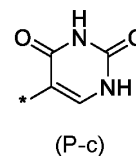
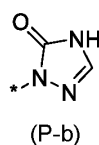
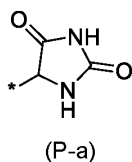
17



In preferred embodiments the compound is of general formula (II-P1c):

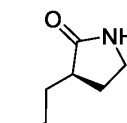
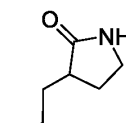
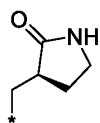
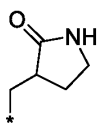


Particularly preferred embodiments of general formula (P1) are:

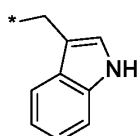
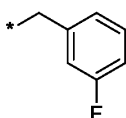


- In some embodiments general formula (P1) is (P-a), (P-c), (P-e), or (P-f). In some embodiments general formula (P1) is (P-a), (P-b), (P-d), or (P-e). In some embodiments general formula (P1) is (P-c) or (P-f). In some embodiments general formula (P1) is (P-b) or (P-e). In some embodiments general formula (P1) is (P-a), (P-c), or (P-f). Preferably, **n** is 0 when **P1** is (P-a), (P-c), (P-d), or (P-f). Preferably **n** is 1 when **P1** is (P-b) or (P-e).

Further preferred **sc1** are shown below, wherein these **sc1** are particularly preferred when head is of general formula (H):



- Further preferred **sc** are shown below, wherein these **sc** are particularly preferred when head is of general formula (H):



The two embodiments on the right are suitable for when **sc** forms a bridging moiety with **h***. In these cases it is preferred that the * depicted on the left be linked to **h*** and the * depicted on the right be linked to **sc**.

Tail

The **tail** moiety of the compounds is found at the N-terminus of the central peptide analogues as defined above. When **t** is 1, the **tail** is connected via a carbonyl moiety as depicted in general formula (I), which thus forms an amide bond with the backbone amine of the N-terminal instance of **aa**. Compounds wherein **t** is 1 can be more readily synthetically accessible. Compounds wherein **t** is 0 have tail directly attached to the backbone amine of the N-terminal instance of **aa**, which can make the compounds more stable. In preferred embodiments **t** is 1. In some embodiments **t** is 0. The **tail** is generally hydrophobic and can be a long aliphatic moiety, although it can vary as defined.

tail is a linear, branched, or cyclic C1-22 alkyl, alkoxy, or alkylamine, wherein the alkyl, alkoxy, or alkylamine is optionally unsaturated and optionally substituted with halogen, alkylamine, alkylamide, aminoalkylamine, aminoalkylamide, alkoxy, or haloalkoxy, or is an independently chosen instance of **sc**¹; optionally **tail** is C1-24alkyl, -O-C1-24alkyl, 5-20-membered (hetero)aryl, or 3-20-membered (hetero)cycloalkyl, wherein **tail** is optionally unsaturated, wherein **tail** is optionally substituted with halogen, C1-3(halo)alkyl, or C1-3(halo)alkoxy; in some embodiments **tail** is not optionally unsaturated and not optionally substituted; in some embodiments **tail** is not optionally unsaturated; in some embodiments **tail** is not optionally substituted; preferred optional substitutions are -F, -Cl, -CH₃, and -OCH₃, or aminoalkylamide as defined later herein, more preferably -CH₃ and -OCH₃, most preferably -OCH₃. **tail** preferably comprises at least 8 carbon atoms, more preferably at least 10, even more preferably at least 13, most preferably at least 15. A highly preferred tail is pentadecyl (which is - (CH₂)₁₄CH₃) thus forming a palmitic acid amide with the carbonyl moiety of general formula 1 and the N-terminus of the central peptide analogue.

In **tail** C1-24alkyl is preferably C1-22 alkyl, which is preferably C1-20 alkyl, more preferably C2-19 alkyl, even more preferably C6-C15 alkyl or C4-15 alkyl, wherein alkyl is preferably straight and unbranched. The alkyl preferably comprises at least 2 carbon atoms, more preferably at least 4, more preferably at least 6, even more preferably at least 9, more preferably at least 13, most preferably at least 15. The alkyl preferably comprises at most 21 carbon atoms, more preferably at most 19, even more preferably at most 17, more preferably at most 15. Another preferred embodiment is where tail is CH₃. In some embodiments shorter chains are preferred for **tail**. In such embodiments C1-24alkyl is preferably C1-16 alkyl, more preferably C1-14 alkyl, still more preferably C1-12 alkyl, still more preferably C1-10alkyl, still more preferably C1-8alkyl, still more preferably C1-6alkyl, still more preferably C1-4alkyl. In other preferred embodiments where shorter tails are preferred, C1-24alkyl is preferably C2-16 alkyl, more preferably C4-14 alkyl, still more preferably C6-12 alkyl, still more preferably C6-10alkyl, still more preferably C4-8alkyl, still more preferably C2-6alkyl, still more preferably C2-4alkyl.

Alkoxy is generally -O-alkyl, preferably -O-C1-24alkyl. In **tail** -O-C1-24alkyl is preferably -O-C1-20 alkyl, more preferably -O-C2-19 alkyl, even more preferably -O-C2-C8 alkyl, wherein alkyl is preferably branched. The -O-alkyl preferably comprises at least 2 carbon atoms, more preferably at least 3, even more preferably at least 4. The -O-alkyl preferably comprises at most

20 carbon atoms, more preferably at most 16, even more preferably at most 10, more preferably at most 4. Highly preferred -O-alkyl are -O-CH₃ and -O-C(CH₃)₃, or which the latter is most preferred. In some embodiments shorter chains are preferred for **tail**. In such embodiments -O-C1-24alkyl is preferably -O-C1-16 alkyl, more preferably -O-C1-14 alkyl, still more preferably -O-C1-12 alkyl, still more preferably -O-C1-10alkyl, still more preferably -O-C1-8alkyl, still more preferably -O-C1-6alkyl, still more preferably -O-C1-4alkyl. In other preferred embodiments where shorter tails are preferred, -O-C1-24alkyl is preferably -O-C2-16 alkyl, more preferably -O-C4-14 alkyl, still more preferably -O-C6-12 alkyl, still more preferably -O-C6-10alkyl, still more preferably -O-C4-8alkyl, still more preferably -O-C2-6alkyl, still more preferably -O-C2-4alkyl.

Alkylamine is generally -NH-alkyl, preferably -NH-C1-24alkyl. In **tail** -NH-C1-24alkyl is preferably -NH-C1-20 alkyl, more preferably -NH-C2-19 alkyl, even more preferably -NH-C2-C8 alkyl, wherein alkyl is preferably branched. The -NH-alkyl preferably comprises at least 2 carbon atoms, more preferably at least 3, even more preferably at least 4. The -NH-alkyl preferably comprises at most 20 carbon atoms, more preferably at most 16, even more preferably at most 10, more preferably at most 4. Highly preferred -NH-alkyl are -NH-CH₃ and -NH-C(CH₃)₃, or which the latter is most preferred. In some embodiments shorter chains are preferred for **tail**. In such embodiments -NH-C1-24alkyl is preferably -NH-C1-16 alkyl, more preferably -NH-C1-14 alkyl, still more preferably -NH-C1-12 alkyl, still more preferably -NH-C1-10alkyl, still more preferably -NH-C1-8alkyl, still more preferably -NH-C1-6alkyl, still more preferably -NH-C1-4alkyl. In other preferred embodiments where shorter tails are preferred, -NH-C1-24alkyl is preferably -NH-C2-16 alkyl, more preferably -NH-C4-14 alkyl, still more preferably -NH-C6-12 alkyl, still more preferably -NH-C6-10alkyl, still more preferably -NH-C4-8alkyl, still more preferably -NH-C2-6alkyl, still more preferably -NH-C2-4alkyl. Aminoalkylamine is preferably as defined for alkylamine, with an -NH₂ or -NHp' moiety replacing a hydrogen atom at the carbon atom that is at the distant terminus of **tail**. The distant terminus is the carbon atom that is most distant from the remainder of the inhibitor.

Alkylamide is generally -NHC(=O)-alkyl, preferably -NHC(=O)-C1-23alkyl. In **tail** -NHC(=O)-C1-23alkyl is preferably -NHC(=O)-C1-19 alkyl, more preferably -NHC(=O)-C1-18 alkyl, even more preferably -NHC(=O)-C1-C7 alkyl, wherein alkyl is preferably not branched or cyclic. The alkylamide preferably comprises at least 2 carbon atoms, more preferably at least 3, even more preferably at least 4. The alkylamide preferably comprises at most 20 carbon atoms, more preferably at most 16, even more preferably at most 10, more preferably at most 6. Highly preferred alkylamides are -NHC(=O)-CH₃ and -NHC(=O)-(CH₂)₄CH₃, or which the latter is most preferred. Aminoalkylamide is preferably as defined for alkylamide, with an -NH₂ or -NHp' moiety replacing a hydrogen atom at the carbon atom that is at the distant terminus of **tail**. In preferred embodiments, **tail** is substituted with aminoalkylamide, more preferably with -NHC(=O)-(CH₂)₄CH₂-NHp', most preferably with -NHC(=O)-(CH₂)₄CH₂-NH-C(=O)O-C(CH₃)₃. A preferred such **tail** is -(CH₂)₅-NHC(=O)-(CH₂)₄CH₂-NH-C(=O)O-C(CH₃)₃, wherein t is preferably 1. Comparable preferred embodiments of **tail** are -(CH₂)₃₋₇-NHC(=O)-(CH₂)₃₋₇-NHp'.

Unsaturated cyclic alkyl can be (hetero)aryl. In **tail** 5-20-membered (hetero)aryl is preferably 5-12 membered, or 5-10 membered. It preferably comprises at least one 5-6 membered (hetero)aryl ring, and optionally it further comprises a second 5-6 membered (hetero)aryl ring. In some embodiments the (hetero)aryl is aryl. In some embodiments the (hetero)aryl is heteroaryl. Preferred 5-20-membered (hetero)aryl are optionally substituted phenyl, optionally substituted thiophene, and optionally substituted indole. A preferred substitution of phenyl herein is a second phenyl, preferably forming biphenyl, which is preferably linked to the compound of general formula (I) at a 4-position. A preferred substitution of thiophene herein is a second thiophene, preferably forming 2-2'-thiophene, which is preferably linked to the compound of general formula (I) at a 5-position. Indole is preferably linked to the compound of general formula (I) at its 2-position, and a preferred substitution of indole is -OCH₃, preferably forming -[4-methoxy-1H-indole-2-yl].

In **tail** 3-20-membered (hetero)cycloalkyl is preferably 3-15 membered, more preferably 3-10 membered, even more preferably 5-10 membered, most preferably 5-6 membered or 10-membered. Preferred examples of 3-20-membered (hetero)cycloalkyl are cyclohexyl and adamantanyl, preferably adamantan-2-yl.

In preferred embodiments **tail** is C1-20alkyl such as methyl or ethyl or butyl or hexanyl or heptyl or nonanyl or undecanyl or tridecanyl or pentadecanyl or nonadecanyl, -O-C1-16alkyl such as -O-butyl, 3-12-membered (hetero)cycloalkyl such as adamantanyl or cyclohexyl, or 5-12-membered (hetero)aryl such as phenyl or biphenyl or bithiophene or substituted indole such as methoxyindole. More preferably **tail** is C1-20alkyl such as methyl or ethyl or butyl or hexanyl or heptyl or nonanyl or undecanyl or tridecanyl or pentadecanyl or nonadecanyl, 3-12-membered (hetero)cycloalkyl such as adamantanyl or cyclohexyl, or 5-12-membered (hetero)aryl such as phenyl or biphenyl or bithiophene or substituted indole such as methoxyindole. Even more preferably **tail** is C1-20alkyl such as methyl or ethyl or butyl or hexanyl or heptyl or nonanyl or undecanyl or tridecanyl or pentadecanyl or nonadecanyl or 5-12-membered (hetero)aryl such as phenyl or biphenyl or bithiophene or substituted indole such as methoxyindole.

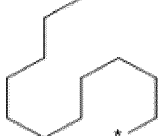
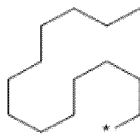
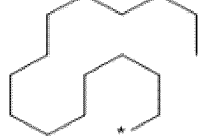
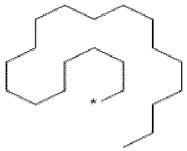
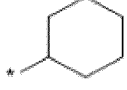
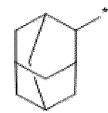
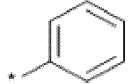
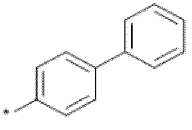
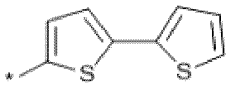
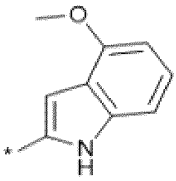
In preferred embodiments, wherein **tail** is a linear or branched C1-20 alkyl or alkoxy, wherein the alkyl or alkoxy is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, alkyl, alkoxy, or -C(=O)O(CH₂)₀₋₄H. Further preferred instances of **tail** are linear C12-18 alkyl such as -C₁₅H₃₁, is optionally substituted heteroaryl such as indolyl, preferably 1H-indol-3-yl, more preferably methoxyl-1H-indol-3-yl, or is optionally substituted C4-11 alkoxy such as -O-CH₂-phenyl, wherein optional substitutions are preferably selected from halogen or alkoxy, preferably alkoxy, more preferably methoxy.

In other preferred embodiments, **tail** is -CH₃, -CH₂CH₃, -(CH₂)₅CH₃, -(CH₂)₆CH₃, -(CH₂)₈CH₃, -(CH₂)₁₀CH₃, -(CH₂)₁₂CH₃, -(CH₂)₁₄CH₃, -(CH₂)₁₆CH₃, -cyclohexanyl, -adamantan-2-yl, -phenyl, -biphen-4-yl, -[2,2'-bithiophen-5-yl], -[4-methoxy-1H-indole-2-yl], or -O-C(CH₃)₃. In other preferred embodiments, **tail** is -CH₃, -CH₂CH₃, -(CH₂)₅CH₃, -(CH₂)₆CH₃, -(CH₂)₈CH₃, -(CH₂)₁₀CH₃,

$-(CH_2)_{12}CH_3$, $-(CH_2)_{14}CH_3$, $-(CH_2)_{19}CH_3$, -cyclohexanyl, -adamantan-2-yl, -phenyl, or -biphen-4-yl. In other preferred embodiments, **tail** is $-CH_3$, $-CH_2CH_3$, $-(CH_2)_5CH_3$, $-(CH_2)_6CH_3$, $-(CH_2)_8CH_3$, $-(CH_2)_{10}CH_3$, $-(CH_2)_{12}CH_3$, $-(CH_2)_{14}CH_3$, or $-(CH_2)_{19}CH_3$. In other preferred embodiments, **tail** is $-(CH_2)_8CH_3$, $-(CH_2)_{10}CH_3$, $-(CH_2)_{12}CH_3$, $-(CH_2)_{14}CH_3$, or $-(CH_2)_{19}CH_3$. In other preferred

 5 embodiments, **tail** is -cyclohexanyl, -adamantan-2-yl, -phenyl, -biphen-4-yl, -[2,2'-bithiophen-5-yl], -[4-methoxy-1H-indole-2-yl], or $-O-C(CH_3)_3$. In other preferred embodiments, **tail** is -phenyl, -biphen-4-yl, -[2,2'-bithiophen-5-yl], or -[4-methoxy-1H-indole-2-yl]. In other preferred embodiments, **tail** -phenyl, -biphen-4-yl, or -[2,2'-bithiophen-5-yl]. In other preferred embodiments, **tail** is -biphen-4-yl or -[2,2'-bithiophen-5-yl]. Highly preferred embodiments of **tail** are linear C1-22

 10 alkyl as defined above or -[4-methoxy-1H-indole-2-yl], wherein the alkyl is preferably not unsaturated, wherein the alkyl is preferably not substituted. Representative embodiments of **tail** are shown below.

			
-CH ₃	-CH ₂ CH ₃	-(CH ₂) ₅ CH ₃	-(CH ₂) ₆ CH ₃
			
-(CH ₂) ₈ CH ₃	-(CH ₂) ₁₀ CH ₃	-(CH ₂) ₁₂ CH ₃	-(CH ₂) ₁₄ CH ₃
			
-(CH ₂) ₁₉ CH ₃	-cyclohexanyl	-adamantan-2-yl	-phenyl
			
-biphen-4-yl	-[2,2'-bithiophen-5-yl]	-[4-methoxy-1H-indole-2-yl]	$-O-C(CH_3)_3$
			
-(CH ₂) ₃ CH ₃			

p' is a protecting group. Preferably **p'** is a protecting group that is suitable for SPPS strategies or for solution chemistry. The protecting group is preferably acid-labile. As used herein,

 15 a protecting group has its customary meaning of a group that is linked to a heteroatom in such a

way as to reduce its reactivity, or to protect it from certain conditions. After having served their function, protecting groups can generally be removed again using routine chemistry. A skilled person is well aware of protecting groups and can select useful protecting groups for **p'** for particular groups or atoms to be protected, reaction conditions, storage conditions, purification techniques, or available deprotection techniques. A helpful reference in this regard is Greene's Protective Groups in Organic Synthesis, Wuts & Greene, DOI:10.1002/0470053488.

In preferred embodiments **p'**, when linked to a nitrogen atom, is 9-fluorenylmethyl carbamate (Fmoc), *t*-Butyl carbamate (Boc), benzyl carbamate (Cbz), acetamide (Ac), trifluoroacetamide, phthalimide, benzylamine (Bn), triphenylmethylamine (Tr, or Trt), or tosylamide (Ts). Preferred base-labile embodiments for **p'** are Fmoc, Cbz, Ac, trifluoroacetamide, and phthalimide, more preferably Fmoc or Cbz, most preferably Fmoc. In **tail** a highly preferred **p'** is acid labile, most preferably Boc.

Head

The **head** moiety of the compounds is a warhead and is found at the C-terminus of the central peptide analogues as defined above. Warheads are known moieties for protease inhibitors (see for instance Huang et al., Molecules, 2022, 27(22): 7728 doi: 10.3390/molecules27227728). It is generally a moiety that can help improve the potency of the compounds, or it can optionally be a secondary, shorter tail. **head** can comprise a carbonyl moiety that is connected to the carbon atom that also bears **sc1**, resembling an amino acid moiety. Depending on the nature of **head** it can thus form moieties ranging from an aldehyde (**head** is -C(=O)-H) or a free carboxylic acid (**head** is -C(=O)OH) to an octyl ester (**head** is -C(=O)O-(CH₂)₇-CH₃) or a ketoamide (**head** is -C(=O)C(=O)-NH₂). The inventors found general formula (H) as surprisingly effective new **head** moieties. Although compounds of the invention preferably have a **head** that is general formula (H), other instances of **head** can also be used, particularly when the compound is of general formula (I-p).

The below definitions are particularly relevant when **head** is not of general formula (H). In some embodiments **head** is -C(=O)H, -C(=O)C₁₋₂₄alkyl, -C(=O)(NH)₀₋₁₋₅₋₂₀-membered (hetero)aryl, or -C(=O)(NH)₀₋₁₋₃₋₂₀-membered (hetero)cycloalkyl, wherein **head** is optionally unsaturated, wherein **head** is optionally substituted with halogen, C₁₋₃(halo)alkyl, or C₁₋₃(halo)alkoxyl; in some embodiments **head** is not optionally unsaturated; in some embodiments **head** is not optionally substituted; in some embodiments **head** is not optionally unsaturated and not optionally substituted; preferred optional substitutions of **head** are halogen, C₁₋₂(halo)alkyl, or C₁₋₂(halo)alkoxyl, more preferably halogen, -CF₃, -CH₃, or -OCH₃, wherein halogen is preferably F.

When **head** is -C(=O)H, the C-terminus of the central peptide forms an aldehyde. When **head** is -C(=O)OH it forms a carboxylic acid, and when it is -C(=O)O-C₁₋₂₀alkyl it forms an ester. When **head** is -C(=O)NH₂ it forms a terminal amide, and when **head** is -C(=O)NH-C₁₋₂₀ alkyl it forms an amide moiety that is further substituted with alkyl. When **head** is -C(=O)C(=O)-C₁₋

20alkyl it forms an 1,2-diketo moiety that is further substituted with alkyl, and when head is -C(=O)C(=O)-N(H)C1-20alkyl is forms a ketoamide that is further substituted with alkyl.

In preferred embodiments **head** is - C1-14alkyl, -5-6-membered (hetero)aryl, -3-8-membered cycloalkyl, -(NH)-5-6-membered (hetero)aryl, -NH-3-8-membered cycloalkyl, or -
 5 C(=O)-C1-14alkyl, -C(=O)-5-6-membered (hetero)aryl, -C(=O)-3-8-membered cycloalkyl, -C(=O)-(NH)-5-6-membered (hetero)aryl, or -C(=O)-NH-3-8-membered cycloalkyl. More preferably **head** is -OH, -OCH₃, -CH₃, or -5-6-membered aryl, or - C(=O)-OH, - C(=O)-OCH₃, - C(=O)-CH₃, or - C(=O)-5-6-membered aryl.

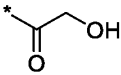
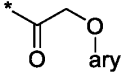
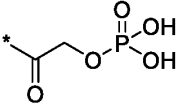
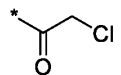
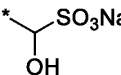
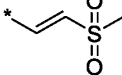
In preferred embodiments **head** is -C(=O)-H or -C(=O)-OH; - C(=O)-C(O)-N(H)C1-20alkyl
 10 that is optionally cyclic and optionally unsaturated such as -C(O)NH-Bn, -C(=O)-C(=O)-NH₂; - C(=O)-5-10-membered (hetero)aryl such as -C(=O)-phenyl, -C(=O)-C1-4alkyl such as -C(=O)-methyl, -C(=O)-NH₂, -C(=O)-NH(C1-6alkyl) such as -C(=O)-NH(propyl) or -C(=O)-NH(hexyl), - C(=O)-NH-S(O)₂-C1-6(cyclo)alkyl such as -C(=O)-NH-S(O)₂-cyclopropyl, -C(=O)-NH-CH₂-(5-6-membered aryl) such as -C(=O)-NH-CH₂-methoxyphenyl or -C(=O)-NH-CH₂-
 15 trifluoromethylphenyl or -C(=O)-NH-benzyl; or such as -C(=O)-N(C1-6alkyl)₂ such as -C(=O)-N(CH₃)₂, or such as -C(=O)-N(C1-6alkyl)(C1-6alkoxy) such as -C(=O)-N(CH₃)(OCH₃); or -B(OH)₂, or optionally substituted 5-membered heteroaryl such as 5-methyl-4-aza-oxazol-2-yl.

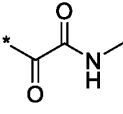
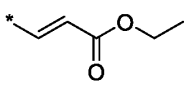
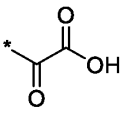
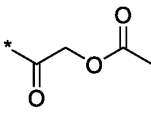
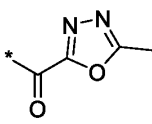
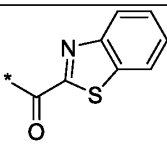
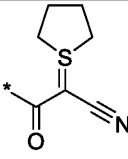
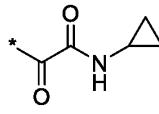
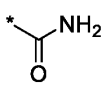
In other preferred embodiments **head** comprises at most 1-10 total of carbon atoms and heteroatoms, preferably 1-8, more preferably 1-7, or is-B(OH)₂ or is optionally methylated 5-
 20 membered heteroaryl such as 5-methyl-4-aza-oxazol-2-yl. More preferably **head** comprises at most 1-7 total of carbon atoms and heteroatoms, or is-B(OH)₂ or is optionally methylated 5-membered heteroaryl such as 5-methyl-4-aza-oxazol-2-yl. Highly preferably **head** is -C(=O)-NH₂, -C(=O)-phenyl, or -C(=O)-N-SO₂-cyclopropyl; or -(5-methyl-1,3,4-oxadiazol-2-yl) or -B(OH)₂. Most preferably **head** is -C(=O)-NH₂, -C(=O)-phenyl, or -B(OH)₂.

25 Preferred embodiments of **head** are -C(=O)H, -C(=O)CH₃, -C(=O)phenyl, -C(=O)OH, -C(=O)OCH₃, -C(=O)C(O)-NH₂, -C(=O)C(O)-NH-CH₂-phenyl, -C(=O)NH₂, -C(=O)N(CH₃)₂, -C(=O)N-(OCH₃)(CH₃), -C(=O)N-SO₂-cyclopropyl, -C(=O)NH-CH₂-(methoxyphen-3-yl), -C(=O)NH-CH₂-(trifluoromethylphen-4-yl), -C(=O)NH-CH₂-phenyl, -C(=O)N-(CH₂)₁₅CH₃, -C(=O)N-(CH₂)₇CH₃, -C(=O)-N-(CH₂)₅CH₃, -C(=O)-N-(CH₂)₂CH₃, -(5-methyl-1,3,4-oxadiazol-2-yl), and -B(OH)₂.

30

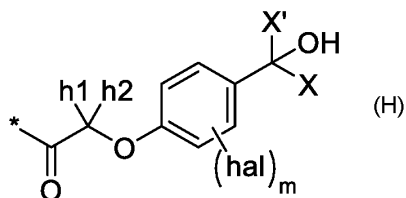
In preferred embodiments **head** is a warhead selected from:

			
(wh1)	(wh2)	(wh3)	(wh4)
			
(wh5)	(wh6)	(wh7)	(wh8)

			
(wh9)	(wh10)	(wh11)	(wh12)
			
(wh13)	(wh14)	(wh15)	(wh16)
			
(wh17)	(wh18)	(wh19)	(wh20)

wherein **aryl** is a 5-6-membered aromatic (hetero)cycle that is optionally substituted with 1, 2, 3, 4, or 5 halogen atoms and that is optionally substituted with 1, 2, or 3 C1-C3 (hydroxy)alkyl moieties, and wherein (wh2) is preferably of general formula (H). In preferred embodiments head is selected from (wh1) through (wh-18), more preferably from (wh1) and (wh2) and (wh16), still more preferably from (wh1) and (wh2), most preferably it is (wh2).

In preferred embodiments **head** is a warhead of general formula (H)



wherein

h1 and **h2** are each independently H, halogen, or C1-4 alkyl, or **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure;

hal is a halogen;

m is 0, 1, 2, 3, or 4;

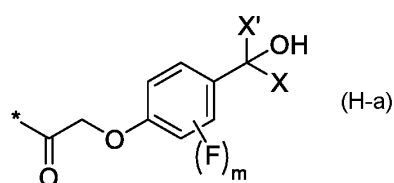
X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure.

hal is a halogen. Examples of halogens are F, Cl, and Br. Preferred halogens are F and Cl, and F is most preferred. **m** is 0, 1, 2, 3, or 4, preferably **m** is 1, 2, 3, or 4, more preferably it is 2, 3, or 4, still more preferably it is 3 or 4, most preferably it is 4. In preferred embodiments **hal** is F and **m** is 3 or 4, more preferably 4.

h1 and **h2** are each independently H, halogen, or C1-4 alkyl, or **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure. Preferably when one of **h1** and **h2** is C1-4 alkyl, the other is H or halogen. Preferred halogen for **h1** and **h2** is F or Cl, more preferably F. For **h1** and **h2** a preferred C1-4 alkyl is a C1-2 alkyl, more preferably methyl. When **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure, the structure preferably forms a spiro-cycloalkyl. In this case, **h1** and **h2** can be said to represent a bridging moiety, preferably selected from $-(CH_2)_2-$, $-(CH_2)_3-$, $-(CH_2)_4-$, and $-(CH_2)_5-$, more preferably selected from $-(CH_2)_2-$ and $-(CH_2)_3-$, most preferably it is $-(CH_2)_2-$. This can be said to form a spiro cyclopropyl moiety. For **h1** and **h2** it is most preferred that **h2** is H. Most preferably both **h1** and **h2** are H.

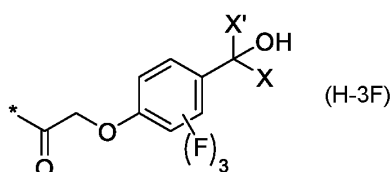
X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure. Preferred halogen for **X** and **X'** is F or Cl, more preferably F. For **X** and **X'** a preferred C1-4 alkyl is a C1-2 alkyl, more preferably methyl. When **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure, the structure preferably forms a spiro-cycloalkyl. In this case, **X** and **X'** can be said to represent a bridging moiety, preferably selected from $-(CH_2)_2-$, $-(CH_2)_3-$, $-(CH_2)_4-$, and $-(CH_2)_5-$, more preferably selected from $-(CH_2)_2-$ and $-(CH_2)_3-$, most preferably it is $-(CH_2)_2-$, wherein the bridging moiety can be optionally substituted with 1, 2, or 3 moieties selected from methyl or F or -OH, preferably methyl or F, preferably 1 or 2 such moieties, more preferably 1, most preferably no such substitutions. Preferably **X** and **X'** are each independently H or C1-2 alkyl, preferably H or -CH₃. **X** is preferably H or -CH₃. **X'** is preferably H or -CH₃. More preferably both **X** and **X'** are H or -CH₃. In some preferred embodiments **X** is -CH₃ and **X'** is H or -CH₃. In some preferred embodiments **X** is -CH₃ and **X'** is -CH₃.

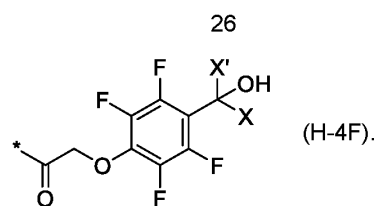
In preferred embodiments **head** is a warhead of general formula (H-a)



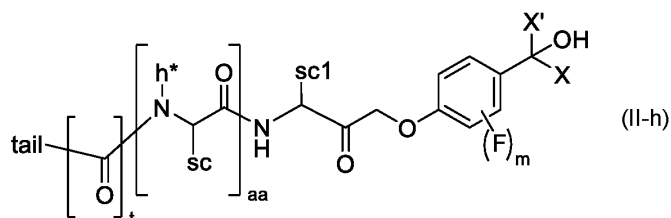
25

In preferred embodiments **head** is a warhead of general formula (H-3F) or (H-4F), more preferably of general formula (H-4F):





Preferred compounds of general formula (I) are of general formula (II-h)



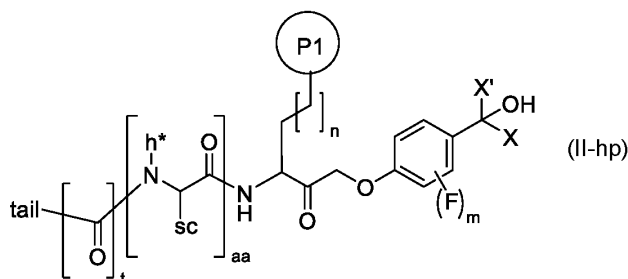
wherein

t is preferably 1;

m is 1, 2, 3, or 4; preferably 2, 3, or 4, more preferably 3 or 4, most preferably 4;

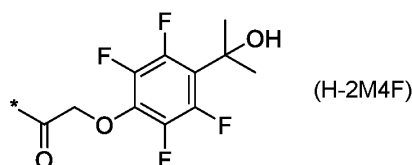
X and X' are each independently H, halogen, or C1-4 alkyl, or X and X' together with the carbon atom to which they are attached form a 3-4 membered cyclic structure, preferably X and X' are each independently H or $-CH_3$, more preferably X is $-CH_3$ and X' is H or $-CH_3$.

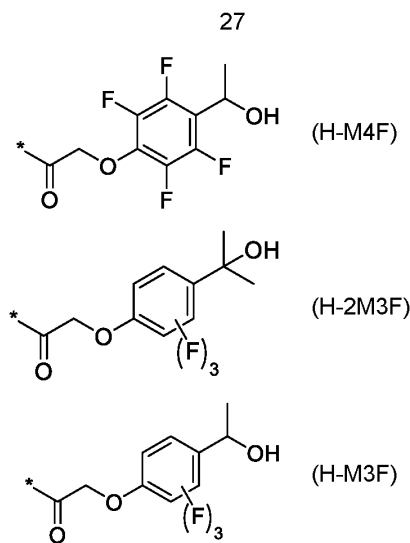
More preferred compounds of general formula (I) are of general formula (II-hp)



wherein t is preferably 1; m is 1, 2, 3, or 4; preferably 2, 3, or 4, more preferably 3 or 4, most preferably 4; X and X' are each independently H, halogen, or C1-4 alkyl, or X and X' together with the carbon atom to which they are attached form a 3-4 membered cyclic structure, preferably X and X' are each independently H or $-CH_3$, more preferably X is $-CH_3$ and X' is H or $-CH_3$.

The following are preferred embodiments of general formula (H):





Of these, (H-2M4F) and (H-M4F) are most preferred. In preferred compounds of the invention **head** is (H-2M4F) or (H-M4F). When both X and X' are identical, compounds have one fewer chiral centre, which can be attractive.

5 Further definitions

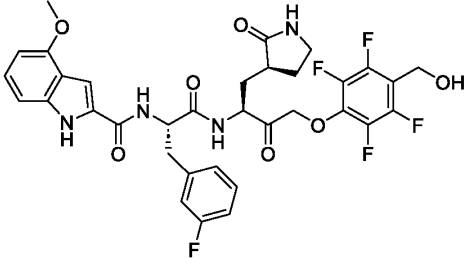
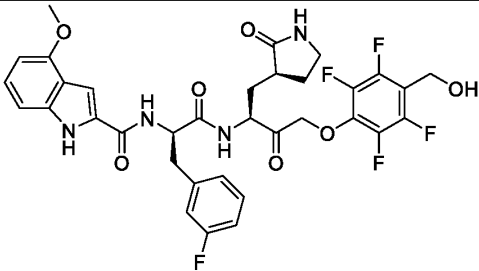
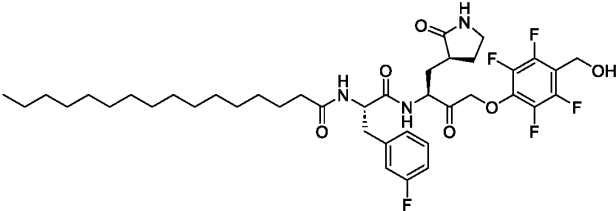
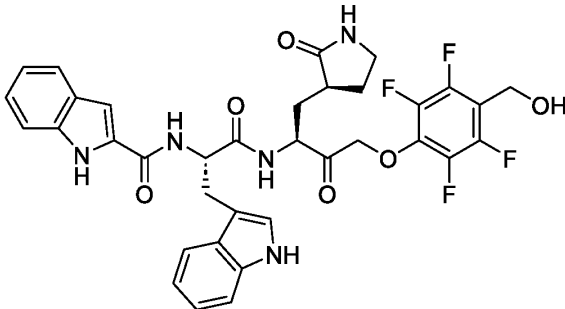
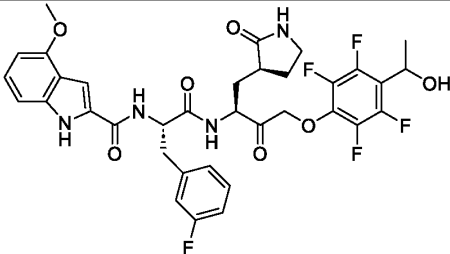
 The following numbered features are preferred:

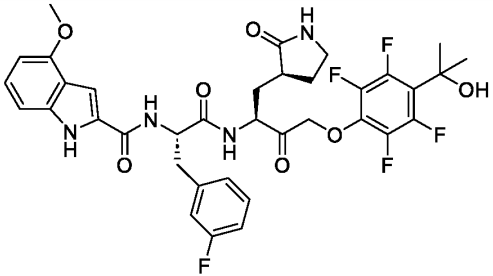
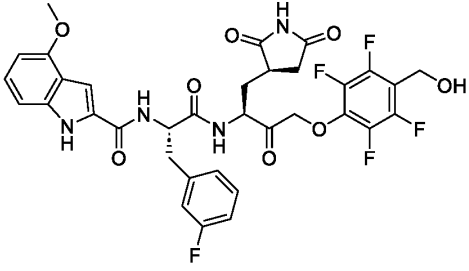
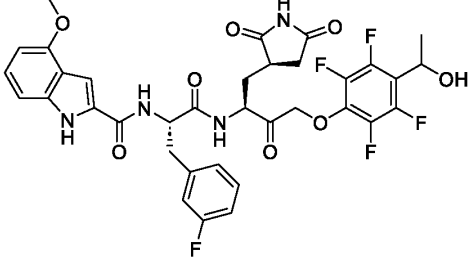
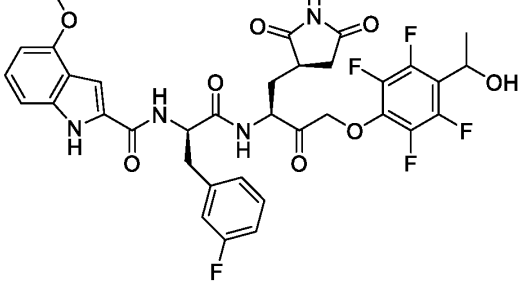
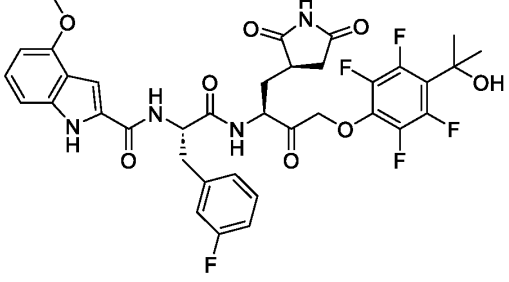
- 1) **aa** is 1 or 2, preferably 1; 2) **t** is 0 or 1, preferably 1; 3) **h*** is H or together with its adjacent **sc** forms a cyclic amino acid side chain; 4) **sc** is butyl such as 2-methyl-n-propyl, is optionally substituted -CH₂-aryl such as -CH₂-fluorophenyl, or is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, wherein optional substitutions are preferably selected from halogen of alkoxy; 5) **tail** is linear C12-18 alkyl such as -C₁₅H₃₁, is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, more preferably methoxyl-1*H*-indol-3-yl, or is optionally substituted C4-11 alkoxy such as -O-CH₂-phenyl, wherein optional substitutions are preferably selected from halogen or alkoxy, preferably alkoxy, more preferably methoxyl.
- 10

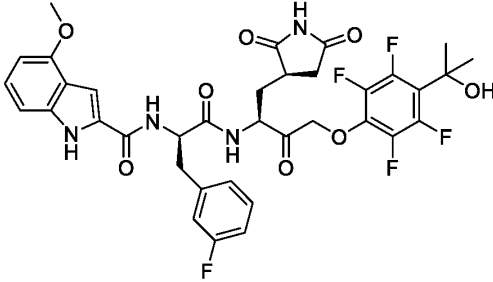
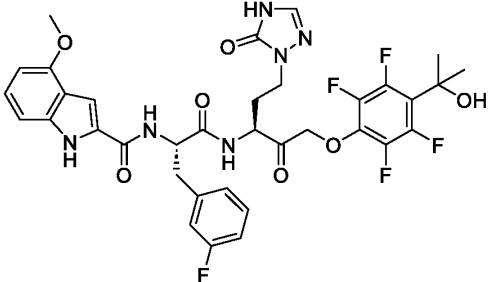
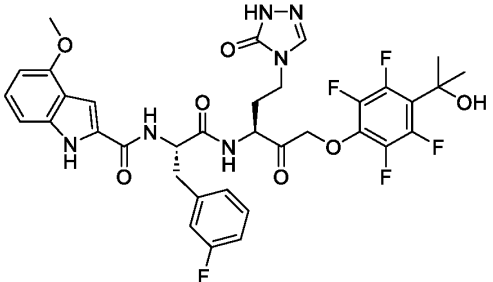
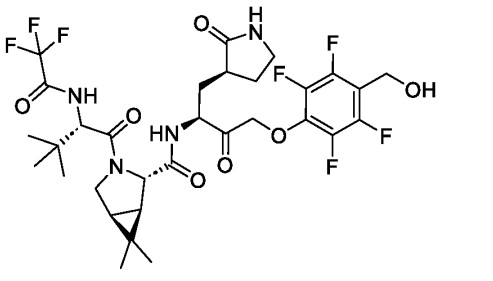
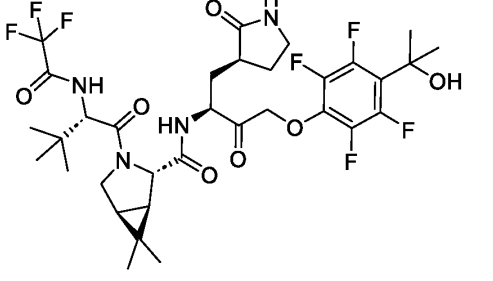
15 The below table show preferred embodiments where the numbered features as described above apply:

Embodiment	Features	Embodiment	Features	Embodiment	Features
A	12	J	45	S	245
B	13	K	123	T	345
C	14	L	124	U	1234
D	15	M	125	V	1235
E	23	N	134	W	1245
F	24	O	135	X	1345
G	25	P	145	Y	2345
H	34	Q	234	Z	12345
I	35	R	235		

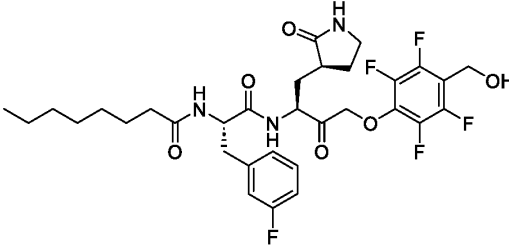
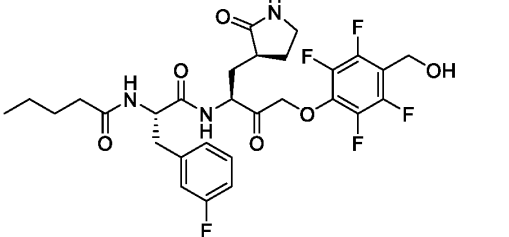
The following are preferred embodiments of the invention:

#	Structure
1	
2	
3	
4	
5	

6	
7	
8	
9	
10	

11	
12	
13	
14	
15	

31

16	
20	

Of the above, more preferred are compounds 1, 5, 6, 12, 13, 14, 15, 16, and 20, preferably 1, 5, 6, 12, 13, 14, 15, and 16, even more preferred are compounds 1, 5, 6, 14, 15, and 16, still more preferred are compounds 1, 5, 6, 14, and 15. Highly preferred compounds are compounds 1, 5, and 6. Other highly preferred compounds are compounds 14 and 15.

5

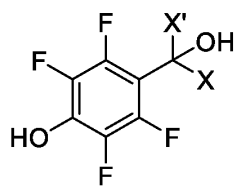
In preferred embodiments of the compound, **tail** comprises at least 8 carbon atoms, preferably it is pentadecyl; and/or **head** comprises at most 1-10 total of carbon atoms and heteroatoms, preferably 1-8, more preferably 1-7, or **head** is -B(OH)_2 or 5-membered heteroaryl such as 5-methyl-4-aza-oxazol-2-yl. When tail is pentadecyl, it can contribute to the formation of a palmitic acid amide, which was found to provide good results. Other attractive amides are C14 and C18, wherein tail is respectively tridecyl or heptadecyl.

In some embodiments, **tail** comprises at least 8 carbon atoms, preferably it is pentadecyl. In some embodiments, **tail** comprises at least 10 carbon atoms. In some embodiments, **tail** comprises at least 11, 12, or preferably 13 carbon atoms. In some embodiments **head** comprises at most 1-10 total of carbon atoms and heteroatoms, preferably 1-8, more preferably 1-7, or **head** is -B(OH)_2 or methylated 5-membered heteroaryl such as 5-methyl-4-aza-oxazol-2-yl. In some embodiments **head** comprises at most 1-10 total of carbon atoms and heteroatoms, preferably 1-8, more preferably 1-7. In some embodiments **head** is -B(OH)_2 or optionally methylated 5-membered heteroaryl such as 5-methyl-4-aza-oxazol-2-yl.

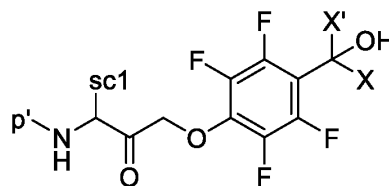
20

Certain compounds are provided that are useful in the synthesis of the inhibitors wherein **head** is of general formula (H). Provided are compounds according to general formula (intermediate-1) or (intermediate-2):

32

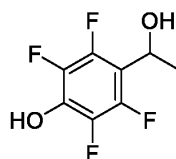


(intermediate-1)

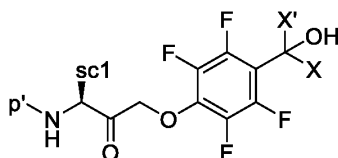


(intermediate-2)

wherein

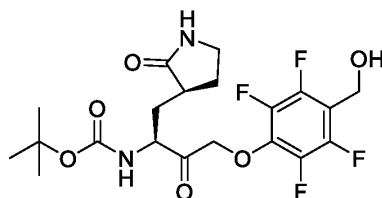
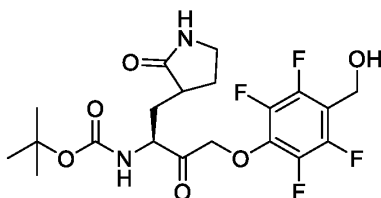
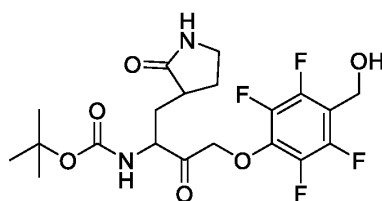
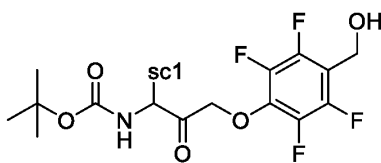
p' is H or a protecting group;**sc1** is an amino acid side chain;**X** is H, halogen, or C1-4 alkyl; preferably **X** is H, halogen, or $-(CH_2)_{0-3}-CH_3$;5 **X'** is H, halogen, or C1-4 alkyl; preferably **X'** is H, halogen, or $-(CH_2)_{0-3}-CH_3$;or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure;wherein for the compound of general formula (intermediate-1) **X'** is halogen or C1-4 alkyl when **X** is H.10 Definitions for **p'** and **sc1** and **X** and **X'** are preferably as defined for compounds of general formula (I). Preferred compounds of general formula (intermediate-1) are:

Compounds of general formula (intermediate-2) are preferably of general formula (intermediate-2L).



(intermediate-2L)

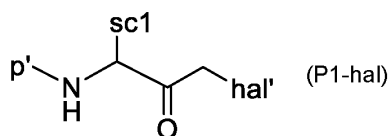
Preferred compounds of general formula (intermediate-2) are:



15

The intermediated can be usefully applied in a method for providing a compound wherein **head** is of general formula (H), the method comprising the step of

- i) contacting a compound according to general formula (intermediate-1) with a compound of general formula (P1-hal):



5

wherein

p' is a protecting group;

sc1 is an amino acid side chain;

hal' is a halogen, preferably chloride;

to obtain a compound of general formula (intermediate-2); and/or

10

- ii) removing the protecting group from a compound according to general formula (intermediate-2), preferably by using a carboxylic acid such as trifluoroacetic acid.

Definitions for **p'** and **sc1** are preferably as defined for compounds of general formula (I).

Finding suitable reaction conditions are within the ambit of the skilled person, and examples of the above synthetic strategy are provided in the Examples.

15

Inhibitors of the invention demonstrated attractive lipophilicity. A suitable means for describing lipophilicity is the LogD value, representing the partitioning of a compound over octanol and an aqueous buffer, preferably at about pH 7.4, preferably as described in the examples.

20

Preferred inhibitors have a LogD value of at most 5, more preferably at most 4, still more preferably at most 3.5, even more preferably at most 3, still more preferably at most 2.6, most preferably at most about 1.8, or at most about 1.5.

Inhibitors of the invention demonstrated attractive solubility. Solubility is preferably kinetic solubility. Solubility is preferably in PBS buffer, more preferably at pH 7.4, preferably as described in the examples. Preferred inhibitors have a solubility of over 3 μ M, more preferably over 10 μ M, still more preferably over 15 μ M, even more preferably over 25 μ M, still more preferably over 50 μ M, most preferably over 70 μ M.

25

Inhibitors of the invention demonstrated attractive plasma stability. A suitable means for describing plasma stability is the percentage of original inhibitor remaining after 2 hours of storage in plasma, preferably as described in the examples. Preferred inhibitors have at least 80% plasma stability, more preferably at least 85% plasma stability, still more preferably at least 88% plasma stability.

30

Inhibitors of the invention demonstrated attractive plasma stability. A suitable means for describing plasma stability is the percentage of original inhibitor remaining after 2 hours of storage in plasma, preferably as described in the examples. Preferred inhibitors have at least 80% plasma

35

stability, more preferably at least 85% plasma stability, still more preferably at least 88% plasma stability.

Inhibitors of the invention demonstrated attractive plasma protein binding values. Suitable means for determining plasma protein binding values are as described in the examples Preferred compounds have a bound fraction of at most 99.5%. Preferred inhibitors have at least 95% plasma stability, more preferably at least 98% plasma stability, still more preferably at least 98.5% plasma stability, most preferably at least 99%.

Inhibitors of the invention demonstrated attractive EC50 values for inhibiting M^{pro}, preferably SARS-CoV-2 M^{pro}. The main proteases (M^{pro}), also termed 3-chymotrypsin-like proteases (3CL^{pro}), are a class of highly conserved cysteine hydrolases in β -coronaviruses. Suitable means for determining M^{pro} inhibition values are as described in the examples Preferred compounds have an EC50 value for inhibiting M^{pro} of below 20 μ M, more preferably below 3 μ M, still more preferably below 1 μ M, even more preferably below 35 nM, still more preferably below 10 nM. Surprisingly, superior values were found for compounds wherein the amino acid residue adjacent to the amino acid bearing **sc1** is an L-amino acid. Preferred inhibitors show good inhibition of M^{pro} while showing low inhibition of cathepsin.

Inhibitors of the invention demonstrated attractive microsomal stability. Suitable means for determining microsomal stability values are as described in the examples Preferred compounds have a microsomal half-life of at least 30 seconds, preferably at least 60 seconds. Preferred compounds have a microsomal half-life of at most 2 minutes.

Inhibitors of the invention demonstrated attractive hepatocyte stability. Suitable means for determining hepatocyte stability values are as described in the examples Preferred compounds have a hepatocyte half-life of at least 11.5 minutes, preferably at least 15 minutes.

Viral proteases

Compounds of the invention are suitable as protease inhibitors, advantageously as inhibitors of viral proteases, particularly of flaviviral proteases. Flavivirus is a genus of the family *Flaviviridae*. This genus includes the West Nile virus, dengue virus, Tick-borne Encephalitis Virus, Japanese Encephalitis virus, Yellow Fever Virus, and several other viruses that may cause encephalitis. In addition, Compounds of the invention are suitable as protease inhibitors, advantageously as inhibitors of proteases associated with coronaviral infection. This encompasses inhibition of host proteases that are associated with the viral infection. For instance, several coronaviral infections are known to progress through activity of host proteases such as cathepsin or TMPRSS2. Such host proteases can assist a virus in entering a cell, or can otherwise assist a virus in its replication cycle.

Dengue virus and its various strains and isolates are members of the genus *Flavivirus*. The genus *Flavivirus* is a genera of the *Flaviviridae* family and includes the viral groups of Yellow Fever virus group, Tick-borne encephalitis virus group, Rio Bravo Group, Japanese encephalitis Group, Tyuleniy Group, Ntaya Group, Uganda S Group, Dengue Group, and Modoc Group. Members of the Flavivirus genus may produce a wide variety of disease states, such as fever,

arthralgia, rash, hemorrhagic fever, and/or encephalitis. The outcome of infection is influenced by both the virus and host-specific factors, such as age, sex, genetic susceptibility, and/or pre-exposure to the same or a related agent. Some of the various diseases associated with members of the genus *Flavivirus* are yellow fever; dengue fever; and West Nile, Japanese, and St. Louis encephalitis.

The existence of a trypsin-like serine protease domain in the N-terminal region of the flaviviral NS3 proteins was originally predicted by sequence comparisons between cellular and virus-encoded proteases. The NS2B-NS3 endopeptidases of the *Flavivirus* genus, which at present comprises at least 68 known members, are now commonly designated as flavivirin (EC 3.4.21.91). The dengue virus 69 kDa NS3 protein is a multifunctional protein with a serine protease domain located within the N-terminal 167 amino acid residues and activities of a nucleoside triphosphatase (NTPase) and RNA helicase in the C-terminal moiety. A catalytic triad consisting of residues His51, Asp75 and Ser135 was identified by site-directed mutagenesis experiments and replacement of the catalytic serine by alanine resulted in an enzymatically inactive NS3 protein. The NS3 protease is an essential component for maturation of the virus and viable virus was never recovered from infectious cDNA clones carrying mutations in the NS3 sequence that abolished protease activity. Interaction of the helicase portion of NS3 with the viral RNA-dependent RNA polymerase NS5 may promote the association of the viral replicase complex to the membranes of the ER.

The DENV NS3 is a serine protease, as well as an RNA helicase and RTPase/NTPase. The protease domain consists of six β -strands arranged into two β -barrels formed by residues 1-180 of the protein. The catalytic triad (His-51, Asp-75 and Ser-135), is found between these two β -barrels, and its activity is dependent on the presence of the NS2B cofactor. This cofactor wraps around the NS3 protease domain and becomes part of the active site. The remaining NS3 residues (180-618), form the three subdomains of the DENV helicase. A six-stranded parallel β -sheet surrounded by four α -helices make up subdomains I and II, and subdomain III is composed of 4 α -helices surrounded by three shorter α -helices and two antiparallel β -strands

The presence of a small activating protein or co-factor is a prerequisite for optimal activity of the flaviviral NS3 proteases with their natural polyprotein substrates. Although the dengue virus NS3 protease exhibits NS2B-independent activity with model substrates for serine proteases, enzymatic cleavage of dibasic peptides is markedly enhanced with the NS2B-NS3 co-complex and the presence of the NS2B activation sequence is important for the cleavage of polyprotein substrates in vitro. The initial characterization of the co-factor requirement for the dengue virus NS3 protease had revealed that the minimal region necessary for protease activation was located in a 40-residue hydrophilic segment of NS2B.

A typical enzyme of interest in the present invention is a serine protease from viruses of the flaviviridae family (especially the flavivirus genus), such as dengue protease or West Nile protease. For example, dengue proteases from the four different dengue virus serotypes have all been well known and characterized in the art. See, e.g., Chambers et al., Proc. Nat. Acad. Sci. USA 87: 8898-902, 1990; Chambers et al., J. Virol. 67: 6797-807, 1993, Ryan et al. J. Gen.

Virology 79: 947-959, 1998; Murthy et al., J Biol Chem 274: 5573-5580, 1999; Murthy et al., J Mol Biol 301: 759-767, 2000; Brinkworth et al., J Gen Virol 80, 1167-1177, 1999 (Den2), and Leung et al., J Biol Chem 276: 45762-71, 2001 (Den2). The virus-encoded dengue protease comprises the amino-terminal 180 amino acids of NS3 (NS3pro) of the polyprotein. It is responsible for cleavage
5 both in cis and in trans to generate viral proteins that are essential for viral replication and maturation of infectious dengue virions. In addition to its protease activity, the carboxyl-terminal region of NS3 encodes both nucleoside triphosphatase and helicase activities (see, e.g., Li et al., J. Virol. 73: 3108-3116, 1999). Similarly, other flaviviruses (e.g., West Nile virus and yellow fever virus) and their proteases have also been well characterized in the art. See, e.g., Anderson et al.,
10 Ann N Y Acad. Sci. 951:328-31, 2001; Nall et al., J Biol Chem. 79:48535-42, 2004; J Biol Chem. 280:2896-903, 2005; Hillyer et al., Histochem Cell Biol. 117:431-40, 2002; Chambers et al., J Gen Virol. 86:1403-13, 2005; Scaramozzino et al., Biochem Biophys Res Commun. 294:16-22, 2002; and Bessaud et al., Virus Res. 120:79-90, 2006. The viral genome of the Flavivirus genus is translated as a single polyprotein and is subsequently cleaved into mature proteins.

15 Compounds of the invention are preferably inhibitors of viral protease, preferably of flaviviral protease, more preferably of NS3 proteases such as flaviviral NS3 proteases, most preferably of flavivirin NS2B-NS3 endopeptidases such as dengue NS2B-NS3 protease. Throughout this document preferred proteases to be inhibited are viral protease, preferably flaviviral protease, more preferably NS3 proteases such as flaviviral NS3 proteases, most
20 preferably of flavivirin NS2B-NS3 endopeptidase such as dengue NS2B-NS3 protease.

Compounds of the invention can additionally be useful for being specific. Preferably the compounds are specific protease inhibitors. Preferably the compounds do not inhibit or do not strongly inhibit other proteases such as host proteases or proteases of the subject to be treated. More preferably the compounds do not inhibit or do not strongly inhibit other serine proteases
25 such as trypsin. As used herein, compounds that do not strongly inhibit a protease preferably inhibit it with a K_i of 200, 300, 400, 500 μ M or greater, or do not detectably inhibit it.

In embodiments, the compounds according to the invention inhibit their target proteases at least 10% more than that they inhibit trypsin, preferably human trypsin, under the same experimental conditions. More preferably this is 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%,
30 95%, 96%, 97%, 98%, 99%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450%, or 500% more, or even more. In other embodiments the compounds according to the invention inhibit trypsin, preferably human trypsin, at most 90% as compared to their inhibition of a viral protease, under the same experimental conditions. More preferably at most 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, or less. Methods for
35 determining inhibition are known in the art, and are exemplified in the examples.

Therapeutic compounds and compositions of the invention are optionally tested in one or more appropriate in vitro and/or in vivo animal model of disease, to confirm efficacy, tissue metabolism, and to estimate dosages, according to methods well known in the art. In particular, dosages can initially be determined by activity, stability or other suitable measures of the
40 formulation.

Methods and uses of the compounds

Because the compounds of the invention are good inhibitors of proteases associated with viral infection, they can be used in related methods. The invention provides a method for inhibiting a protease associated with viral infection, preferably a coronaviral infection, the method comprising the step of contacting the protease with a compound according to the invention, or with a composition according to the invention. This method is preferably for the treatment of a disease or condition associated with viral, preferably coronaviral, infection or a condition related to a viral, preferably coronaviral, infection, and comprises administration to the subject of an effective dose of a compound or composition according to the invention. Preferably the protease is a flaviviral protease, more preferably it is a dengue virus protease, a West Nile virus protease, or a tick-borne encephalitis virus protease, most preferably it is a protease associated with coronaviral infection, which can be a host protease such as cathepsin or TMPRSS2. In some embodiments it is a West Nile virus protease. In some embodiments it is a tick-borne encephalitis virus protease. Highly preferably it is a dengue virus protease. Most preferably it is a protease associated with coronaviral infection such as described elsewhere herein. In some embodiments it is a dengue virus protease or a tick-borne encephalitis virus protease. In some embodiments it is a dengue virus protease or a West Nile virus protease. In some embodiments it is a West Nile virus protease or a tick-borne encephalitis virus protease.

In one embodiment, the use is provided of either a compound according to the invention, or with a composition according to the invention. Said use is preferably for the treatment of a disease or condition associated with viral, preferably coronaviral, infection or a condition related to a viral, preferably coronaviral, infection, and comprises administration to the subject of an effective dose of a compound or composition according to the invention. Further features and definitions are preferably as defined elsewhere herein, particularly for diseases or conditions to be treated. The methods and uses described here can be in vitro, in vivo, or ex vivo.

Medical use of the compounds

In another aspect, the invention provides the compound or composition according to the invention for use as a medicament. The medicament is preferably for use in the treatment of a viral infection or a condition related to a viral infection, more preferably for use in the treatment of a coronaviral infection or a condition related to a coronaviral infection. The medicament is very suitable for use in a method of treating, preventing, or delaying a viral infection or a condition related to a viral infection in a subject in need thereof, the method comprising the step of administering to the subject an effective amount of a compound or composition according to the invention. Herein, the viral infection is preferably a coronaviral infection. More preferably, the viral infection is a betacoronaviral infection. In preferred embodiments a betacoronaviral infection is a sarbecoviral infection. Sarbecovirus is the subgenus of the genus betacoronavirus that comprises the species SARS-CoV-1 and SARS-CoV-2. In preferred embodiments a betacoronaviral infection is a merbecoviral infection. Merbecovirus is the subgenus of the genus betacoronavirus that

comprises the species MERS-CoV. Highly preferably the viral infection is an infection by SARS-CoV-1, SARS-CoV-2, or MERS-CoV. In some embodiments the viral infection is an infection by SARS-CoV-1 or SARS-CoV-2.

5 Examples of viral infections are flaviviral infections. Flavivirus is a genus of the family *Flaviviridae*. This genus includes the West Nile virus, dengue virus, Tick-borne Encephalitis Virus, Japanese Encephalitis virus, Yellow Fever Virus, and several other viruses that may cause encephalitis. In certain aspects, the flavivirus infection may be a Dengue virus infection or a West Nile virus infection. Additional flaviviruses that can be treated or prevented include other
10 mosquito-borne flaviviruses, such as Japanese encephalitis, Murray Valley encephalitis, St. Louis encephalitis, Kunjin, Rocio encephalitis, and Ilheus viruses; tick-borne flaviviruses, such as Central European encephalitis, Siberian encephalitis, Russian Spring-Summer encephalitis, Kyasanur Forest Disease, Omsk Hemorrhagic fever, Louping ill, Powassan, Negishi, Absettarov, Hansalova, Apoi, and Hypr viruses.

15 The compounds of the present invention can inhibit flaviviral NS3 proteases and therefore can have applications in treating or preventing diseases caused by these viruses. For example, they are useful in the treatment of West Nile encephalitis, yellow fever, Japanese encephalitis, dengue fever, dengue hemorrhagic fever, and dengue shock syndrome associated with the different dengue virus serotypes. In some embodiments the treatment is of serotype DENV-1,
20 serotype DENV-2, serotype DENV-3, or serotype DENV-4. In some embodiments the treatment is of serotype DENV-1. In preferred embodiments the treatment is of serotype DENV-2. In some embodiments the treatment is of serotype DENV-3. In some embodiments the treatment is of serotype DENV-4. In preferred embodiments the treatment is of infection by Dengue virus or of infection by West Nile Virus, more preferably by DENV-2 or by West Nile Virus.

25 Preferred examples of coronaviral infections are respiratory coronaviral infections. A respiratory infection can lead to pulmonary dysfunction. Acute respiratory distress syndrome (ARDS) is a preferred type of pulmonary dysfunction, and a type of respiratory failure associated with fast onset of widespread inflammation in the lungs. Symptoms can include shortness of breath, rapid breathing, and blue skin coloration (central and/or peripheral cyanosis). Survivors
30 often experience a decreased quality of life. Coronaviral infections, particularly respiratory coronaviral infections, may cause ARDS. A preferred treatment as described herein is a treatment or amelioration or prevention of coronaviral infection-associated ARDS. A subject to be treated in this context is preferably a subject suffering from a coronaviral infection, or suspected of suffering from a coronaviral infection, while not suffering ARDS, or while not suffering severe ARDS. For
35 ARDS in general, the underlying mechanism involves diffuse injury to cells which form the barrier of the microscopic air sacs of the lungs (alveoli), surfactant dysfunction, activation of the immune system, and dysfunction of the body's regulation of blood clotting. In effect, ARDS impairs the lungs' ability to exchange oxygen and carbon dioxide. Diagnosis is based on a $\text{PaO}_2/\text{FiO}_2$ ratio (ratio of partial pressure arterial oxygen and fraction of inspired oxygen) of less than 300 mm Hg
40 despite a positive end-expiratory pressure (PEEP) above 5 cm H_2O . The coronaviral infection

associated with and/or causing the ARDS preferably is an infection by a respiratory virus, such as SARS-CoV (or SARS-CoV-1), MERS-CoV, or SARS-CoV-2. Most preferably, the viral infection is an infection by SARS-CoV-2, i.e. is COVID-19.

The compounds of the present invention can inhibit protease associated with SARS-CoV-1 infection, protease associated with SARS-CoV-2 infection, protease associated with MERS-infection, TMPRSS2, and Cathepsins. In preferred embodiments the compounds inhibit SARS-CoV-1 associated protease, SARS-CoV-2 associated protease, or MERS-CoV associated protease. In preferred embodiments the compounds inhibit SARS-CoV-2 associated protease or cathepsins or TMPRSS2. In some embodiments the compounds inhibit SARS-CoV-1 associated protease. In some embodiments the compounds inhibit SARS-CoV-2 associated protease. In some embodiments the compounds inhibit MERS-CoV associated protease. In some embodiments the compounds inhibit TMPRSS2. In some embodiments the compounds inhibit Cathepsins. In some embodiments the compounds inhibit TMPRSS2 or Cathepsins.

Transmembrane protease, serine 2 is an enzyme that in humans is encoded by the TMPRSS2 gene. It belongs to the TMPRSS family of proteins, whose members are transmembrane proteins which have a serine protease activity. Several viruses, including SARS-CoV-2, use the protease activity of the TMPRSS2 protein in the process of entering cells. Cathepsins are proteases commonly found in animals, including humans, as well as other organisms.

Cathepsins can also be used by viruses such as coronaviruses to enter cells, or to escape endosomes. Cathepsins are a well-known family of proteases, and most of the members become activated at the low pH found in lysosomes. Thus, the activity of this family predominantly lies within those organelles. There are exceptions such as cathepsin K, which works extracellularly after secretion by osteoclasts in bone resorption. Cathepsins have an important role in mammalian cellular turnover. Several cathepsins exist, for instance cathepsin A, cathepsin B, cathepsin C, cathepsin D, cathepsin E, cathepsin F, cathepsin G, cathepsin H, cathepsin K, cathepsin L1, cathepsin L2, cathepsin O, cathepsin S, cathepsin W, or cathepsin Z. In some embodiments the protease associated with coronaviral infection is a cysteine protease. In preferred embodiments, the compounds are for inhibiting cathepsin B, cathepsin C, cathepsin F, cathepsin H, cathepsin L1, cathepsin L2, cathepsin O, cathepsin S, cathepsin W, or cathepsin Z, and optionally cathepsin K. These cathepsins are cysteine proteases. Most preferably a cathepsin is cathepsin B or cathepsin L.

In preferred embodiments the compounds inhibit TMPRSS2 with an IC₅₀ of at most 30 µM, more preferably at most 25 µM, still more preferably at most 20 µM, still more preferably at most 15 µM. It is highly preferred that the compounds inhibit TMPRSS2 with an IC₅₀ of at most 12 µM, more preferably at most 5 µM, most preferably at most 2 µM. Most preferred compounds inhibit TMPRSS2 with an IC₅₀ of at most 1 µM, more preferably at most 0.5 µM, still more preferably at most 0.2 µM.

In preferred embodiments the compounds inhibit cathepsins with an IC₅₀ of at most 10 µM, more preferably at most 5 µM, still more preferably at most 4 µM, still more preferably at most

3 μM . It is highly preferred that the compounds inhibit cathepsins with an IC_{50} of at most 1 μM , more preferably at most 0.5 μM , most preferably at most 0.4 μM . Most preferred compounds inhibit cathepsins with an IC_{50} of at most 0.05 μM , more preferably at most 0.01 μM , still more preferably at most 0.005 μM .

5

Flavivirus is a genus of the family *Flaviviridae*. This genus includes the West Nile virus, dengue virus, Tick-borne Encephalitis Virus, Japanese Encephalitis virus, Yellow Fever Virus, and several other viruses that may cause encephalitis. In certain aspects, the flavivirus infection may be a Dengue virus infection or a West Nile virus infection. Additional flaviviruses that can be treated or prevented include other mosquito-borne flaviviruses, such as Japanese encephalitis, Murray Valley encephalitis, St. Louis encephalitis, Kunjin, Rocio encephalitis, and Ilheus viruses; tick-borne flaviviruses, such as Central European encephalitis, Siberian encephalitis, Russian Spring-Summer encephalitis, Kyasanur Forest Disease, Omsk Hemorrhagic fever, Louping ill, Powassan, Negishi, Absettarov, Hansalova, Apoi, and Hypr viruses.

The compounds of the present invention can inhibit flaviviral NS3 proteases and therefore can have applications in treating or preventing diseases caused by these viruses. For example, they are useful in the treatment of West Nile encephalitis, yellow fever, Japanese encephalitis, dengue fever, dengue hemorrhagic fever, and dengue shock syndrome associated with the different dengue virus serotypes. In some embodiments the treatment is of serotype DENV-1, serotype DENV-2, serotype DENV-3, or serotype DENV-4. In some embodiments the treatment is of serotype DENV-1. In preferred embodiments the treatment is of serotype DENV-2. In some embodiments the treatment is of serotype DENV-3. In some embodiments the treatment is of serotype DENV-4. In preferred embodiments the treatment is of infection by Dengue virus or of infection by West Nile Virus, more preferably by DENV-2 or by West Nile Virus.

25

The protease inhibitors or prodrugs of the present invention can be used alone or in combination with any known antiviral drugs to treat these infections. In preferred embodiments, treatment of a coronaviral infection comprises preventing, ameliorating, suppressing, or curing a symptom or syndrome associated with the coronaviral infection.

30

The methods described herein are useful for both prophylactic and therapeutic treatment of viral infections such as Flavivirus infections or preferably coronavirus infections. For prophylactic use, a therapeutically effective amount of one or more compounds as described herein is administered to a subject prior to exposure (e.g., before or when traveling to a location where Flavivirus infections are possible), during a period of potential exposure to Flavivirus infections, or after a period of potential exposure to Flavivirus infections. Prophylactic administration can occur for several days to weeks prior to potential exposure, during a period of potential exposure, and for a period of time, e.g., several days to weeks, after potential exposure. Therapeutic treatment involves administering to a subject a therapeutically effective amount of one or more compounds or compositions according to the invention.

40

The term "treat" or "treatment" or "treating" refers to the medical management of a patient with the intent to cure, ameliorate, stabilize, or prevent a disease, pathological condition, or disorder. In certain aspect treating reduces viral load, ameliorate symptoms, delays progression of disease, etc. This term includes active treatment that includes treatment directed specifically
5 toward the improvement of a disease, pathological condition, or disorder, and also includes causal treatment, that is, treatment directed toward removal of the cause of the associated disease, pathological condition, or disorder. In addition, this term includes supportive treatment including treatment employed to supplement another specific therapy directed toward the improvement of the associated disease, pathological condition, or disorder.

10 As used herein the terms treatment, treat, or treating refer to a method of reducing or delaying one or more symptoms of viral infections such as a Flavivirus infection, more preferably such as a coronaviral infection. Thus in the disclosed method, treatment can refer to a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% reduction in the severity or progression of one or more symptoms of the disease or condition. For example, a method for treating a disease is
15 considered to be a treatment if there is a 10% reduction in one or more symptoms or signs of the viral infection in a subject as compared to a control. As used herein, control refers to the untreated condition. Thus, the reduction can be a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or any percent reduction in between 10% and 100% as compared to native or control levels. It is understood that treatment does not necessarily refer to a cure or complete ablation of the
20 disease, condition, or symptoms of the disease or condition. Examples of symptoms are fever and fatigue.

As used herein, the terms prevent, preventing, and prevention of a disease or disorder refer to an action, for example, administration of a composition or therapeutic agent, that occurs before or at about the same time a subject begins to show one or more symptoms of the disease
25 or disorder, which inhibits or delays onset or severity of one or more symptoms of the disease or disorder. For example, the method is considered to be a prevention if there is a reduction or delay in onset, incidence, severity, or recurrence of a virus infection, preferably a coronavirus infection. The reduction or delay in onset, incidence, severity, or recurrence of a virus infection can be a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or any percent reduction in between
30 10% and 100% as compared to native or control levels.

As used herein the terms treatment, treat, or treating refer to a method of reducing or delaying one or more symptoms of a virus infection. Thus in the disclosed method, treatment can refer to a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% reduction in the severity or
35 progression of one or more symptoms of the disease or condition. For example, a method for treating a disease is considered to be a treatment if there is a 10% reduction in one or more symptoms or signs of the virus infection in a subject as compared to a control. As used herein, control refers to the untreated condition. Thus the reduction can be a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or any percent reduction in between 10% and 100% as compared to native or control levels. It is understood that treatment does not necessarily refer to a cure or
40 complete ablation of the disease, condition, or symptoms of the disease or condition.

Compositions, formulation, and administration

In a further aspect, the invention provides a composition comprising a pharmaceutically acceptable excipient and a compound according to the invention, preferably wherein the composition is a pharmaceutical composition. Such a composition is referred to herein as a composition according to the invention. Preferred compositions according to the invention are pharmaceutical compositions and are for use in treatment as defined elsewhere herein. In preferred embodiments, the composition according to the invention is formulated for oral, sublingual, parenteral, intravascular, intravenous, subcutaneous, inhaled, or transdermal administration; preferably for oral administration. More features and definitions of administration methods are provided below.

Oral drug administration is a commonly used route for many drugs. Compounds of the invention can be seen as peptidomimetics, for which intravenous injections and intramuscular injections each are clinically commonly used administration routes. Surprisingly, subcutaneous delivery (another well-accepted administration method) was found to improve the bioavailability and pharmacokinetic behaviour of compounds of the invention. Subcutaneous administration of compounds has great efficacy, and that the compounds show improved pharmacokinetic behaviour when administered subcutaneously. This can enable prophylactic and therapeutic use, for example in the case of viral infection such as dengue infection, in comparison to other routes of intake such as oral intake.

The increased efficiency allows subcutaneous drug delivery to achieve additional advantages for the three main target populations related to viral infections such as dengue, which are (1) therapeutic use for early treatment of for instance dengue; this first group are the patients that have recently been infected by the virus and need effective treatment as soon as possible and preferentially with a fast onset of action, to immediately suppress the viral load built-up and associated disease symptoms, thereby reducing the risk of progression to severe disease and progression of the virus to others through cross-infection. The two other groups comprise healthy subjects that are at risk of infection and therefore require access to prophylactic use: (2) prophylactic use for subjects at risk of contracting a viral infection such as dengue, and (3) prophylactic use for subjects visiting regions where the viral infection such as dengue is endemic.

To control epidemics such as dengue epidemics (e.g., outbreak management), it is important that all target populations have access to treatment. In preferred embodiments, the compound for use or the composition for use is for therapeutic use for early treatment of a coronaviral infection, or optionally of a viral infection such as dengue, for prophylactic use for subjects at risk of contracting coronaviral infection or optionally dengue, or for prophylactic use for subjects visiting regions where coronaviral infection or optionally dengue is endemic. In some embodiments the compound for use or the composition for use is for therapeutic use for early treatment of coronaviral infection or optionally dengue. In some embodiments the compound for use or the composition for use is for prophylactic use, preferably for subjects at risk of contracting coronaviral infection or optionally dengue, or for subjects visiting regions where coronaviral

infection or optionally dengue is endemic. In some embodiments the compound for use or the composition for use is for prophylactic use for subjects at risk of contracting coronaviral infection or optionally dengue. In some embodiments the compound for use or the composition for use is for prophylactic use for subjects visiting regions where coronaviral infection or optionally dengue is endemic.

Subcutaneous administration was found to have specific advantages such as dengue-relevant advantages for efficient systemic delivery and patient/traveller compliance and convenience. Pharmacokinetic data can show slow elimination of the compounds, indicating that a single dose subcutaneous administration can be sufficient for one or two weeks of treatment in a therapeutic and prophylactic setting. Based on preclinical data, a single dose subcutaneous administration can be sufficient for one or two weeks of therapeutic and/or prophylactic treatment with compounds according to the invention. This has relevant advantages for patient compliance such as dengue patient compliance and convenience since patients with severe viral infection such as dengue may not be able to take oral drugs. This approach also yields advantages for visitors to areas where there is an increased risk of infection, such as travellers who visit risk areas, who need only a (bi)weekly shot (similar to a vaccination) instead of an oral (twice) daily prophylaxis pill.

Subcutaneous administration can be either via direct injection, via autoinjectors, or via patches. Preferably it is via direct injection or via patches, more preferably via direct injection.

Vector-mediated transmission of DENV is initiated when a blood-feeding female Aedes mosquito injects saliva, together with the virus, into the skin of its mammalian host. The cells that first encounter the infection are skin resident macrophages, dendritic cells and keratinocytes (Martina, Koraka, and Osterhaus. 2009. Clinical Microbiology Reviews 22(4)). These infected cells migrate to the lymph nodes where macrophages and monocytes are exposed to the virus and get infected. The DENV infection progresses to viremia (primary viremia; virus in blood) due to the presence of the virus leaving the draining and remote lymph nodes. Infection of DENV is observed in the spleen, kidneys, lungs, and liver, although spleen and liver are the key target organs. Specifically, the macrophages in these organs are the major sites of viral replication (Martina et al., 2009). The secondary viremia occurs when the virus leaves the organs, which can be detected as early as 24–48 h (4–7 days after infection) before the onset of clinical symptoms (such as fever) and can last up to 10–12 days. High concentration of virus in the blood is related to the risk to develop severe disease (haemorrhage and shock). Low to moderate concentration of virus in the blood is associated with mild disease (Martina et al., 2009).

Following subcutaneous administration, molecules reach the systemic circulation via either the blood capillaries or the lymphatic system. Data showed that efficacy against DENV increased substantially when compounds were administered via subcutaneous injection over other routes of administration. The compounds reached more than three times higher concentrations in the main target sites of for instance dengue (liver and spleen) when compared to blood plasma. Accordingly, in preferred embodiments the compound for use or composition for use is for use in a method of targeting a protease inhibitor to a target organ of a subject, the

method comprising injected, preferably subcutaneous administration of the protease inhibitor. Injected administration can be intravenous, intramuscular, or subcutaneous. Preferred targeted organs are liver, kidney, lungs, and spleen, more preferably liver, kidney, and spleen, still more preferably liver and spleen, most preferably liver. In some embodiments the target organs are liver and kidney. In some embodiments the target organ is spleen. Preferably the protease inhibitor is a DENV protease inhibitor.

The invention also provides combinations of compounds according to the invention with further measures known for treating or ameliorating diseases or conditions related to a viral infection, preferably a coronaviral infection. In preferred embodiments of such combinations is provided a combination of a compound according to the invention and an antiviral agent. Antiviral agents are widely known. In another preferred combination, the compound according to the invention is combined with an anti-inflammatory agent. In some preferred combinations the compound may be combined with clinical management, for example involving physical therapy, aerobic exercise, respiratory function therapy, or orthopedic interventions.

Antiviral agents are known in the art and a skilled can select a suitable antiviral agent in light of this disclosure. Examples of suitable antiviral agents are abacavir, acyclovir, adefovir, amantadine, amplitgen, amprenavir, umifenovir, atazanavir, atipla, baloxavir marboxil, biktavir, boceprevir, bulevirtide, cidofovir, cobicistat, combivir, daclatasvir, darunavir, delavirdine, descovy, didanosine, docosanol, dolutegravir, doravirine, edoxudine, efavirenz, elvitegravir, emtricitabine, enfuvirtide, entecavir, etravirine, famciclovir, fomivirsen, fosamprenavir, foscarnet, ganciclovir, ibacitabine, ibalizumab, idoxuridine, imiquimod, imunovir, indinavir, lamivudine, letermovir, lopinavir, loviride, maraviroc, methisazone, moroxydine, nelfinavir, nevirapine, nexavir, nitazoxanide, norvir, oseltamivir, penciclovir, peramivir, penciclovir, peramivir, pleconaril, podophyllotoxin, raltegravir, remdesivir, ribavirin, rilpivirine, rilpivirine, rimantadine, ritonavir, squinavir, simeprevir, sofosbuvir, stavudine, taribavirin, telaprevir, telbivudine, tenofovir alafenamide, tenofovir disoproxil, tipranavir, trifluridine, trizivir, tromantadine, truvada, umifenovir, valaciclovir, valganciclovir, vicriviroc, vidarabine, zalcitabine, zanamivir, and zidovudine. In preferred embodiments, the protease inhibitor is for coadministration with another antiviral agent, more preferably for coadministration with ritonavir. In other preferred embodiments the protease inhibitor is not for coadministration with another antiviral agent, particularly not for coadministration with ritonavir.

Anti-inflammatory agents are known in the art and a skilled can select a suitable anti-inflammatory agent in light of this disclosure. Examples of suitable antiviral agents are nonsteroidal anti-inflammatory drugs (NSAIDs), and corticosteroids. Examples of NSAIDs are pyrazolones (such as aminophenazone ampyrone azapropazone clofezone difenamizole famprofazone feprazone kebuzone metamizole mofebutazone morazonen ifenazone oxyphenbutazone phenazone phenylbutazone propyphenazone sulfinpyrazone and suxibuzone), salicylates (such as acetylsalicylic acid, aloxiprin, benorylate, carbasalate calcium, diflunisal, dipyrrocetyl, ethenzamide, guacetisal, magnesium salicylate, methyl salicylate, salsalate, salicin, salicylamide, salicylic acid, and sodium salicylate), acetic acid derivatives (such as aceclofenac,

acemetacin, alclofenac, amfenac, bendazac, bromfenac, bumadizone, bufexamac, diclofenac, difenpiramide, etodolac, felbinac, fenclozic acid, fentiazac, indomethacin, indometacin farnesyl, isoxepac, ketorolac, lonazolac, oxametacin, prodolic acid, proglumetacin, sulindac, tiopinac, tolmetin, and zomepirac), oxicams (such as ampiroxicam, droxicam, isoxicam, lornoxicam, meloxicam, piroxicam, and tenoxicam), propionic acid derivatives (profens, such as alminoprofen, benoxaprofen, carprofen, dexibuprofen, dexketoprofen, fenbufen, fenoprofen, flunoxaprofen, flurbiprofen, ibuprofen, ibuproxam, indoprofen, ketoprofen, loxoprofen, miroprofen, naproxen, oxaprozin, piketoprofen, pirprofen, suprofen, tarenfluril, tepoxalin, tiaprofenic acid, vedaprofen, zaltoprofen, and naproxcinod), N-arylanthranilic acids (fenamates, such as azapropazone, clonixin, etofenamate, floctafenine, flufenamic acid, flunixin, glafenine, meclofenamic acid, mefenamic acid, morniflumate, niflumic acid, tolfenamic acid, and flutiazin) coxibs (such as apricoxib, celecoxib, cimicoxib, deracoxib, etoricoxib, firocoxib, lumiracoxib, mavacoxib, parecoxib, robenacoxib, rofecoxib, and valdecoxib), aminopropionitrile, benzydamine chondroitin sulfate, diacerein, fluproquazone. Glucosamine, glycosaminoglycan, hyperforin, nabumetone, nimesulide, oxaceprol, proquazone, superoxide dismutase/orgotein, and tenidap. Examples of corticosteroids are hydrocortisone, hydrocortisone acetate, cortisone acetate, tixocortol pivalate, prednisolone, methylprednisolone, prednisone, amcinonide, budesonide, desonide, fluocinolone acetonide, fluocinonide, halcinonide, triamcinolone acetonide, beclometasone, betamethasone, dexamethasone, fluocortolone, halometasone, mometasone, alclometasone dipropionate, betamethasone dipropionate, betamethasone valerate, clobetasol propionate, clobetasone butyrate, fluprednidene acetate, mometasone furoate, ciclesonide, cortisone acetate, hydrocortisone aceponate, hydrocortisone acetate, hydrocortisone buteprate, hydrocortisone butyrate, hydrocortisone valerate, prednicarbate, and tixocortol pivalate.

The compositions comprising the compounds as described above, can be prepared as a medicinal or cosmetic preparation or in various other media, such as foods for humans or animals, including medical foods and dietary supplements. A "medical food" is a product that is intended for the specific dietary management of a disease or condition for which distinctive nutritional requirements exist. By way of example medical foods may include vitamin and mineral formulations fed through a feeding tube (referred to as enteral administration). A "dietary supplement" shall mean a product that is intended to supplement the human diet and is typically provided in the form of a pill, capsule, tablet or like formulation. By way of example a dietary supplement may include one or more of the following ingredients: vitamins, minerals, herbs, botanicals; amino acids, dietary substances intended to supplement the diet by increasing total dietary intake, and concentrates, metabolites, constituents, extracts or combinations of any of the foregoing. Dietary supplements may also be incorporated into food, including food bars, beverages, powders, cereals, cooked foods, food additives and candies; or other functional foods designed to promote health or to prevent or halt the progression of a treatment of a viral infection or a condition related to a viral infection.

The subject compounds and compositions may be compounded with other physiologically acceptable materials that can be ingested including foods. In addition, or alternatively, the compositions as described herein may be administered orally in combination with (the separate) administration of food.

5 The compositions or compound according to the invention may be administered alone or in combination with other pharmaceutical or cosmetic agents and can be combined with a physiologically acceptable carrier thereof. In particular, the compounds described herein can be formulated as pharmaceutical or cosmetic compositions by formulation with additives such as pharmaceutically or physiologically acceptable excipients carriers, and vehicles. Suitable
10 pharmaceutically or physiologically acceptable excipients, carriers and vehicles include processing agents and drug delivery modifiers and enhancers, such as, for example, calcium phosphate, magnesium stearate, talc, monosaccharides, disaccharides, starch, gelatin, cellulose, methyl cellulose, sodium carboxymethyl cellulose, dextrose, hydroxypropyl-P-cyclodextrin, polyvinylpyrrolidinone, low melting waxes, ion exchange resins, and the like, as well as
15 combinations of any two or more thereof. Other suitable pharmaceutically acceptable excipients are described in "Remington's Pharmaceutical Sciences, " Mack Pub. Co., New Jersey (1991), and "Remington: The Science and Practice of Pharmacy, " Lippincott Williams & Wilkins, Philadelphia, 22nd edition (2012).

20 Compositions for use according to the invention may be manufactured by processes well known in the art; e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes, which may result in liposomal formulations, coacervates, oil-in-water emulsions, nanoparticulate/microparticulate powders, or any other shape or form. Compositions for use in accordance with the invention thus
25 may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent on the route of administration chosen.

For injection, the compounds and compositions for use according to the invention may be
30 formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

Oral and parenteral administration may be used where the compounds and compositions
35 for use are formulated by combining them with pharmaceutically acceptable carriers well known in the art, or by using them as a food additive. Such strategies enable the compounds and compositions for use according to the invention to be formulated as tablets, pills, dragées, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a subject to be treated. Preparations or pharmacological preparations for oral use may be made with the use
40 of a solid excipient, optionally grinding the resulting mixture, and processing the mixture of

granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragée cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Additionally, coformulations may be made with uptake enhancers known in the art.

Dragée cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, PVP, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solution, and suitable organic solvents or solvent mixtures. Polymethacrylates can be used to provide pH-responsive release profiles so as to pass the stomach. Dyestuffs or pigments may be added to the tablets or dragée coatings for identification or to characterize different combinations of active compound doses.

Compounds and compositions which can be administered orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules may contain the active ingredients in admixture with a filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration.

For buccal administration, the compounds and compositions for use according to the invention may be administered in the form of tablets or lozenges formulated in a conventional manner.

The compounds and compositions for use according to the invention may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. In this way it is also possible to target a particular organ, tissue, tumor site, site of inflammation, etc. Formulations for infection may be presented in unit dosage form, e.g., in ampoules or in multi-dose container, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Compositions for parenteral administration include aqueous solutions of the compositions in water soluble form. Additionally, suspensions may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compositions to allow for the preparation of highly concentrated solutions.

Alternatively, one or more components of the composition may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compositions for use according to the invention may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional
5 suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds and compositions for use according to the invention may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, they may be formulated with
10 suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil), or as part of a solid or semi-solid implant that may or may not be auto-degrading in the body, or ion exchange resins, or one or more components of the composition can be formulated as sparingly soluble derivatives, for example, as a sparingly soluble salt. Examples of suitable polymeric materials are known to the person skilled in the art and include PLGA and polylactones such as
15 polycaproic acid.

The compositions for use according to the invention also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

20 The compositions for use according to the invention may also be comprised in a transdermal patch. Preferred transdermal patches for use according to the invention are selected from single-layer drug-in-adhesive patch, or multi-layer drug-in-adhesive patch, or reservoir patch, or matrix patch, or vapour patch.

Compositions for use according to the invention include compounds and compositions
25 wherein the active ingredients are contained in an amount effective to achieve their intended purposes. More specifically, a therapeutically effective amount means an amount of compound effective to prevent, stabilize, alleviate, revert, or ameliorate causes or symptoms of disease, or prolong the survival, mobility, or independence of the subject being treated. Determination of a therapeutically effective amount is within the capability of those skilled in the art, especially in light
30 of the detailed disclosure provided herein. For any compounds and compositions used in the invention, the therapeutically effective amount or dose can be estimated initially from cell culture assays. Dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g., Fingl, et al., 1975,
35 in "The Pharmacological Basis of Therapeutics" Ch. 1 p. 1). The amount of compound and compositions administered will, of course, be dependent on the subject being treated, on the subject's weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

A composition for use according to the invention may be supplied such that a compound
40 for use according to the invention and one or more of the other components as defined herein are

in the same container, either in solution, in suspension, or in powder form. A composition for use according to the invention may also be provided with all components provided separately from one another, for example to be mixed with one another prior to administration, or for separate or sequential administration. Various packaging options are possible and known to the ones skilled
5 in the art, depending, among others, on the route and mechanism of administration. In light of the methods of administration described above, the invention provides a compound for use according to the invention, or a composition for use according to the invention, characterized in that it is administered orally, sublingually, intravascularly, intravenously, subcutaneously, transdermally, or optionally by inhalation; preferably orally; more preferably subcutaneously.

10

An "effective amount" of a compound or composition is an amount which, when administered to a subject, is sufficient to reduce or eliminate either one or more symptoms of a disease, or to retard the progression of one or more symptoms of a disease, or to reduce the severity of one or more symptoms of a disease, or to suppress the manifestation of a disease, or
15 to suppress the manifestation of adverse symptoms of a disease. An effective amount can be given in one or more administrations.

The "effective amount" of that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host to which the active ingredient is administered and the particular mode of administration. The unit dosage chosen is usually
20 fabricated and administered to provide a desired final concentration of the compound in the blood.

The effective amount (i.e. the effective total daily dose), preferably for adults, is herein defined as a total daily dose of about 0.01 to 2000 mg, or about 0.01 to 1000 mg, or about 0.01 to 500 mg, or about 5 to 1000 mg, or about 20 to 800 mg, or about 30 to 800 mg or about 30 to 700 mg, or about 20 to 700 mg or about 20 to 600 mg, or about 30 to 600 mg, or about 30 to 500 mg,
25 about 30 to 450 mg or about 30 to 400 mg, or about 30 to 350 mg or about 30 to 300 mg or about 50 to 600 mg, or about 50 to 500 mg, or about 50 to 450 mg, or about 50 to 400 mg or about 50 to 300 mg, or about 50 to 250 mg, or about 100 to 250 mg or about 150 to 250 mg. In the most preferred embodiment, the effective amount is about 200 mg. In preferred embodiments, the invention provides a compound for use according to the invention, or a composition for use
30 according to the invention, characterized in that it is administered to a subject in an amount ranging from 0.1 to 1500 mg/day, preferably from 0.1 to 1000 mg/day, more preferably from 0.1 to 400 mg/day, still more preferably from 0.25 to 150 mg/day, such as about 100 mg/day.

Alternatively, the effective amount of the compound, preferably for adults, preferably is administered per kg body weight. The total daily dose, preferably for adults, is therefore about
35 0.05 to about 40 mg/kg, about 0.1 to about 20 mg/kg, about 0.2 mg/kg to about 15 mg/kg, or about 0.3 mg/kg to about 15 mg/kg or about 0.4 mg/kg to about 15 mg/kg or about 0.5 mg/kg to about 14 mg/kg or about 0.3 mg/kg to about 14 mg/kg or about 0.3 mg/kg to about 13 mg/kg or about 0.5 mg/kg to about 13 mg/kg or about 0.5 mg/kg to about 11 mg/kg.

The total daily dose for children is preferably at most 200 mg. More preferably the total daily
40 dose is about 0.1 to 200 mg, about 1 to 200 mg, about 5 to 200 mg about 20 to 200 mg about 40

to 200 mg, or about 50 to 200 mg. Preferably, the total daily dose for children is about 0.1 to 150 mg, about 1 to 150 mg, about 5 to 150 mg about 10 to 150 mg about 40 to 150 mg, or about 50 to 150 mg. More preferably, the total daily dose is about 5 to 100 mg, about 10 to 100 mg, about 20 to 100 mg about 30 to 100 mg about 40 to 100 mg, or about 50 to 100 mg. Even more preferably, the total daily dose is about 5 to 75 mg, about 10 to 75 mg, about 20 to 75 mg about 30 to 75 mg about 40 to 75 mg, or about 50 to 75 mg.

Alternative examples of dosages which can be used are an effective amount of the compounds for use according to the invention within the dosage range of about 0.1 μg /kg to about 300 mg/kg, or within about 1.0 μg /kg to about 40 mg/kg body weight, or within about 1.0 μg /kg to about 20 mg/kg body weight, or within about 1.0 μg /kg to about 10 mg/kg body weight, or within about 10.0 μg /kg to about 10 mg/kg body weight, or within about 100 μg /kg to about 10 mg/kg body weight, or within about 1.0 mg/kg to about 10 mg/kg body weight, or within about 10 mg/kg to about 100 mg/kg body weight, or within about 50 mg/kg to about 150 mg/kg body weight, or within about 100 mg/kg to about 200 mg/kg body weight, or within about 150 mg/kg to about 250 mg/kg body weight, or within about 200 mg/kg to about 300 mg/kg body weight, or within about 250 mg/kg to about 300 mg/kg body weight. Other dosages which can be used are about 0.01 mg/kg body weight, about 0.1 mg/kg body weight, about 1 mg/kg body weight, about 10 mg/kg body weight, about 20 mg/kg body weight, about 30 mg/kg body weight, about 40 mg/kg body weight, about 50 mg/kg body weight, about 75 mg/kg body weight, about 100 mg/kg body weight, about 125 mg/kg body weight, about 150 mg/kg body weight, about 175 mg/kg body weight, about 200 mg/kg body weight, about 225 mg/kg body weight, about 250 mg/kg body weight, about 275 mg/kg body weight, or about 300 mg/kg body weight.

The protease inhibitor can be given in a single dose, or multiple doses. Continuous administration also may be applied where appropriate. The dose of a therapeutic composition via continuous perfusion may be equivalent to that given by a single or multiple injections, adjusted over a period of time during which the perfusion occurs. The amount of protease inhibitor administered may be dependent on the subject being treated, the subject's weight, the manner of administration, and the judgment of the physician. Treatment regimens may vary as well, and often depend on the type and location of the damage or symptoms, disease progression, and health and age of the patient.

In some embodiments, a protease inhibitor may be administered to a patient systemically or by local injection. Systemic administration can be by intravenous or intraperitoneal delivery. The protease inhibitor can be administered to reach a circulating level of about 2 to 20 mg/ml in blood, or a dose of about 100-300 mg can be delivered to a patient.

Compounds or compositions for use according to the present invention may be administered in a single daily dose, or the total daily dosage may be administered in divided dosage of two, three or four times daily.

In a preferred embodiment of the invention, "subject", "individual", or "patient" is understood to be an individual organism, preferably a vertebrate, more preferably a mammal, even more preferably a primate and most preferably a human. A subject is preferably a subject who has been

diagnosed with a disease or condition as described herein, or who is suspected of suffering from a disease or condition as described herein, or who resides or has resided in an area where a disease or condition as described herein is prevalent, more preferably the subject has been diagnosed.

5 Compounds of the invention were found to not inhibit cytochrome p450. Accordingly, the invention is suitable for treating subjects who have comedication that is known to inhibit cytochrome p450. A skilled person knows which drugs inhibit cytochrome p450, and a subject may be aware of the fact and can inform the treating physician.

10 In a further preferred embodiment of the invention, the human is an adult, e.g. a person that is 18 years or older. In addition, it is herein understood that the average weight of an adult person is 62 kg, although the average weight is known to vary between countries. In another embodiment of the invention the average weight of an adult person is therefore between about 50 – 90 kg. It is herein understood that the effective dose as defined herein is not confined to subjects having an average weight. Preferably, the subject has a BMI (Body Mass Index) between 18.0 to 40.0
15 kg/m², and more preferably a BMI between 18.0 to 30.0 kg/m².

 Alternatively, the subject to be treated is a child, e.g. a person that is 17 years or younger. In addition, the subject to be treated may be a person between birth and puberty or between puberty and adulthood. It is herein understood that puberty starts for females at the age of 10 -11 years and for males at the age of 11 – 12 year. Furthermore, the subject to be treated may be a
20 neonate (first 28 days after birth), an infant (0-1 year), a toddler (1-3 years), a preschooler (3–5 years); a school-aged child (5–12 years) or an adolescent (13–18 years).

 To maintain an effective range during treatment, the compound or composition may be administered once a day, or once every two, three, four, or five days. However preferably, the
25 compound may be administered at least once a day. Hence in a preferred embodiment, the invention pertains to a compound for use according to the invention, or a composition for use according to the invention, characterized in that it is administered to a subject 4, 3, 2, or 1 times per day or less, preferably 1 time per day. The total daily dose may be administered as a single daily dose. Alternatively, the compound is administered at least twice daily. Hence, the compound
30 as defined herein may be administered once, twice, three, four or five times a day. As such, the total daily dose may be divided over the several doses (units) resulting in the administration of the total daily dose as defined herein. In a preferred embodiment, the compound is administered twice daily. It is further understood that the terms “twice daily”, “bid” and “*bis in die*” can be used interchangeable herein.

35 In preferred embodiments the compound or composition is administered once every week, more preferably once every 8, 9, 10, 11, 12, 13, or 14 days, most preferably once every 14 days. In other embodiments the compound or composition is administered once every 15, 16, 17, 18, 19, 20, or 21 days, such as once every 21 days.

 In a preferred embodiment, the total daily dose is divided over several doses per day.
40 These separate doses may differ in amount. For example, for each total daily dose, the first dose

may have a larger amount of the compound than the second dose or *vice versa*. However preferably, the compound is administered in similar or equal doses. Therefore, in a most preferred embodiment, the compound is administered twice daily in two similar or equal doses.

In a further preferred embodiment of the invention, the total daily dose of the compound as defined herein above is administered in at least two separate doses. The interval between the administration of the at least two separate doses is at least about 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 hours, preferably the interval between the at least two separate doses is at least about 4, 5, 6, 7, 8, 9, 10, 11 or 12 hours and more preferably the interval between the at least two separate doses is at least about 8, 9, 10, 11 or 12 hours.

10

Also provided herein are kits, suitable for and for treating a disease or condition as described herein. A kit can include any of the compounds or compositions described herein. A kit can further include one or more additional agents, such as a protease inhibitor or an antiviral agent as described elsewhere herein. A kit can additionally include directions for use of the kit (e.g., instructions for treating a disease or condition as described elsewhere herein in a subject), a container, a means for administration, or a means for pain relief.

General definitions

When a structural formula or chemical name is understood by the skilled person to have chiral centers, yet no chirality is indicated, for each chiral center individual reference is made to all three of either the racemic mixture (having any enantiomeric excess), the pure R enantiomer, and the pure S enantiomer.

Compounds and compounds for use provided in this invention can be optionally substituted. Suitable optional substitutions are replacement of -H by a halogen. Preferred halogens are F, Cl, Br, and I. Further suitable optional substitutions are substitution of one or more -H by -NH₂, -OH, =O, alkyl, alkoxy, haloalkyl, haloalkoxy, alkene, haloalkene, alkyne, haloalkyn, and cycloalkyl. Alkyl groups have the general formula C_nH_{2n+1} and may alternately be linear or branched. Unsubstituted alkyl groups may also contain a cyclic moiety, and thus have the concomitant general formula C_nH_{2n-1}. Optionally, the alkyl groups are substituted by one or more substituents further specified in this document. Examples of alkyl groups include methyl, ethyl, propyl, 2-propyl, t-butyl, 1-hexyl, 1-dodecyl, etc. Throughout this application, the valency of atoms should always be fulfilled, and H can be added or removed as required.

Unless stated otherwise, -H may optionally be replaced by one or more substituents independently selected from the group consisting of C₁ – C₁₂ alkyl groups, C₂ – C₁₂ alkenyl groups, C₂ – C₁₂ alkynyl groups, C₃ – C₁₂ cycloalkyl groups, C₅ – C₁₂ cycloalkenyl groups, C₈ – C₁₂ cycloalkynyl groups, C₁ – C₁₂ alkoxy groups, C₂ – C₁₂ alkenyloxy groups, C₂ – C₁₂ alkynyloxy groups, C₃ – C₁₂ cycloalkyloxy groups, halogens, amino groups, oxo and silyl groups, wherein the silyl groups can be represented by the formula (R²)₃Si-, wherein R² is independently selected from the group consisting of C₁ – C₁₂ alkyl groups, C₂ – C₁₂ alkenyl groups, C₂ – C₁₂ alkynyl groups, C₃ – C₁₂ cycloalkyl groups, C₁ – C₁₂ alkoxy groups, C₂ – C₁₂ alkenyloxy groups, C₂ – C₁₂ alkynyloxy

groups and C₃ – C₁₂ cycloalkyloxy groups, wherein the alkyl groups, alkenyl groups, alkynyl groups, cycloalkyl groups, alkoxy groups, alkenyloxy groups, alkynyloxy groups and cycloalkyloxy groups are optionally substituted, the alkyl groups, the alkoxy groups, the cycloalkyl groups and the cycloalkoxy groups being optionally interrupted by one of more hetero-atoms selected from the group consisting of O, N and S. Preferably, these optional substitutions comprise no more than twenty atoms, more preferably no more than fifteen atoms.

Whenever a parameter of a substance is discussed in the context of this invention, it is assumed that unless otherwise specified, the parameter is determined, measured, or manifested under physiological conditions. Physiological conditions are known to a person skilled in the art, and comprise aqueous solvent systems, atmospheric pressure, pH-values between 6 and 8, a temperature ranging from room temperature to about 37 °C (from about 20 °C to about 40 °C), and a suitable concentration of buffer salts or other components. It is understood that charge is often associated with equilibrium. A moiety that is said to carry or bear a charge is a moiety that will be found in a state where it bears or carries such a charge more often than that it does not bear or carry such a charge. As such, an atom that is indicated in this disclosure to be charged could be non-charged under specific conditions, and a neutral moiety could be charged under specific conditions, as is understood by a person skilled in the art.

In the context of this invention, a decrease or increase or inhibition of a parameter to be assessed can mean a change of at least 5% of the value corresponding to that parameter. More preferably, a decrease or increase of the value means a change of at least 10%, even more preferably at least 20%, at least 30%, at least 40%, at least 50%, at least 70%, at least 90%, or 100%. In this latter case, it can be the case that there is no longer a detectable value associated with the parameter. Such terms can include, but do not necessarily include, complete elimination.

The use of a compound or composition as a medicament as described in this document can also be interpreted as the use of said compound or composition in the manufacture of a medicament. Similarly, whenever a compound or composition is used for as a medicament, it can also be used for the manufacture of a medicament, or in a method.

In this document and in its claims, the verb "to comprise" and its conjugations is used in its non-limiting sense to mean that items following the word are included, but items not specifically mentioned are not excluded. In addition, reference to an element by the indefinite article "a" or "an" does not exclude the possibility that more than one of the element is present, unless the context clearly requires that there be one and only one of the elements. The indefinite article "a" or "an" thus usually means "at least one". The word "about" or "approximately" when used in association with a numerical value (e.g. about 10) preferably means that the value may be the given value (of 10) more or less 10% of the value, or more or less 1% of the value.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

In the context of this invention, a cell or a sample can be a cell or a sample from a sample obtained from a subject. Such an obtained sample can be a sample that has been previously

obtained from a subject. Such a sample can be obtained from a human subject. Such a sample can be obtained from a non-human subject.

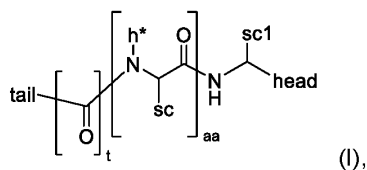
A salt is preferably a pharmaceutically acceptable salt. A salt is preferably a base addition salt wherein a cationic counterion is present. Examples of suitable salts are non-metallic salts such as ammonia salts, and metallic salts such as sodium salts and potassium salts. A skilled person can select suitable salt forms, and their means of production are well known (see e.g. "Occurrence of pharmaceutically acceptable anions and cations in the Cambridge Structural Database" Haynes et al., DOI: 10.1002/jps.20441). A salt can also be an acid addition salt. Acid addition salts are known in the art and examples are HCl salts and acetic acid salts. It should be recognized that the particular anion or cation forming a part of any salt of this invention is not critical, so long as the salt, as a whole, is pharmacologically acceptable. Additional examples of pharmaceutically acceptable salts and their methods of preparation and use are presented in Handbook of Pharmaceutical Salts: Properties, Selection and Use (2002). These salts can be prepared in situ during the isolation and purification of the compounds or by separately reacting the purified compound in its free base form with a suitable organic or inorganic acid and isolating the salt thus formed. Representative salts include the hydrobromide, hydrochloride, sulfate, bisulfate, nitrate, acetate, oxalate, valerate, oleate, palmitate, stearate, laurate, borate, benzoate, lactate, phosphate, tosylate, citrate, maleate, fumarate, succinate, tartrate, naphthylate mesylate, glucoheptonate, lactobionate, methane sulphonate, and laurylsulphonate salts, and the like. These may include cations based on the alkali and alkaline earth metals, such as sodium, lithium, potassium, calcium, magnesium, and the like, as well as non-toxic ammonium, quaternary ammonium, and amine cations including, but not limited to ammonium, tetramethylammonium, tetraethylammonium, methylamine, dimethylamine, trimethylamine, triethylamine, ethylamine, and the like (See S.M. Barge et al., J. Pharm. Sci. (1977) 66, 1).

The phrase "pharmaceutically-acceptable" or "pharmacologically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. As used herein, unless stated otherwise, an EC_{50} value generally represents a value determined in a cell-based assay and an IC_{50} value represents a value determined in a cell-free biochemical assay.

Herein, boldface can be used in variables to assist the reader; it does not imply further definition. As used herein, subject means both mammals and non-mammals. Mammals include, for example, humans; non-human primates, e.g., apes and monkeys; cattle; horses; sheep; rats; mice; pigs; and goats. Non-mammals include, for example, fish and birds. It is contemplated that any embodiment of a method or composition described herein can be implemented with respect to any other method or composition described herein. The examples below are intended to further illustrate certain aspects of the methods and compositions described herein, and are not intended to limit the scope of the claims. The following are embodiments of the invention:

55

1. A compound of general formula (I), or a salt thereof:



wherein

sc1 and each instance of **sc** are independently an amino acid side chain;

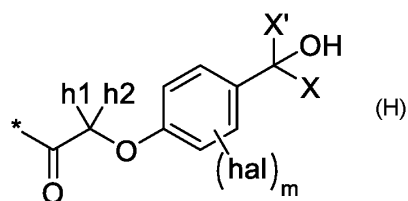
h* is H or -CH₃ or -CH₂CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain;

aa is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

t is 0 or 1;

tail is a linear, branched, or cyclic C1-22 alkyl, alkoxy, or alkylamine, wherein the alkyl, alkoxy, or alkylamine is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or is an independently chosen instance of **sc1**;

head is a warhead of general formula (H)



wherein

h1 and **h2** are each independently H, halogen, or C1-4 alkyl, or **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure;

hal is a halogen;

m is 0, 1, 2, 3, or 4;

X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure.

2. The compound according to embodiment 1, wherein

h1 and **h2** are H;

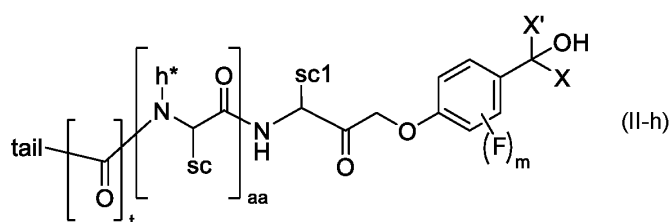
hal is F;

m is 3 or 4, preferably 4; and/or

X and **X'** are each independently H or C1-2 alkyl, preferably H or -CH₃.

3. The compound according to embodiment 1 or 2, wherein it is of general formula (II-h)

56



wherein

m is 1, 2, 3, or 4;

X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-4 membered cyclic structure, preferably **X** and **X'** are each independently H or -CH₃.

4. The compound according to any one of embodiments 1-3, wherein each instance of **sc** is independently

a linear, branched, or cyclic C1-12 alkyl, wherein alkyl is optionally unsaturated and optionally substituted with halogen, or

-(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, or

sc forms a cyclic amino acid side chain together with its adjacent **h***, wherein the cyclic amino acid side chain together with the carbon and nitrogen atoms to which it is attached forms a 4-7 membered ring.

5. The compound according to any one of embodiments 1-4, wherein **tail** is

a linear or branched C1-20 alkyl or alkoxy, wherein the alkyl or alkoxy is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or

-(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, alkyl, alkoxy, or -C(=O)O(CH₂)₀₋₄H.

6. The compound according to any one of embodiments 1-5, wherein

aa is 1 or 2, preferably 1;

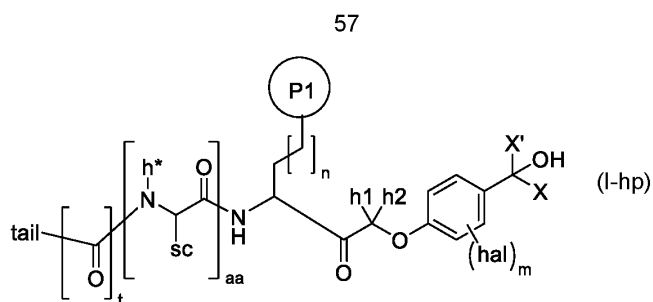
t is 0 or 1, preferably 1;

h* is H or together with its adjacent **sc** forms a cyclic amino acid side chain;

sc is butyl such as 2-methyl-n-propyl, is optionally substituted -CH₂-aryl such as -CH₂-fluorophenyl, or is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, wherein optional substitutions are preferably selected from halogen of alkoxy; and/or

tail is linear C12-18 alkyl such as -C₁₅H₃₁, is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, more preferably methoxyl-1*H*-indol-3-yl, or is optionally substituted C4-11 alkoxy such as -O-CH₂-phenyl, wherein optional substitutions are preferably selected from halogen or alkoxy, preferably alkoxy, more preferably methoxyl.

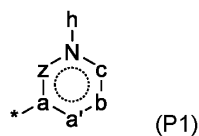
7. The compound according to any one of embodiments 1-6, wherein it is of general formula (I-hp)



wherein

n is 0 or 1;

P1 is a cyclic structure of general formula (P1):



5

wherein

a is C, CH, or N;

a' is absent or is CH or -C(=O)- or N or NH;

b is CH or N or NH;

c is CH or -C(=O)- or N or NH;

10

h is H or is absent; and

z is -C(=O)- or N.

8. The compound according to embodiment 7, wherein

when **a** is CH, **n** is 0;

when **a'** is CH, **n** is 0;

15

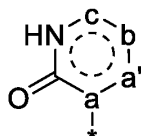
when **a'** is absent and **z** is N, **n** is 0;

when **a'** is absent, **n** is 1; or

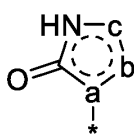
when **a'** is absent and **a** is N and **z** is -C(=O)-, **n** is 1;

or wherein **a** is N or C.

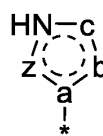
9. The compound according to embodiment 7 or 8, wherein **P1** is selected from



(P1a)



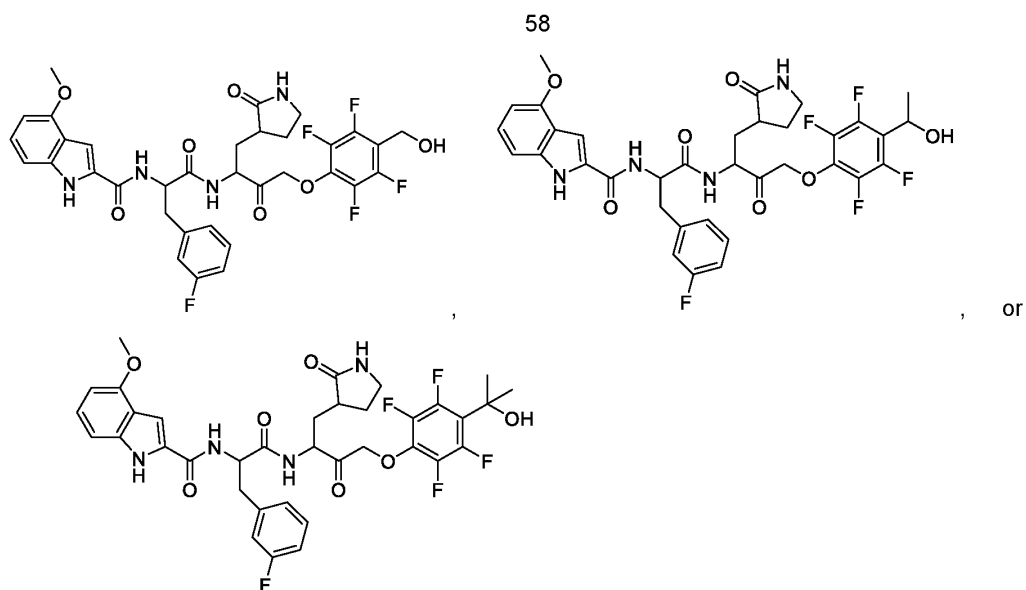
(P1b)



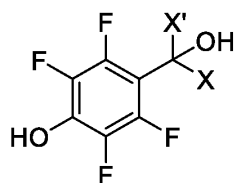
(P1c).

20

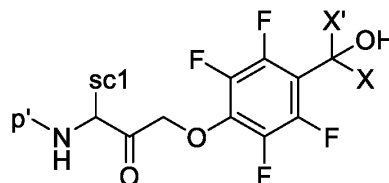
10. The compound according to any one of embodiments 1-9, wherein it is selected from



11. A composition comprising
 a compound as defined in any one of embodiments 1-10, and
 a pharmaceutically acceptable excipient,
 preferably wherein the composition is a pharmaceutical composition.
12. A compound as defined in any one of embodiments 1-10, or a composition as defined in
 embodiment 11, for use as a medicament.
13. A compound according to general formula (intermediate-1) or (intermediate-2):



(intermediate-1)



(intermediate-2)

wherein

p' is H or a protecting group;

sc1 is an amino acid side chain;

X is H, halogen, or C1-4 alkyl; preferably **X** is H, halogen, or $-(CH_2)_{0-3}-CH_3$;

X' is H, halogen, or C1-4 alkyl; preferably **X'** is H, halogen, or $-(CH_2)_{0-3}-CH_3$;

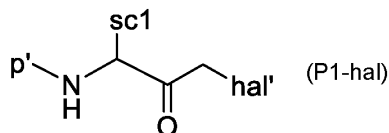
or **X** and **X'** together with the carbon atom to which they are attached form a 3-6
 membered cyclic structure;

wherein for the compound of general formula (intermediate-1) **X'** is halogen or
 C1-4 alkyl when **X** is H.

14. Method for providing a compound according to any one of embodiments 1-10, the method
 comprising the step of

59

- i) contacting a compound according to general formula (intermediate-1) according to embodiment 13 with a compound of general formula (P1-hal):



wherein

5

p' is a protecting group;

sc1 is an amino acid side chain;

hal' a halogen, preferably chloride;

to obtain a compound of general formula (intermediate-2) according to embodiment 13; and/or

10

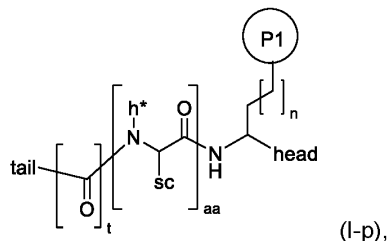
- ii) removing the protecting group from a compound according to general formula (intermediate-2) according to embodiment 13, preferably by using a carboxylic acid such as trifluoroacetic acid.

15. A method for inhibiting a viral protease, the method comprising the steps of:

15

- i) providing a compound as defined in any one of embodiments 1-10, or a composition as defined in embodiment 11;
- ii) contacting the viral protease with the provided compound or composition.

16. A compound of general formula (I-p), or a salt thereof:



wherein

20

sc is in each instance independently an amino acid side chain;

h* is H or -CH₃ or -CH₂CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain;

aa is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

t is 0 or 1;

25

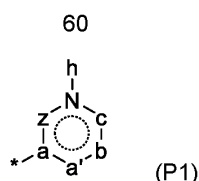
tail is a linear, branched, or cyclic C1-22 alkyl, alkoxy, or alkylamine, wherein the alkyl, alkoxy, or alkylamine is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or is an independently chosen instance of **sc¹**;

head is a warhead;

30

n is 0 or 1;

and wherein **P1** is a cyclic structure of general formula (P1):

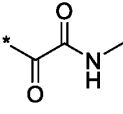
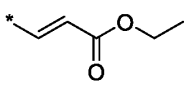
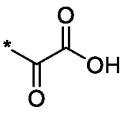
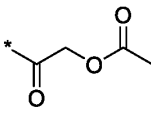
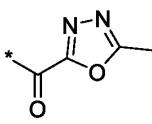
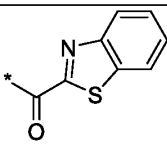
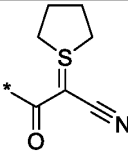
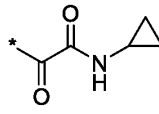
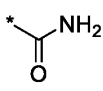


wherein

- a** is C, CH, or N;
a' is absent or is CH or -C(=O)- or N or NH;
b is CH or N or NH;
c is CH or -C(=O)- or N or NH;
h is H or is absent; and
z is -C(=O)- or N.

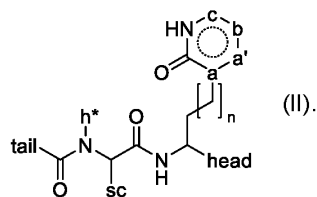
17. The compound according to embodiment 16, wherein **sc** is
 10 a linear, branched, or cyclic C1-12 alkyl, wherein alkyl is optionally unsaturated and optionally substituted with halogen, or
 -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen.
18. The compound according to embodiment 16 or 17, wherein **tail** is
 15 a linear or branched C1-20 alkyl or alkoxy, wherein the alkyl or alkoxy is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or
 -(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, alkyl, alkoxy, or
 -C(=O)O(CH₂)₀₋₄H.
19. The compound according to any one of embodiments 16-18, wherein
 20 **sc** is butyl such as 2-methyl-n-propyl, is optionally substituted -CH₂-aryl such as -CH₂-fluorophenyl, or is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, wherein optional substitutions are preferably selected from halogen of alkoxy; and/or
tail is linear C12-18 alkyl such as -C₁₅H₃₁, is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, more preferably methoxyl-1*H*-indol-3-yl, or is
 25 optionally substituted C4-11 alkoxy such as -O-CH₂-phenyl, wherein optional substitutions are preferably selected from halogen or alkoxy, preferably alkoxy, more preferably methoxyl.
20. The compound according to any one of embodiments 16-19, wherein **head** is a warhead selected from:

(wh1)	(wh2)	(wh3)	(wh4)
(wh5)	(wh6)	(wh7)	(wh8)

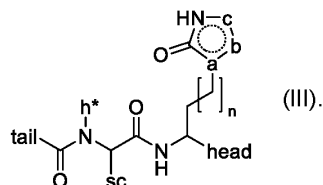
			
(wh9)	(wh10)	(wh11)	(wh12)
			
(wh13)	(wh14)	(wh15)	(wh16)
			
(wh17)	(wh18)	(wh19)	(wh20)

wherein **aryl** is a 5-6-membered aromatic (hetero)cycle that is optionally substituted with 1, 2, 3, 4, or 5 halogen atoms and that is optionally substituted with 1, 2, or 3 C1-C3 (hydroxy)alkyl moieties, and wherein (wh2) is preferably of general formula (H).

21. The compound according to any one of embodiments 16-20, wherein
- 5 when **a** is CH, **n** is 0;
when **a'** is CH, **n** is 0;
when **a'** is absent and **z** is N, **n** is 0;
when **a'** is absent, **n** is 1; or
when **a'** is absent and **a** is N and **z** is -C(=O)-, **n** is 1.
- 10 22. The compound according to any one of embodiments 16-21, wherein it is of general formula (II):

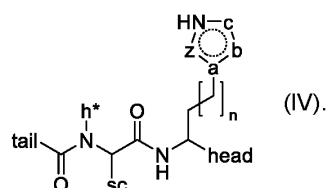


23. The compound according to any one of embodiments 16-22, wherein it is of general formula (III):

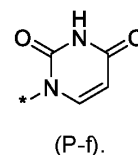
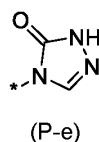
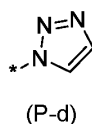
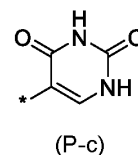
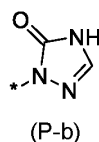
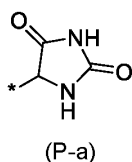


24. The compound according to any one of embodiments 16-21, wherein it is of general formula (IV):
- 15

62



25. The compound according to any one of embodiments 16-25, wherein **a** is N or C.
26. The compound according to any one of embodiments 16-21, wherein P1 is selected from:



27. The compound according to embodiment 26, wherein
n is 0 when **P1** is (P-a), (P-c), (P-d), or (P-f); and wherein
n is 1 when **P1** is (P-b) or (P-e).

28. A composition comprising
 a compound as defined in any one of embodiments 16-27, and
 a pharmaceutically acceptable excipient,
 preferably wherein the composition is a pharmaceutical composition.

29. A compound as defined in any one of embodiments 16-27, or a composition as defined in embodiment 28, for use as a medicament.
30. A method for inhibiting a viral protease, the method comprising the steps of:
 i) providing a compound as defined in any one of embodiments 16-27, or a composition as defined in embodiment 28;
 ii) contacting the viral protease with the provided compound or composition.

Examples

Example 1 - synthetic procedures

1.1.1 General procedure 1: Chloromethylketone coupling to phenol

Chloromethylketone (1 equiv) was dissolved in dry DMF (20 vol), followed by addition of phenol (1.2 equiv) and potassium fluoride (1.5 equiv). The reaction mixture was stirred for 16 h at 60 °C. The reaction mixture was then diluted with water and the aqueous layer was extracted with EtOAc (3 x 25 mL). The combined organic layer was washed with sat. NaCl, dried over Na₂SO₄, filtered, and concentrated *in vacuo*. Unless stated otherwise, the crude product was purified by flash column chromatography (0→5% MeOH in DCM).

1.1.2 General procedure 2: Boc deprotection

Boc-protected amino acid (1.1 equiv) was dissolved in dry DCM (20 vol) and the solution was cooled in an ice bath, followed by the dropwise addition of TFA (20%/DCM). The reaction mixture was stirred for 1 h at room temperature. Subsequently, the reaction mixture was co-
5 evaporated with diethyl ether (5 x 5 mL), dried *in vacuo* and the obtained TFA salt was used without further purification.

1.1.3 General procedure 3: Peptide coupling

Carboxylic acid (1 equiv) was introduced in a three-necked round-bottomed flask, after
10 which three vacuum-backfill cycles were performed. Subsequently, peptide-grade DMF (20 vol) was added, and the solution was cooled in an ice bath. Then, HATU (1.5 equiv) and DIPEA (3 equiv) were added, and the reaction mixture was stirred for 20 min at 0 °C. Lastly, Boc-deprotected TFA salt (1.1 equiv) was added and the reaction mixture was stirred at room temperature for 3 h. The reaction mixture was poured into sat. aq. NH₄Cl (20 mL) and extracted
15 with EtOAc (3 x 20 mL). The combined organic layers were washed with sat. aqueous NaHCO₃ (2 x 50 mL), water (2 x 50 mL) and brine (2 x 50 mL), dried over Na₂SO₄, filtered and concentrated *in vacuo*. Unless stated otherwise, the crude product was purified by flash column chromatography (0→5% MeOH in DCM).

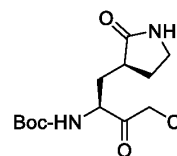
1.1.4 General procedure 4: Solid Phase Peptide Synthesis of dipeptide acids

A solution of Fmoc-protected amino acid (3.0 equiv) and DIPEA (5.0 equiv) in DMF (10 vol) was added to 2-chloro CTC resin (10 g, original loading rate: 1.2 mmol/g) and gently agitated under nitrogen bubbling for 16 h. Reagents were drained, and the resin was washed with DMF (2 x 10 vol), IPA (2 x 10 vol) and DMF (2 x 10 vol) sequentially for each 5 min. The unreacted
25 chlorides in 2-CTC resin were capped with MeOH/DIPEA/DMF (15/5/80) for 15 min. The reagents were drained and washed with DCM. Subsequent Fmoc deprotection was performed with a mixture of 20 % piperidine in DMF (10 vol) for 2 x 10 min by gently agitating under nitrogen bubbling. The mixture was drained and washed with DMF (2 x 10 vol), IPA (2 x 10 vol) and finally with DMF (2 x 10 vol). A solution of Fmoc-protected amino acid (3.0 equiv), Oxyma (3.0 equiv)
30 and DIC (4.0 equiv) in DMF (10 vol) was added to the above resin and gently agitated under nitrogen bubbling for 2 h. Completion of coupling was monitored by Kaiser test. Upon a negative test result, reagents were drained, and the resin was washed with DMF (2 x 10 vol), IPA (2 x 10 vol) and finally with DCM (2 x 10 vol) for each 5 min. The dipeptide was cleaved off the resin using 30% HFIP in DCM as cleavage cocktail (10 vol) for 2 x 30 min, filtered and washed the resin
35 with DCM (10 vol). The combined filtrates were concentrated under vacuum. The obtained crude peptide mass was triturated with diethyl ether to get off-white solid. It was filtered, washed two times with ether followed by drying under vacuum for 2 h, yielding the desired dipeptide.

1.2 Intermediates for inhibitor synthesis

tert-Butyl ((S)-4-chloro-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)carbamate (i1): A three-necked flame-dried flask (100 mL) equipped with a nitrogen inlet and internal thermometer was charged with (S)-methyl 2-((tert-

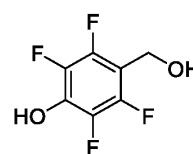
5 butoxycarbonyl)amino-3-((S)-2-oxopyrrolidin-3-yl)propanoate (1 equiv, 500 mg, 1.75 mmol), chloriodomethane (4 equiv, 507 μ L, 6.98 mmol), and dry THF, and the solution was cooled to -77 °C. Lithium diisopropylamide (6 equiv, 2M, 5.24 mL, 10.5 mmol) in THF/hexane was added via a pressure-equalizing dropping funnel at such a rate to keep the internal temperature below -70 °C. After complete addition, the reaction mixture was stirred for another hour, before quenching with acetic acid (8 equiv, 800 μ L, 14.0 mmol) in THF (5 mL), over 20 min, while maintaining the temperature below -70 °C. The reaction mixture was diluted with EtOAc and water. The organic layer was collected and washed with water, sat. NaHCO₃ and brine. The combined organic layers were dried over Na₂SO₄, filtered, and concentrated *in vacuo*. The crude product was purified by flash column chromatography (20% acetone in toluene) to give **i1** as a brown oil (176 mg, 33%). *R*_f = 0.3 (acetone/toluene 1:1). ¹H NMR (500 MHz, *d*-DMSO) δ 7.65 (s, 1H), 7.52 (d, *J* = 7.6 Hz, 1H), 4.61 (ABq, *J* = 23.6 Hz, 2H), 4.16 (ddd, *J* = 11.3, 7.6, 4.0 Hz, 1H), 3.20–3.09 (m, 2H), 2.29–2.19 (m, 1H), 2.18–2.09 (m, 1H), 1.87 (ddd, *J* = 13.9, 10.9, 4.6 Hz, 1H), 1.71–1.56 (m, 2H), 1.38 (s, 9H). HRMS (*m/z*): [M+H]⁺ calcd for C₁₃H₂₁ClN₂O₄: 305.1263; found: 305.1247.



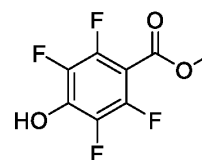
20

2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenol (i2): **2,3,5,6-tetrafluoro-4-hydroxybenzoic acid** (1 equiv, 500 mg, 2.38 mmol) was dissolved in anhydrous THF (2 mL). BH₃THF (4 equiv, 1M, 9.52 mmol) was added dropwise and the reaction was heated to reflux for 16 h. The reaction

25 mixture was quenched with 2N HCl, diluted with water (50 mL) and extracted with EtOAc (3 x 50 mL). The organics were combined, washed with sat. NaCl, dried over Na₂SO₄, filtered, and concentrated *in vacuo*. The crude product was purified by flash column chromatography (10→50% EtOAc in *n*-heptane) to yield **i2** as a white solid (392 mg, 84%). *R*_f = 0.3 (EtOAc/heptane 2:1). ¹H NMR (400 MHz, *d*-MeCN) δ 4.61 (t, *J* = 1.8 Hz, 2H). HRMS (*m/z*): [M-H]⁻ calcd for C₇H₃F₄O₂: 195.0075; found: 195.0066.



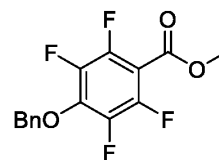
Methyl 2,3,5,6-tetrafluoro-4-hydroxybenzoate (i3): In a 250 mL round-bottom flask, **2,3,5,6-tetrafluoro-4-hydroxybenzoic acid** (10.0 g, 47.6 mmol) was dissolved in MeOH (50 mL, 5 vol) and cooled to 0 °C. To this, H₂SO₄ (5 mL, 0.5 vol) was added drop wise over 2 min. The reaction mixture was stirred at 70 °C for 4 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with ice cold water (200 mL) and extracted with EtOAc (3 x 500 mL). The combined organic layer was washed with sat. NaHCO₃ (200 mL) followed by brine (200 mL), dried over Na₂SO₄ and concentrated under vacuum to give crude material as brown oil (11 g). The crude material triturated with Et₂O to give **i3** (9.0 g, 85%) as an off-white solid, which



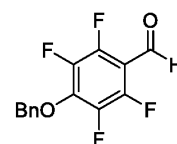
40

was used without further purification. $R_f = 0.5$ (MeOH/DCM 5:95). ^1H NMR (400 MHz, d -DMSO) δ 3.78 (s, 3H). ESI-MS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_8\text{H}_4\text{F}_4\text{O}_3$: 225.02; found: 225.10.

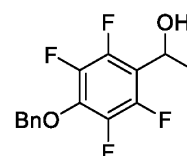
Methyl 4-(benzyloxy)-2,3,5,6-tetrafluorobenzoate (i4): In a 250 mL round-bottom flask, **methyl 2,3,5,6-tetrafluoro-4-hydroxybenzoate** (1.0 equiv, 9.0 g, 40.0 mmol) was dissolved in DMF (90 mL, 5 vol) and cooled to 0 °C. To this, was added K_2CO_3 (2.0 equiv, 11.08 g, 80.0 mmol) and benzyl bromide (1.5 equiv, 7.15 mL, 60.0 mmol). The reaction mixture was stirred at 25 °C for 6 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with ice cold water (200 mL) and extracted with EtOAc (3 x 500 mL). The combined organic layer was washed with ice cold water (2x200 mL) followed by brine (200 mL), dried over Na_2SO_4 and concentrated under vacuum to give crude material as brown oil (~15 g). The crude was purified by silica flash chromatography (20→25% EtOAc in pet-ether) to give **i4** (10 g, 79%) as an off-white solid. $R_f = 0.7$ (EtOAc/Pet-ether 6:4). ^1H NMR (400 MHz, CDCl_3) δ 7.45–7.34 (m, 5H), 5.34 (s, 2H), 3.94 (s, 3H). ESI-MS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{15}\text{H}_{10}\text{F}_4\text{O}_3$: 315.06; found: 315.18.



4-Benzyloxy-1,2,5,6-tetrafluorobenzaldehyde (i5): In a 250 mL round-bottom flask, **methyl 2,3,5,6-tetrafluoro-4-hydroxybenzoate** (1.0 equiv, 10 g, 31.84 mmol) was dissolved in THF (200 mL, 20 vol) and cooled to -78 °C. To this was added, LAH (2M in THF) (1.5 equiv, 24 mL, 47.7 mmol) dropwise over 10 min. The reaction mixture was stirred at -78 °C for 30 min. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with EtOAc (500 mL), quenched with sat. NH_4Cl solution (200 mL), filtered through celite and extracted with EtOAc (3 x 500 mL). The combined organic layer was washed with saturated brine solution (500 mL), dried over Na_2SO_4 and concentrated under vacuum to give crude material as brown gum (~10 g). The obtained crude was purified by silica flash chromatography (20→25% EtOAc in pet-ether) to give **i5** (6.0 g, 66%) as an off-white solid. $R_f = 0.6$ (EtOAc/Pet-ether 3:7). ^1H NMR (400 MHz, CDCl_3) δ 10.20 (t, $J = 1.18$ Hz, 1H), 7.45–7.35 (m, 5H), 5.42 (s, 2H). ESI-MS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{14}\text{H}_8\text{F}_4\text{O}_2$: 285.05; found: 285.20.

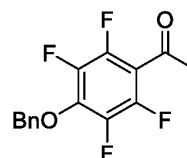


1-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)ethan-1-ol (i6): To a 250 mL round-bottom flask, was added MeMgBr (1M in THF) (5.0 equiv, 105 mL, 105.6 mmol) and THF (30 mL, 5 vol). To this, **4-(benzyloxy)-2,3,5,6-tetrafluorobenzaldehyde** (1.0 equiv, 6.0 g, 21.12 mmol) in THF (30 mL, 5 vol) was added dropwise over 10 min at 25 °C. The resulting reaction mixture was stirred at 25 °C for 2 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with saturated NH_4Cl (100 mL) and extracted with EtOAc (3 x 200 mL). The combined organic layer was washed with saturated brine solution (100 mL), dried over Na_2SO_4 and concentrated under vacuum to give crude material as brown gum. The obtained crude was purified by silica flash chromatography (30→40% EtOAc in pet-ether) to give **i6** (4.5 g, 71%) as an off-white solid. $R_f =$

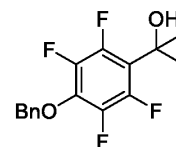


0.3 (EtOAc/Pet-ether 2:3). ^1H NMR (400 MHz, CDCl_3) δ 7.45–7.33 (m, 5H), 5.23 (s, 2H), 5.26–5.17 (m, 1H), 2.11 (dt, J = 7.87, 1.33 Hz, 1H), 1.63 (d, J = 6.75 Hz, 3H). ESI-MS (m/z): $[\text{M}-\text{OH}]^+$ calcd for $\text{C}_{15}\text{H}_{12}\text{F}_4\text{O}_2$: 283.07; found: 283.14.

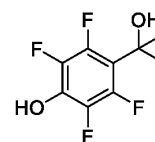
- 5 **1-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)ethan-1-one (i7)**: To a 250 mL round-bottom flask, was added PCC (5.0 equiv, 7.1 g, 33.33 mmol), 4 Å molecular sieves (7.1 g) and DCM (60 mL, 30 vol). The reaction mixture was cooled to 0 °C and a solution of **1-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)ethan-1-ol** (1.0 equiv, 2.0 g, 6.66 mmol) in DCM (20 mL, 10 vol) dropwise over 10 min. The resulting reaction mixture was stirred at 25 °C for 2 h. Progress of the reaction was monitored by TLC. Upon completion, it was filtered through 230-400 mesh silica and the filtrate was concentrated to vacuum, to give crude material as brown gum. The obtained crude was purified by silica flash chromatography (5→10% EtOAc in pet-ether) to give **i7** (1.5 g, 75%) as an off-white solid. R_f = 0.6 (EtOAc/Pet-ether 2:3). ^1H NMR (400 MHz, CDCl_3) δ 7.45–7.34 (m, 5H), 5.34 (s, 2H), 2.58 (t, J = 2.1 Hz, 3H). ESI-MS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{15}\text{H}_{10}\text{F}_4\text{O}_2$: 299.07; found: 299.18.



- 20 **2-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)propan-2-ol (i8)**: To a 100 mL round-bottom flask, was added MeMgBr (1M in THF) (5.0 equiv, 25 mL, 25.16 mmol) and THF (30 mL, 5 vol). To this, **1-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)ethan-1-one** (1.0 equiv, 1.5 g, 5.03 mmol) in THF (7.5 mL, 5 vol) was added drop wise over 10 min at 25 °C. The resulting reaction mixture was stirred at 25 °C for 1 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with saturated NH_4Cl (50 mL) and extracted with EtOAc (3 x 100 mL). The combined organic layer was washed with sat. brine solution (50 mL), dried over Na_2SO_4 and concentrated under vacuum to give crude material as brown gum. The crude was purified by silica flash chromatography (30→40% EtOAc in pet-ether) to give **i8** (1.1 g, 70%) as an off-white solid. R_f = 0.3 (EtOAc/Pet-ether 2:3). ^1H NMR (400 MHz, CDCl_3) δ 7.46–7.33 (m, 5H), 5.24 (s, 2H), 2.69 (t, J = 4.32 Hz, 1H), 1.71 (t, J = 2.08 Hz, 6H). ESI-MS (m/z): $[\text{M}-\text{OH}]^+$ calcd for $\text{C}_{16}\text{H}_{14}\text{F}_4\text{O}_2$: 297.09; found: 297.23.

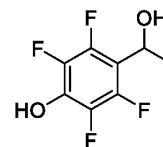


- 30 **2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenol (i9)**: In a 250 mL round-bottom flask, **2-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)propan-2-ol** (1.2 g, 3.82 mmol) was dissolved in MeOH (20 mL, 20 vol). The reaction mixture was flushed with H_2 and added 10% Pd-C (240 mg, 20%). The heterogenous mass was stirred at 25 °C under 1 atm H_2 pressure for 2 h. Progress of the reaction was monitored by TLC. Upon completion, it was filtered through celite and the filtrate was concentrated to vacuum to give crude material as yellow oil. The obtained crude was purified by silica flash chromatography (30→40% EtOAc in pet-ether) to give **i9** (0.7 g, 82%) as an off-white solid. R_f = 0.2 (EtOAc/Pet-ether 2:3). ^1H NMR (400 MHz, CDCl_3) δ 5.89 (bs, 1H), 2.72 (bs, 1H),

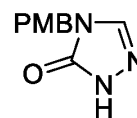


1.72 (t, $J = 2.07$ Hz, 6H), 1.67 (t, $J = 2.39$ Hz, 1H). ESI-MS (m/z): $[M-OH]^+$ calcd for $C_9H_8F_4O_2$: 207.04; found: 207.14.

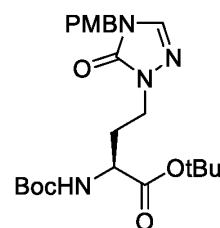
2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenol (i10): In a 250 mL round-bottom flask, **1-(4-(benzyloxy)-2,3,5,6-tetrafluorophenyl)ethan-1-ol** (1.0 g, 3.33 mmol) was dissolved in MeOH (20 mL, 20 vol). The reaction mixture was flushed with H_2 and 10 % Pd-C (200 mg, 20 %) was added. The reaction mixture was stirred at 25 °C under 1 atm H_2 pressure for 2 h. Progress of the reaction was monitored by TLC. Upon completion, it was filtered through celite and the filtrate was concentrated to vacuum to give crude material as yellow oil. The crude was purified by silica flash chromatography (30→40% EtOAc in pet-ether) to give **i10** (0.5 g, 71%) as an off-white solid. $R_f = 0.2$ (EtOAc/Pet-ether 2:3). 1H NMR (400 MHz, $CDCl_3$) δ 5.24 (q, $J = 6.74$ Hz, 1H), 1.65 (dt, $J = 6.71, 0.79$ Hz, 3H). ESI-MS (m/z): $[M-OH]^+$ calcd for $C_8H_6F_4O_2$: 193.03; found: 193.09.



4-(4-methoxybenzyl)-2,4-dihydro-3H-1,2,4-triazol-3-one (i11): In a round-bottom flask, **2,4-dihydro-3H-1,2,4-triazol-3-one 1** (2.5 g, 29.41 mmol) was dissolved in DMF (50 mL, 20 vol) and cooled to 0 °C. To this, was added K_2CO_3 (6.08 g, 44.1 mmol) followed by PMB-Cl (1.99 mL, 14.7 mmol). The reaction mixture was stirred at 25 °C for 4 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with ice cold water (120 mL) and extracted with EtOAc (3 x 100 mL). The combined organic layer was washed with ice cold water (3 x 50 mL) followed by brine (100 mL) and dried over Na_2SO_4 . It was concentrated under vacuum to give crude material as brown oil (~2.8 g). The crude was purified by silica flash chromatography (100% EtOAc in pet-ether) to give **i11** (1.0 g, 17%) as an off-white solid. $R_f = 0.4$ (EtOAc/Pet-ether 3:2). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{10}H_{11}N_3O_2$: 206.09; found: 206.07.

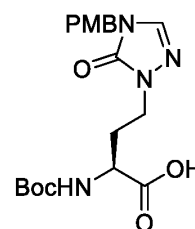


tert-Butyl (S)-2-((tert-butoxycarbonyl)amino)-4-(4-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)butanoate (i12): In a round-bottom flask, **4-(4-methoxybenzyl)-2,4-dihydro-3H-1,2,4-triazol-3-one** (1.0 g, 4.87 mmol) was dissolved in DMF (10 mL, 10 vol) and cooled to 0 °C. To this, was added Cs_2CO_3 (2.38 g, 7.30 mmol) and stirred for 5 min followed by the drop-wise addition of **tert-butyl (S)-2-((tert-butoxycarbonyl)amino)-4-iodobutanoate** (1.877 g, 4.873 mmol) in DMF (10 mL). The reaction mixture was stirred at 25 °C for 4 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with ice cold water (120 mL) and extracted with EtOAc (3 x 50 mL). The combined organic layer was washed with ice cold water (3 x 30 mL) followed by brine (60 mL), dried over Na_2SO_4 and concentrated under vacuum to give crude material as brown oil (~1.4 g). The crude was purified by silica flash chromatography (30% EtOAc in pet-ether) to give **i12** (2.0 g, 88 % yield) as an off-white solid. $R_f = 0.4$ (EtOAc/pet-ether 2:3). 1H NMR (400 MHz, $CDCl_3$) δ 7.25–7.21 (m, 3H), 6.92–6.86 (m, 2H), 5.31 (d, $J = 8.6$ Hz, 1H), 4.70 (s, 2H), 4.27 (d, $J =$



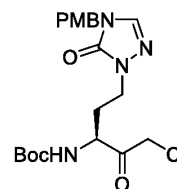
7.0 Hz, 1H), 3.89 (t, J = 7.4 Hz, 2H), 3.80 (s, 3H), 2.29–2.20 (m, 1H), 2.13–2.01 (m, 1H), 1.44 (s, 18H). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{23}H_{34}N_4O_6$: 463.26; found: 463.27.

5 **(S)-2-((tert-butoxycarbonyl)amino)-4-(4-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)butanoic acid (i13)**: According to GP2, **tert-butyl (S)-2-((tert-butoxycarbonyl)amino)-4-(4-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)butanoate** (2.0 g, 4.32 mmol) was deprotected and the crude reaction mixture was directly used without further purification. ESI-MS (m/z): $[M+H]^+$ calcd for $C_{14}H_{18}N_4O_4$: 307.14; found: 307.19.



A solution of **((S)-2-amino-4-(4-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)butanoic acid** (2.0 g, 6.52 mmol) in dioxane:H₂O (1:1) (40 mL, 20 vol) was cooled to 0 °C and NaHCO₃ was added (2.19 g, 26.11 mmol), followed by (Boc)₂O (4.49 mL, 19.58 mmol). The contents were stirred for 2 h at 25 °C. Progress of the reaction was monitored by TLC. After completion of reaction, it was diluted with water (40 mL, 20 vol) and extracted with EtOAc (3 x 60 mL). The combined organic layer was washed with brine solution (100 mL) and dried over Na₂SO₄ and concentrated under vacuum to give **i13** (2.0 g, 75 %) as an off-white solid. The crude reaction mixture was directly used without further purification. ESI-MS (m/z): $[M+H]^+$ calcd for $C_{19}H_{26}N_4O_6$: 407.19; found: 407.24.

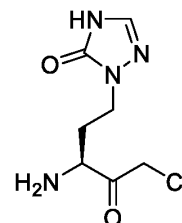
tert-Butyl (S)-((1-chloro-5-(4-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)-2-oxopentan-3-yl)carbamate (i14): A solution of **(S)-2-((tert-butoxycarbonyl)amino)-4-(4-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)butanoic acid** (2.0 g, 4.92 mmol) in THF (40 mL, 20 vol) was cooled to -10 °C and added triethylamine (0.887 mL, 6.39 mmol), followed by isobutyl chloroformate (0.76 mL, 5.90 mmol). The resulting reaction mixture was stirred at -10 °C for 30 min. After completion of reaction, the heterogeneous mixture was filtered and washed with THF (10 mL). The filtrate was taken in RBF and was cooled to -15 °C. To this, was added freshly prepared diazomethane in diethyl ether (20 mL) dropwise at -10 °C. The resulting reaction mixture was stirred for 30 min at -10 °C. Progress of the reaction was monitored by TLC. After completion of reaction, it was quenched with acetic acid until colorless and diluted with water (80 mL, 40 vol) and extracted with EtOAc (3 x 100 mL). The combined organic layer was washed with brine solution (100 mL) and dried over Na₂SO₄. The volatiles were removed under vacuum and the resulting yellow semisolid (2.0 g, crude) was used without further purification. R_f = 0.5 (EtOAc/Pet-ether 3:2). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{20}H_{26}N_6O_5$: 431.20; found: 403.70 (-N₂).



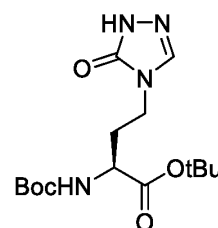
The obtained semisolid intermediate (2.0 g, 4.65 mmol) was dissolved in THF (40 mL, 20 vol) and the solution was cooled to -10 °C and 4M HCl in dioxane (4.65 mL, 18.6 mmol) was added. The reaction mixture was stirred for 30 min at the same temperature. Progress of the reaction was monitored by TLC. After completion of reaction, it was concentrated and triturated

with pentane (2 x 10 mL) to give **i14** (1.8 g, crude) as a light yellow gum. ESI-MS (m/z): [M+H]⁺ calcd for C₂₀H₂₇ClN₄O₅: 439.17; found: 439.58.

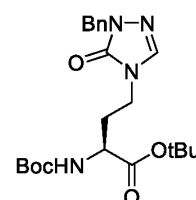
(S)-2-(3-amino-5-chloro-4-oxopentyl)-2,4-dihydro-3H-1,2,4-triazol-3-one (i15): To a solution of (*tert*-butyl (*S*)-(1-chloro-5-(4-methoxybenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)-2-oxopentan-3-yl)carbamate (400 mg, 0.911 mmol) in anisole (8 mL, 20 vol) at 25 °C was added TfOH (0.32 mL, 3.65 mmol). The contents were irradiated in microwave at 100 °C for 20 min. Progress of the reaction was monitored by TLC. After completion of reaction, it was concentrated under reduced pressure. The crude was diluted with water (8 mL, 20 vol) and extracted with EtOAc (3 x 15 mL). The combined organic layer was washed with water (3 x 10 mL), brine solution (10 mL) and dried over Na₂SO₄ and concentrated under vacuum to give **i15** (190 mg, crude) as a pale brown solid. This free amine was immediately coupled with the dipeptide. ESI-MS (m/z): [M+H]⁺ calcd for C₇H₁₁ClN₄O₂: 219.06; found: 219.36.



***tert*-Butyl (S)-2-((*tert*-butoxycarbonyl)amino)-4-(5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)butanoate (i16):** In a round-bottom flask, **2,4-dihydro-3H-1,2,4-triazol-3-one** (1.0 equiv, 3.7 g, 43.3 mmol) was dissolved in DMF (74 mL, 20 vol) and cooled to 0 °C. To this, was added K₂CO₃ (1.5 equiv, 8.6 g, 65.2 mmol) and ***tert*-butyl (S)-4-bromo-2-((*tert*-butoxycarbonyl)amino)butanoate** (0.5 equiv, 7.33 g, 21.7 mmol). The reaction mixture was stirred at 25 °C for 4 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with ice cold water (120 mL) and extracted with EtOAc (3 x 200 mL). The combined organic layer was washed with ice cold water (2 x 100 mL) followed by brine (100 mL), dried over Na₂SO₄ and concentrated under vacuum to give crude material as brown oil (~7.8 g). The crude was purified by silica flash chromatography (20→25% EtOAc in pet-ether) to give **i16** (4.7 g, 32%) as off-white solid. R_f = 0.7 (EtOAc/Pet-ether 3:2). ESI-MS (m/z): [M+H]⁺ calcd for C₁₅H₂₆N₄O₅: 343.20; found: 343.19.

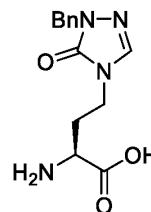


***tert*-Butyl (S)-2-((*tert*-butoxycarbonyl)amino)-4-(5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)butanoate (i17):** In a round-bottom flask, ***tert*-butyl (S)-2-((*tert*-butoxycarbonyl)amino)-4-(5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)butanoate** (1.0 equiv, 4.7 g, 13.7 mmol) was dissolved in DMF (47 mL, 10 vol) and cooled to 0 °C. To this, was added Cs₂CO₃ (1.5 equiv, 6.69 g, 20.6 mmol) and benzyl bromide (1.2 equiv, 2.02 mL, 16.4 mmol). The reaction mixture was stirred at 25 °C for 3 h. Progress of the reaction was monitored by TLC. Upon completion, it was diluted with ice cold water (100 mL) and extracted with EtOAc (3 x 200 mL). The combined organic layer was washed with ice cold water (2 x 100 mL) followed by brine (200 mL), dried over Na₂SO₄ and concentrated under vacuum to give crude material as brown oil (~6.5 g). The crude

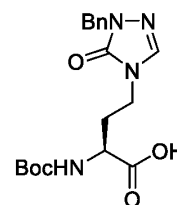


was purified by silica flash chromatography (10→15% EtOAc in pet-ether) to give **i17** (4.3 g, 73%) as off-white solid. R_f = 0.6 (EtOAc/pet-ether 2:3). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{22}H_{32}N_4O_5$: 433.25; found: 433.25.

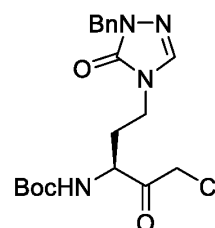
- 5 **(S)-2-amino-4-(1-benzyl-5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)butanoic acid (i18)**: To a stirred solution of **tert-butyl (S)-2-((tert-butoxycarbonyl)amino)-4-(5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)butanoate** (4.3 g, 9.95 mmol), 30 % TFA in DCM (86 mL, 20 vol) was added at 0 °C and brought to room temperature. The resulting reaction mixture was stirred at 25 °C
10 for 16 h. Progress of the reaction was monitored by LCMS. After completion of the reaction, it was concentrated under reduced pressure and triturated with diethyl ether to get **i18** (3.8 g, >99%) as an off-white solid. The crude material was used without further purification. ESI-MS (m/z): $[M+H]^+$ calcd for $C_{13}H_{16}N_4O_3$: 277.13; found: 277.13.



- 15 **(S)-4-(1-benzyl-5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)-2-((tert-butoxycarbonyl)amino)butanoic acid (i19)**: A solution of **(S)-2-amino-4-(1-benzyl-5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)butanoic acid** (3.8 g, 13.7 mmol) in dioxane:H₂O (1:1) (38 mL, 10 vol) was cooled to 0 °C and NaHCO₃ (4.68 g, 55.07 mmol) was added, followed by the addition of (Boc)₂O (9.47 mL, 41.3 mmol). The contents were stirred for 2 h at 25 °C. Progress of the reaction was monitored by TLC. After completion of reaction, it was diluted with water (38 mL, 10 vol) and extracted with EtOAc (3 x 100 mL). The combined organic layer was washed with ice cold water (3 x 60 mL), brine solution (100 mL) and dried over Na₂SO₄ and concentrated under vacuum to give **i19** (4.0 g) as an off-white solid. The crude material was used without further
25 purification. ESI-MS (m/z): $[M+H]^+$ calcd for $C_{18}H_{24}N_4O_5$: 377.18; found: 377.18.



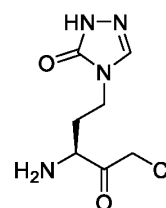
- tert-Butyl (S)-5-(1-benzyl-5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)-1-chloro-2-oxopentan-3-yl)carbamate (i20)**: A solution of **(S)-4-(1-benzyl-5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)-2-((tert-butoxycarbonyl)amino)butanoic acid** (4.0 g, 10.62 mmol) in THF (80 mL, 20 vol) was cooled to -10 °C and added triethylamine (1.915 mL, 13.81 mmol), followed by isobutyl chloroformate (1.65 mL, 12.75 mmol).
30 The resulting reaction mixture was stirred at -10 °C for 30 min. The heterogenous reaction mixture was filtered and washed with THF (20 mL). The filtrate was taken in a RBF and was cooled to -15 °C. To this, was added a freshly prepared diazomethane in diethyl ether (40 ml) dropwise at -10 °C. The resulting reaction mixture was stirred for 30 min at -10 °C. Progress of the reaction was monitored by TLC. After completion of reaction, it was quenched with acetic acid till it turns to colorless and diluted with water (160 mL, 40 vol) and extracted with EtOAc (3 x 200 mL). The combined organic layer was washed with brine solution (200 mL), dried over Na₂SO₄,
40 concentrated under vacuum and the resulting yellow semisolid (4.0 g, crude) was used without



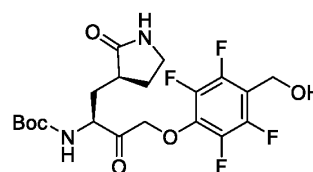
further purification. $R_f = 0.5$ (EtOAc/Pet-ether 3:2). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{19}H_{24}N_6O_4$: 401.19; found: 401.19.

The obtained semisolid intermediate (4.0 g, 9.9 mmol) was dissolved in THF (80 mL, 20 vol) and the solution was cooled to -10°C and 4M HCl in dioxane was added (9.292 mL, 37.169 mmol). The resulting reaction mixture was stirred for 30 min at the same temperature. Progress of the reaction was monitored by TLC. After completion of reaction, it was concentrated and triturated with pentane (2 x 10 mL) to give **i20** (3.0 g) as a light yellow gum. The crude material was used without further purification. ESI-MS (m/z): $[M+H]^+$ calcd for $C_{19}H_{25}ClN_4O_4$: 409.16; found: 409.57.

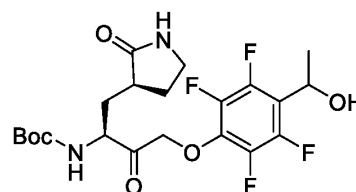
(S)-4-(3-amino-5-chloro-4-oxopentyl)-2,4-dihydro-3H-1,2,4-triazol-3-one (i21): To a solution of *tert*-butyl **(S)-{5-(1-benzyl-5-oxo-1,5-dihydro-4H-1,2,4-triazol-4-yl)-1-chloro-2-oxopentan-3-yl}carbamate** (300 mg, 0.73 mmol) in toluene (6 mL, 20 vol) at 25°C was added $\text{CF}_3\text{SO}_3\text{H}$ (0.21 mL, 2.48 mmol). The contents were irradiated with microwave at 100°C for 20 min. Progress of the reaction was monitored by TLC. After completion of reaction, it was concentrated under reduced pressure. The crude was diluted with water (6 mL, 20 vol) and extracted with EtOAc (3 x 10 mL). The combined organic layer was washed with ice cold water (3 x 10 mL), brine solution (10 mL) and dried over Na_2SO_4 . The volatiles were removed under vacuum to give **i21** (300 mg, crude) as an off-white solid. This free amine was immediately coupled with the dipeptide. ESI-MS (m/z): $[M+H]^+$ calcd for $C_7H_{11}ClN_4O_2$: 219.06; found: 219.12.



tert-Butyl **((S)-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)carbamate (i22):** According to GP1 starting from *tert*-butyl **((S)-4-chloro-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)carbamate** (130 mg, 0.427 mmol) and **2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenol** (100 mg, 0.512 mmol), **i22** was obtained as a white solid (109 mg, 55%). $R_f = 0.5$ (MeOH/DCM 1:9). ^1H NMR (400 MHz, d -DMSO) δ 7.65 (s, 1H), 7.70 (d, $J = 7.7$ Hz, 1H), 5.48 (t, $J = 5.81$ Hz, 1H), 5.22 (ABq, 2H), 4.51 (dt, $J = 5.87, 1.77$ Hz, 2H), 4.12 (ddd, $J = 11.47, 7.61, 4.23$ Hz, 1H), 3.20–3.08 (m, 2H), 2.29–2.18 (m, 1H), 2.18–2.09 (m, 1H), 1.93–1.83 (m, 1H), 1.70–1.53 (m, 2H), 1.39 (s, 9H). ^{19}F NMR (377 MHz, d -DMSO) δ -146.39 (dd, $J = 23.2, 8.8$ Hz), -157.49 (dd, $J = 23.2, 8.8$ Hz). HRMS (m/z): $[M+H]^+$ calcd for $C_{20}H_{24}F_4N_2O_6$: 465.1643; found: 465.1655.

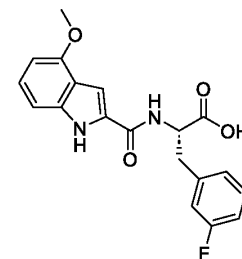


tert-Butyl **((2S)-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenoxy)butan-2-yl)carbamate (i23):** According to GP1 starting from *tert*-butyl **((S)-4-chloro-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)carbamate** (4.0 g, 13.15 mmol) and **2,3,5,6-**

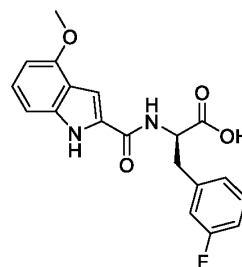


tetrafluoro-4-(1-hydroxyethyl)phenol (3.3 g, 15.78 mmol), **i23** was obtained as an off-white solid (3.5 g, 56%). $R_f = 0.4$ (EtOAc/Pet-ether 3:2). ^1H NMR (400 MHz, CDCl_3) δ 6.17 (d, $J = 7.25$ Hz, 1H), 5.85 (bs, 1H), 5.22 (t, $J = 6.34$ Hz, 1H), 5.07 (ABq, $J = 36.7$ Hz, 2H), 4.57–4.47 (m, 1H), 3.41–3.32 (m, 2H), 2.54–2.42 (m, 2H), 2.40–2.34 (m, 1H), 2.10–2.01 (m, 1H), 1.98–1.81 (m, 2H), 1.63 (d, $J = 6.74$ Hz, 3H), 1.44 (s, 9H). ESI-MS (m/z): calcd for $\text{C}_{21}\text{H}_{26}\text{F}_4\text{N}_2\text{O}_6$: 479.18; found: 479.19.

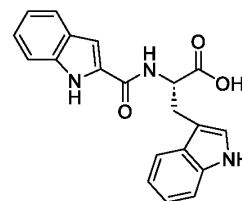
(S)-3-(3-fluorophenyl)-2-(4-methoxy-1H-indole-2-carboxamido)propanoic acid (i24): According to GP4 coupling using 2-chloro CTC resin (5.0 g, 1.2 mmol/g) with **Fmoc-3-fluoro-L-phenylalanine** (3.0 equiv) and **4-methoxy-1H-indole-2-carboxylic acid** (2.0 equiv), dipeptide **i24** was obtained as an off-white solid (2.0 g, 93%), which was used directly without further purification. ^1H NMR (500 MHz, d -DMSO) δ 12.83 (bs, 1H), 11.51 (d, $J = 2.4$, 1H), 8.66 (d, $J = 8.4$ Hz, 1H), 7.33–7.26 (m, 2H), 7.20–7.14 (m, 2H), 7.09 (t, $J = 7.9$ Hz, 1H), 7.02–6.95 (m, 2H), 6.50 (d, $J = 7.7$ Hz, 1H), 4.65 (ddd, $J = 10.8, 8.3, 4.3$ Hz, 1H), 3.88 (s, 3H), 3.22 (dd, $J = 13.8, 4.4$, 1H), 3.07 (dd, $J = 13.9, 10.8$ Hz, 1H). ^{19}F NMR (471 MHz, d -DMSO) δ -113.80. HRMS (m/z): $[\text{M}-\text{H}]^-$ calcd for $\text{C}_{19}\text{H}_{16}\text{FN}_2\text{O}_4$: 355.1100; found: 355.1109.



(R)-3-(3-fluorophenyl)-2-(4-methoxy-1H-indole-2-carboxamido)propanoic acid (i25): According to GP4 coupling using 2-chloro CTC resin (1.0 g, 1.2 mmol/g) with **Fmoc-3-fluoro-D-phenylalanine** (3.0 equiv) and **4-methoxy-1H-indole-2-carboxylic acid** (2.0 equiv), dipeptide **i25** was obtained as an off-white solid (400 mg, 93%), which was used directly without further purification. ^1H NMR (500 MHz, d -DMSO) δ 12.83 (bs, 1H), 11.51 (d, $J = 2.3$ Hz, 1H), 8.66 (d, $J = 8.4$ Hz, 1H), 7.33–7.25 (m, 2H), 7.20–7.13 (m, 2H), 7.09 (t, $J = 7.9$ Hz, 1H), 7.03–6.95 (m, 2H), 6.50 (d, $J = 7.7$ Hz, 1H), 4.65 (ddd, $J = 11.0, 8.4, 4.3$ Hz, 1H), 3.88 (s, 3H), 3.22 (dd, $J = 13.9, 4.4$ Hz, 1H), 3.07 (dd, $J = 13.8, 10.9$ Hz, 1H). ^{19}F NMR (471 MHz, d -DMSO) δ -113.80. HRMS (m/z): $[\text{M}-\text{H}]^-$ calcd for $\text{C}_{19}\text{H}_{16}\text{FN}_2\text{O}_4$: 355.1100; found: 355.1109.

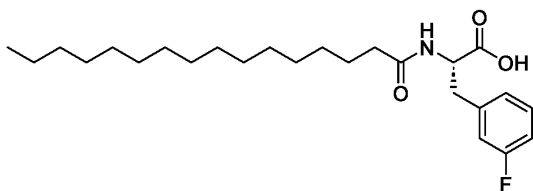


(1H-indole-2-carbonyl)-L-tryptophan (i26): Following GP4, using 2-chloro CTC resin (10.0 g, 1.2 mmol/g), **Fmoc-Trp(Boc)-OH** (3.0 equiv) and **4-methoxy-1H-indole-2-carboxylic acid** (2.0 equiv) were coupled. The dipeptide was cleaved off the resin using 10% TFA in DCM as cleavage cocktail, yielding **i26** as an off-white solid (4.0 g, 100%), which was used directly without further purification. ^1H NMR (500 MHz, d -DMSO) δ 12.30–11.52 (m, 1H), 10.78 (s, 1H), 9.23–8.21 (m, 1H), 7.62 (d, $J = 7.8$ Hz, 1H), 7.59 (d, $J = 8.0$ Hz, 1H), 7.43 (d, $J = 8.2$ Hz, 1H), 7.30 (d, $J = 7.6$ Hz, 1H), 7.23 (d, $J = 18.8$ Hz, 1H), 7.16 (t, $J = 7.6$ Hz, 1H), 7.10 (s, 1H), 7.06–6.99 (m, 2H), 6.95 (t, $J = 7.4$ Hz, 1H), 4.68 (s, 1H), 3.40 (d, $J =$



14.3 Hz, 1H), 3.24 (dd, $J = 14.5, 9.2$ Hz, 1H). HRMS (m/z): $[M+Na]^+$ calcd for $C_{20}H_{17}N_3O_3$: 370.1162; found: 370.1172.

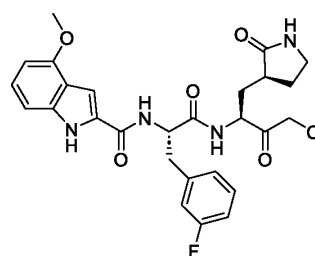
(S)-3-(3-fluorophenyl)-2-

5 **palmitamidopropanoic acid (i27):** 
Following GP4, using 2-chloro CTC resin (5.0 g, 1.2 mmol/g), **Fmoc-3-fluoro-L-phenylalanine** (3.0 equiv) was coupled and subsequently deprotected. **Palmitic acid** (3.0 equiv), PyBOP (3.0 equiv), DIPEA (5.0 equiv),
10 DMF (10 vol) was added to the resin and gently agitated under nitrogen bubbling for 2 h. After cleavage, **i27** was obtained as an off-white solid (2.0 g, 80%), which was used directly without further purification. 1H NMR (400 MHz, $CDCl_3$) δ 9.09 (bs, 1H), 7.26 (td, $J = 7.9, 5.9$ Hz, 1H), 7.01–6.86 (m, 3H), 6.45 (d, $J = 7.6$ Hz, 1H), 4.89 (dt, $J = 7.7, 5.9$ Hz, 1H), 3.19 (ddd, $J = 57.5, 14.0, 5.9$ Hz, 2H), 2.22 (td, $J = 7.4, 2.5$ Hz, 2H), 1.57 (p, $J = 7.2$ Hz, 2H), 1.40–1.19 (m, 24H), 0.90
15 (t, $J = 6.7$ Hz, 3H). ^{19}F NMR (377 MHz, $CDCl_3$) δ -112.90. ESI-MS (m/z): $[M+H]^+$ calcd for $C_{25}H_{40}FNO_3$: 422.31; found: 422.44.

(S)-3-(3-fluorophenyl)-2-octanamidopropanoic acid

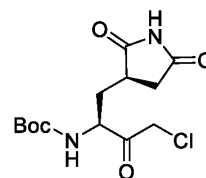
(**i28**): Following GP4, using 2-chloro CTC resin (5.0 g, 1.2 mmol/g), **Fmoc-3-fluoro-L-phenylalanine** (3.0 equiv) and **Octanoic acid** (3.0 equiv) were coupled. The dipeptide was
20 cleaved off the resin using 5% TFA in DCM as cleavage cocktail, yielding **i28** (1.0 g, 56%), which was used directly without further purification. 1H NMR (400 MHz, $CDCl_3$) δ 8.61 (bs, 1H), 7.28–7.18 (m, 1H), 7.00–6.81 (m, 3H), 6.22 (d, $J = 7.5$ Hz, 1H), 4.87 (dt, $J = 7.5, 5.9$ Hz, 1H), 3.17 (ddd, $J = 53.6, 14.0, 5.9$ Hz, 2H), 2.20 (td, $J = 7.4, 2.4$ Hz, 2H), 1.56 (p, $J = 7.1$ Hz, 2H), 1.33–1.16 (m, 8H), 0.90–0.83 (m, 3H). ^{19}F NMR (377 MHz, $CDCl_3$) δ -112.90. HRMS (m/z): $[M+H]^+$ calcd for $C_{17}H_{24}FNO_3$: 310.1805; found: 310.1813.

N-((S)-1-(((S)-4-chloro-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1H-indole-2-carboxamide (i29): **Tert-butyl ((S)-4-chloro-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)carbamate** (5.0 g, 16.4 mmol) was deprotected according to GP2. Subsequently, GP3 was followed using **(S)-3-(3-fluorophenyl)-2-(4-methoxy-1H-indole-2-**
30 **carboxamido)propanoic acid** (3.8 g, 10.67 mmol) and **((S)-3-((S)-2-amino-4-chloro-3-oxobutyl)pyrrolidin-2-one (TFA salt)** (5.12 g, 16.01 mmol). The crude material was purified by C18 reverse phase column chromatography (35→40% ACN in H_2O) to give **i29** as an off-white solid (3.4 g, 34%). 1H NMR (400 MHz, d -DMSO) δ 11.54 (d, $J = 2.38$ Hz, 1H), 8.70 (d, $J = 7.90$ Hz, 1H), 8.63 (d, $J = 8.01$ Hz, 1H), 7.63 (s, 1H), 7.31 (m, 2H), 7.22 (m, 2H), 7.09 (m, 1H), 6.99 (m,

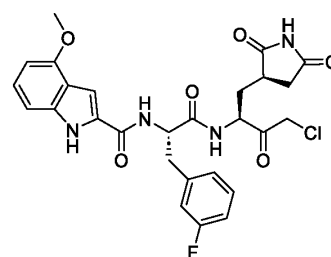


2H), 6.50 (d, $J = 7.75$ Hz, 1H), 4.71 (m, 1H), 4.50 (m, 2H), 4.46 (m, 1H), 3.89 (s, 3H), 3.11 (m, 4H), 2.29 (m, 1H), 2.02 (m, 2H), 1.64 (m, 2H). ESI-MS (m/z): calcd for $C_{27}H_{28}ClFN_4O_5$: 543.18; found: 543.21.

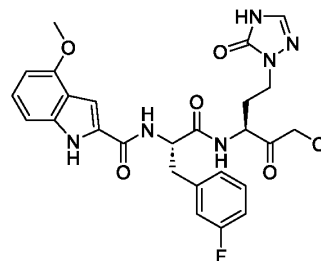
- 5 **tert-Butyl ((S)-4-chloro-1-((R)-2,5-dioxopyrrolidin-3-yl)-3-oxobutan-2-yl)carbamate (i30)**: A solution of of **tert-butyl ((S)-4-chloro-3-oxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)carbamate** (2.5 g, 8.22 mmol, 1.0 equiv) was dissolved in EtOAc/H₂O (1:1, 200 mL, 80 vol) and cooled to 0 °C. To this, NaIO₄ (26.2 g, 123.3 mmol, 15.0 equiv) and RuCl₃ (1.7 g, 8.22 mmol 1.0 equiv) were added. The resulting reaction mixture was stirred at 0-10 °C for 2 h. Progress of the reaction was monitored by TLC. Upon completion of the reaction, it was filtered through celite and extracted with EtOAc (3 x 100 mL). The combined organic layer was washed with brine (100 mL), dried over Na₂SO₄ and concentrated under vacuum to give crude material as brown gum. The obtained crude was purified by flash chromatography, eluting with 30-40% EtOAc in pet-ether to give **i30** (1.6 g, 61 %) as an off-white solid. $R_f = 0.6$ (EtOAc/heptane 6:4). ¹H NMR (400 MHz, CDCl₃) δ 8.16 (bs, 1H), 5.38 (d, $J = 8.5$ Hz, 1H), 4.70 (q, $J = 7.7$ Hz, 1H), 4.31 (d, $J = 3.7$ Hz, 2H), 3.10–3.00 (m, 1H), 2.99–2.91 (m, 1H), 2.55 (dd, $J = 18.0, 5.00$ Hz, 1H), 2.20–2.06 (m, 2H), 1.45 (s, 9H). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{13}H_{19}ClN_2O_5$: 319.11; found: 219.13 (-Boc).



- 20 **N-((S)-1-(((S)-4-chloro-1-((R)-2,5-dioxopyrrolidin-3-yl)-3-oxobutan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1H-indole-2-carboxamide (i31)**: **tert-Butyl ((S)-4-chloro-1-((R)-2,5-dioxopyrrolidin-3-yl)-3-oxobutan-2-yl)carbamate** (1.1 equiv, 1.1 g, 3.45 mmol) was deprotected following GP2. Subsequently, starting from **(R)-3-((S)-2-amino-4-chloro-3-oxobutyl)pyrrolidine-2,5-dione (TFA salt)** (1.0 g, 4.56 mmol) and **(S)-3-(3-fluorophenyl)-2-(4-methoxy-1H-indole-2-carboxamido)propanoic acid** (1.62 g, 4.56 mmol) GP3 was followed and the crude product was purified by flash column chromatography (8→10% MeOH in DCM) to yield **i31** (1.0 g, 64%) as an off-white solid. $R_f = 0.5$ (MeOH/DCM 1:10). ¹H NMR (400 MHz, *d*-DMSO) δ 11.53 (bs, 1H), 11.13 (bs, 1H), 8.67 (dd, $J = 19.6, 8.1$ Hz, 2H), 7.35–6.90 (m, 7H), 6.55–6.44 (m, 1H), 4.75–4.59 (m, 1H), 4.48 (s, 3H), 3.89 (s, 3H), 3.21–3.12 (m, 2H), 3.10–2.99 (m, 2H), 2.10–1.84 (m, 3H). ESI-MS (m/z): $[M+H]^+$ calcd for $C_{27}H_{26}ClFN_4O_6$: 557.16; found: 557.32.

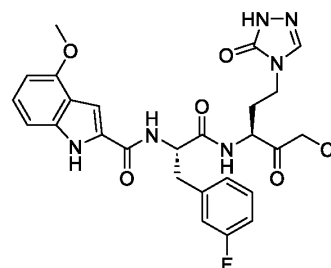


- 35 **N-((S)-1-(((S)-1-chloro-2-oxo-5-(5-oxo-4,5-dihydro-1H-1,2,4-triazol-1-yl)pentan-3-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1H-indole-2-carboxamide (i32)**: To a cooled solution (-5 °C) of **(S)-3-(3-fluorophenyl)-2-(4-methoxy-1H-indole-2-carboxamido)propanoic acid** (0.6 g,

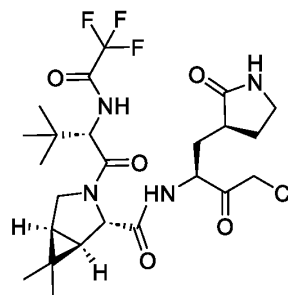


1.68 mmol) in DMF (18 mL, 30 vol), was added **(S)-4-(3-amino-5-chloro-4-oxopentyl)-2,4-dihydro-3H-1,2,4-triazol-3-one** (0.551 g, 2.52 mmol). The contents were stirred for 5 min at the same temperature followed by the addition of HATU (0.96 g, 2.52 mmol) and DIPEA (0.96 mL, 5.05 mmol). The resulting reaction mixture was stirred at -5 °C for 30 min. Progress of the reaction was monitored by TLC. After completion of reaction, it was diluted with water (24 mL, 40 vol) and extracted with EtOAc (3 x 50 mL). The combined organic layer was washed with ice cold water (3 x 30 mL), brine solution (50 mL) and dried over Na₂SO₄. The volatiles were removed under vacuum and the crude was purified by silica flash chromatography (2→3% MeOH in DCM) to give **i32** (33 mg, 4%) as an off-white solid. ¹H NMR (400 MHz, *d*-DMSO) δ 11.54 (d, *J* = 2.3 Hz, 1H), 8.75 (d, *J* = 7.7 Hz, 1H), 8.66 (d, *J* = 7.8 Hz, 1H), 7.82 (s, 1H), 7.34–7.26 (m, 2H), 7.25–7.18 (m, 2H), 7.08 (t, *J* = 8.0 Hz, 1H), 6.97 (d, *J* = 8.3 Hz, 2H), 6.50 (d, *J* = 7.7 Hz, 1H), 4.75–4.65 (m, 1H), 4.48 (d, *J* = 1.4 Hz, 1H), 4.46–4.36 (m, 1H), 3.88 (s, 3H), 3.66 (t, *J* = 7.2 Hz, 2H), 3.18 (dd, *J* = 13.8, 4.5 Hz, 1H), 3.08–2.98 (m, 1H), 2.24–2.14 (m, 1H), 1.95–1.85 (m, 1H). ¹⁹F NMR (377 MHz, *d*-DMSO) δ -113.8 (m). HRMS (*m/z*): [M+H]⁺ calcd for C₂₆H₂₆ClFN₆O₅: 557.1710; found: 557.1711.

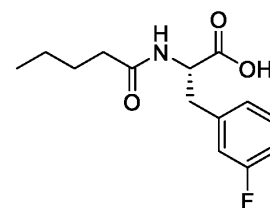
***N*-((*S*)-1-(((*S*)-1-chloro-2-oxo-5-(5-oxo-1,5-dihydro-4*H*-1,2,4-triazol-4-yl)pentan-3-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (**i33**):** A solution of **(S)-3-(3-fluorophenyl)-2-(4-methoxy-1*H*-indole-2-carboxamido)propanoic acid** (0.4 g, 1.12 mmol) in DMF (12 mL, 30 vol) was cooled to -5 °C, followed by the addition of **(S)-4-(3-amino-5-chloro-4-oxopentyl)-2,4-dihydro-3H-1,2,4-triazol-3-one 9** (0.491 g, 2.245 mmol). The contents were stirred for 5 min at the same temperature followed by the addition of HATU (0.64 g, 1.68 mmol) and DIPEA (0.58 mL, 3.36 mmol). The resulting reaction mixture was stirred at -5 °C for 30 min. Progress of the reaction was monitored by TLC. After completion of reaction, it was diluted with water (16 mL, 40 vol) and extracted with EtOAc (3 x 30 mL). The combined organic layer was washed with ice cold water (3 x 20 mL), brine solution (30 mL) and dried over Na₂SO₄. The volatiles were removed under vacuum to give crude material. The crude was purified by silica flash chromatography (2→3% MeOH in DCM) to give **i33** (0.20 g, 32%) as yellow solid. ¹H NMR (400 MHz, *d*-DMSO) δ 11.64 (bs, 1H), 11.53 (bs, 1H), 8.77 (d, *J* = 7.9 Hz, 1H), 8.70 (d, *J* = 7.6 Hz, 1H), 7.75 (d, *J* = 1.3 Hz, 1H), 7.35–7.28 (m, 2H), 7.27–7.18 (m, 2H), 7.09 (t, *J* = 8.0 Hz, 1H), 7.04–6.95 (m, 2H), 6.50 (d, *J* = 7.8 Hz, 1H), 4.72–4.62 (m, 1H), 4.45 (s, 1H), 4.40–4.32 (m, 1H), 3.89 (s, 3H), 3.63–3.48 (m, 2H), 3.18 (dd, *J* = 13.7, 5.0 Hz, 1H), 3.11–3.02 (m, 1H), 2.26–2.15 (m, 1H), 1.89–1.77 (m, 1H). ¹⁹F NMR (377 MHz, *d*-DMSO) δ -113.7 (m). HRMS (*m/z*): [M+H]⁺ calcd for C₂₆H₂₆ClFN₆O₅: 557.1710; found: 557.1706.



(1*R*,2*S*,5*S*)-*N*-((*S*)-4-chloro-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)butan-2-yl)-3-((*S*)-3,3-dimethyl-2-(2,2,2-trifluoroacetamido)butanoyl)-6,6-dimethyl-3-azabicyclo[3.1.0]hexane-2-carboxamide (**i34**): *tert*-Butyl ((*S*)-4-chloro-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)butan-2-yl)carbamate (2.0 g, 6.58 mmol) was deprotected according to GP2. Subsequently, starting from Nirmaltrevir-COOH (1.1 g, 3.14 mmol) and ((*S*)-3-((*S*)-2-amino-4-chloro-3-oxobutyl)pyrrolidin-2-one (TFA salt) (1.0 g, 3.14 mmol) GP3 was followed and the crude product was purified by flash column chromatography to yield **i34** (800 mg, 53%). ¹H NMR (400 MHz, *d*-DMSO) δ 9.39 (d, *J* = 8.5 Hz, 1H), 8.74 (d, *J* = 8.1 Hz, 1H), 7.60 (s, 1H), 4.63 (d, *J* = 0.9 Hz, 2H), 4.51–4.44 (m, 1H), 4.42 (d, *J* = 8.5 Hz, 1H), 4.23 (s, 1H), 3.71–3.67 (m, 2H), 3.15 (t, *J* = 9.0 Hz, 1H), 3.11–3.00 (m, 1H), 2.42–2.34 (m, 1H), 2.16–2.07 (m, 1H), 2.00–1.90 (m, 1H), 1.68–1.58 (m, 2H), 1.54 (dd, *J* = 7.6, 5.3 Hz, 1H), 1.38 (d, *J* = 7.6 Hz, 1H), 1.03 (s, 3H), 0.98 (s, 9H), 0.86 (s, 3H). ESI-MS (*m/z*): [M+H]⁺ calcd for C₂₄H₃₄ClF₃N₄O₅: 551.22; found: 551.22.

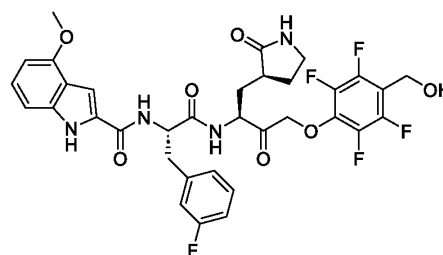


(*S*)-3-(3-fluorophenyl)-2-pentanamidopropanoic acid (**i35**): According to GP4 coupling using 2-chloro CTC resin (253 mg, 1.2 mmol/g) with Fmoc-3-fluoro-L phenylalanine (3.0 equiv) and pentanoic acid (2.0 equiv), dipeptide **i35** was obtained as a yellow oil (103 mg, 67%), which was used directly without further purification. ¹H NMR (400 MHz, CDCl₃) δ 8.69 (bs, 1H), 7.24 (td, *J* = 7.9, 6.0 Hz, 1H), 6.98–6.86 (m, 3H), 6.43 (d, *J* = 7.5 Hz, 1H), 4.88 (q, *J* = 6.4 Hz, 1H), 3.18 (ddd, *J* = 57.4, 10.4, 5.8 Hz, 2H), 2.25–2.18 (m, 2H), 1.55 (p, *J* = 7.5 Hz, 2H), 1.28 (h, *J* = 7.4 Hz, 2H), 0.89 (t, *J* = 7.3 Hz, 3H). ¹⁹F NMR (377 MHz, CDCl₃) δ -113.1 (td, *J* = 9.2, 6.1 Hz). SFC-MS (*m/z*): [M+H]⁺ calcd for C₁₄H₁₈FNO₃: 268.1; found: 268.2.



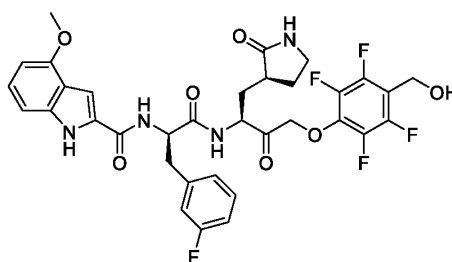
1.3 Synthesis of inhibitors

N-((*S*)-3-(3-fluorophenyl)-1-oxo-1-((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (**1**): *tert*-Butyl ((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)carbamate (1.1 equiv, 287 mg, 0.617 mmol) was deprotected following GP2. Subsequently, GP3 was followed using (*S*)-3-(3-fluorophenyl)-2-(4-methoxy-1*H*-indole-2-carboxamido)propanoic acid (200 mg, 0.561 mmol) and (*S*)-3-((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (295 mg, 0.617 mmol) and the crude product was purified by RP-HPLC to yield **1** (33 mg, 8%). ¹H NMR (400 MHz, *d*-DMSO) δ



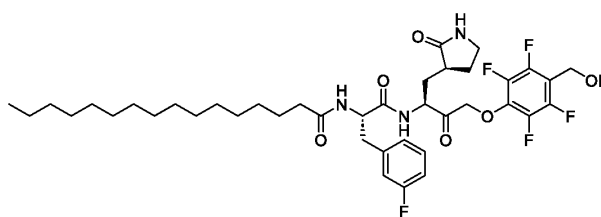
11.52 (d, $J = 2.3$ Hz, 1H), 8.67 (dd, $J = 17.6, 8.1$ Hz, 2H), 7.63 (s, 1H), 7.35–7.16 (m, 4H), 7.08 (t, $J = 8.0$ Hz, 1H), 7.01–6.91 (m, 2H), 6.50 (d, $J = 7.7$ Hz, 1H), 5.12 (ABq, $J = 48.8$ Hz, 2H), 4.76–4.66 (m, 1H), 4.50 (s, 2H), 4.48–4.43 (m, 1H), 3.89 (s, 3H), 3.18–2.99 (m, 4H), 2.31–2.21 (m, 1H), 2.12–1.93 (m, 2H), 1.69–1.55 (m, 2H), 0.89–0.75 (m, 1H). ^{19}F NMR (377 MHz, d -DMSO) δ -115.1 (m), -147.5 (dd, $J = 23.1, 8.7$ Hz), -158.8 (dd, $J = 23.4, 8.8$ Hz). HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{34}\text{H}_{31}\text{F}_5\text{N}_4\text{O}_7$: 703.2186; found: 703.2169.

***N*-((*R*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-3-oxo-1-(((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (2):** *tert*-Butyl ((*S*)-3-oxo-1-(((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)carbamate



(1.1 equiv, 287 mg, 0.617 mmol) was deprotected following GP2. Subsequently, GP3 was followed using (*R*)-3-(3-fluorophenyl)-2-(4-methoxy-1*H*-indole-2-carboxamido)propanoic acid (200 mg, 0.561 mmol) and (*S*)-3-(((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (295 mg, 0.617 mmol) and the crude product was purified by flash column chromatography and SFC to yield **2** (40 mg, 10%). ^1H NMR (400 MHz, d -DMSO) δ 11.49 (d, $J = 2.3$ Hz, 1H), 8.69 (dd, $J = 34.4, 7.8$ Hz, 2H), 7.63 (s, 1H), 7.35–7.26 (m, 2H), 7.24–7.16 (m, 2H), 7.09 (t, $J = 8.0$ Hz, 1H), 7.02–6.93 (m, 2H), 6.54–6.46 (m, 1H), 5.44 (t, $J = 5.8$ Hz, 1H), 5.27 (q, $J = 18.0$ Hz, 2H), 4.72 (q, $J = 7.8$ Hz, 1H), 4.47 (d, $J = 5.8$ Hz, 2H), 4.44–4.35 (m, 1H), 3.89 (s, 3H), 3.19–3.00 (m, 4H), 2.06–1.89 (m, 3H), 1.65–1.50 (m, 2H), 1.27–1.20 (m, 1H). ^{19}F NMR (377 MHz, d -DMSO) δ -115.61 (m), -148.13 (dd, $J = 23.7, 8.7$ Hz), -159.43 (dd, $J = 23.6, 8.6$ Hz). HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{34}\text{H}_{31}\text{F}_5\text{N}_4\text{O}_7$: 703.2186; found: 703.2192.

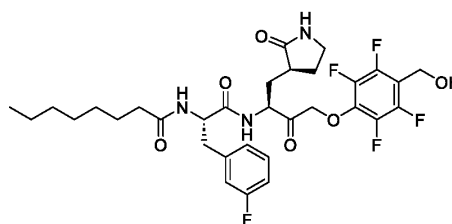
***N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-3-oxo-1-(((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)propan-2-yl)palmitamide**



(3): *tert*-Butyl ((*S*)-3-oxo-1-(((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)carbamate (1.1 equiv, 364 mg, 0.783 mmol) was deprotected following GP2. Subsequently, according to GP3 starting from (*S*)-3-(3-fluorophenyl)-2-palmitamidopropanoic acid (300 mg, 0.711 mmol) and (*S*)-3-(((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (374 mg, 0.782 mmol) the crude product was purified by flash column chromatography and SFC to yield **3** (74 mg, 14%). ^1H NMR (400 MHz, d -DMSO) δ 8.57 (d, $J = 7.8$ Hz, 1H), 8.14 (d, $J = 8.0$ Hz, 1H), 7.63 (bs,

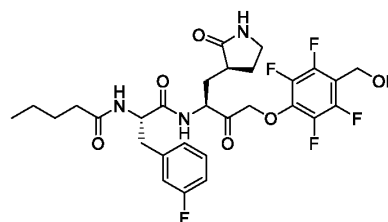
1H), 7.27 (td, $J = 8.0, 6.2$ Hz, 1H), 7.13–7.06 (m, 2H), 7.00–6.91 (m, 1H), 5.48 (t, $J = 5.8$ Hz, 1H), 5.03 (ABq, $J = 48.6$ Hz, 2H), 4.55–4.46 (m, 3H), 4.39 (ddd, $J = 11.7, 8.0, 4.0$ Hz, 1H), 3.19–3.02 (m, 2H), 3.00 (dd, $J = 13.8, 5.4$ Hz, 1H), 2.80 (dd, $J = 13.7, 9.7$ Hz, 1H), 2.25–2.14 (m, 1H), 2.12–2.04 (m, 1H), 2.03 (t, $J = 7.3$ Hz, 2H), 1.99–1.89 (m, 1H), 1.66–1.56 (m, 2H), 1.36 (p, $J = 7.3$ Hz, 2H), 1.29–1.06 (m, 24H), 0.89–0.81 (m, 3H). ^{19}F NMR (377 MHz, d -DMSO) δ -113.98 (m), -146.35 (dd, $J = 23.5, 8.9$ Hz), -157.56 (dd, $J = 23.4, 8.9$ Hz). HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{40}\text{H}_{54}\text{F}_5\text{N}_3\text{O}_6$: 768.4006; found: 768.3990.

***N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)propan-2-yl)octanamide (16): *tert*-Butyl ((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-**



yl)carbamate (1.1 equiv, 0.7 g, 1.508 mmol) was deprotected following GP2. Subsequently, according to GP3 starting from (*S*)-3-(3-fluorophenyl)-2-octanamidopropanoic acid (452 mg, 1.463 mmol) and (*S*)-3-((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (0.7 g, 1.463 mmol) the crude product was purified by RP-HPLC and SFC to yield **16** (80 mg, 8%). ^1H NMR (400 MHz, d -DMSO) δ 8.57 (d, $J = 8.0$ Hz, 1H), 8.14 (d, $J = 8.0$ Hz, 1H), 7.63 (bs, 1H), 7.32–7.23 (m, 1H), 7.13–7.06 (m, 2H), 7.01–6.92 (m, 1H), 5.48 (t, $J = 5.8$ Hz, 1H), 5.03 (ABq, $J = 48.9$ Hz, 2H), 4.54–4.46 (m, 2H), 4.44–4.35 (m, 1H), 3.19–2.97 (m, 4H), 2.80 (dd, $J = 13.7, 9.6$ Hz, 1H), 2.24–2.14 (m, 1H), 2.03 (t, $J = 7.3$ Hz, 2H), 1.98–1.88 (m, 1H), 1.66–1.56 (m, 2H), 1.36 (p, $J = 7.4$ Hz, 2H), 1.29–1.20 (m, 2H), 1.19–1.13 (m, 3H), 1.12–1.04 (m, 2H), 0.94 (bs, 1H), 0.84 (t, $J = 7.0$ Hz, 3H). HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{32}\text{H}_{38}\text{F}_5\text{N}_3\text{O}_6$: 656.2754; found: 656.2710.

***N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)propan-2-yl)pentanamide (20): *tert*-Butyl ((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-**



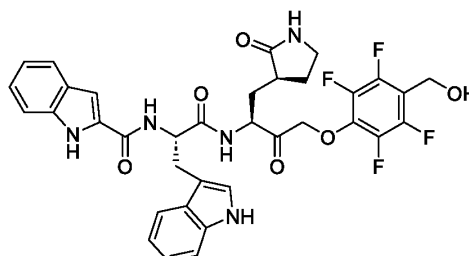
(hydroxymethyl)phenoxy)butan-2-yl)carbamate (75 mg, 0.161 mmol) was deprotected following GP2. Subsequently, GP3 was followed using (*S*)-3-(3-fluorophenyl)-2-pentanamidopropanoic acid (**i35**, 43 mg, 0.139 mmol) and (*S*)-3-((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (77 mg, 0.161 mmol) and the crude product was purified by flash column chromatography and RP-HPLC to yield **20** (5.5 mg, 6%). ^1H NMR (500 MHz, d -DMSO) δ 8.60 (d, $J = 8.0$ Hz, 1H), 7.64 (s, 1H), 7.27 (q, $J = 7.5$ Hz, 1H), 7.12–7.05 (m, 2H), 6.96 (td, $J = 8.7, 2.5$ Hz, 1H), 5.50 (bs, 1H), 5.12–4.92 (m, 2H), 4.51–4.46 (m, 1H), 4.40 (ddd, $J = 11.8, 7.9, 4.0$ Hz, 1H), 3.15 (t, $J = 9.3$ Hz, 1H), 3.07 (q, $J = 8.6, 1\text{H}$), 3.00 (dd, $J = 13.8, 5.5$ Hz, 1H), 2.80 (dd, $J = 13.8, 9.7$ Hz, 1H), 2.19 (qd, $J =$

79

10.1, 3.9 Hz, 1H), 2.10–2.01 (m, 3H), 1.94 (ddd, $J = 14.8, 11.5, 4.3$ Hz, 1H), 1.65–1.55 (m, 2H), 1.35 (p, $J = 7.4$ Hz, 2H), 1.11 (h, $J = 7.4$ Hz, 2H), 0.78 (t, $J = 7.3$ Hz, 3H). ^{19}F NMR (471 MHz, d -DMSO) δ -114.0 (td, $J = 9.7, 6.3$ Hz), -146.3 (dd, $J = 23.2, 8.8$ Hz), -157.5 (dd, $J = 23.3, 8.9$ Hz). HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{29}\text{H}_{32}\text{F}_5\text{N}_3\text{O}_6$: 614.2284; found: 614.2282.

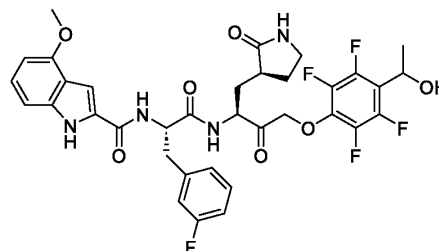
5

***N*-((*S*)-3-(1*H*-indol-3-yl)-1-oxo-1-(((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)propan-2-yl)-1*H*-indole-2-carboxamide (4):** *tert*-Butyl ((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-



yl)carbamate (1.8 g, 3.876 mmol) was deprotected following GP2. Subsequently, GP3 was followed using (1*H*-indole-2-carboxyl)-*L*-tryptophan (700 mg, 2.015 mmol) and (*S*)-3-((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (1.060 g, 2.217 mmol) and the crude product was purified by flash column chromatography and RP-HPLC to yield **4** (24 mg, 2%). ^1H NMR (400 MHz, d -DMSO) δ 11.53 (d, $J = 2.2$ Hz, 1H), 10.81 (d, $J = 2.5$ Hz, 1H), 8.74 (d, $J = 8.1$ Hz, 1H), 8.59 (d, $J = 7.7$ Hz, 1H), 7.71 (d, $J = 7.6$ Hz, 1H), 7.64–7.57 (m, 2H), 7.39 (d, $J = 8.2$ Hz, 1H), 7.32–7.19 (m, 3H), 7.16 (t, $J = 7.6$ Hz, 1H), 7.10–6.94 (m, 4H), 5.05 (ABq, $J = 62.4$ Hz, 2H), 4.80–4.67 (m, 1H), 4.49 (s, 2H), 4.48–4.43 (m, 1H), 3.31–3.23 (m, 1H), 3.20–2.98 (m, 3H), 2.34–2.23 (m, 1H), 2.10–1.94 (m, 2H), 1.69–1.57 (m, 2H). ^{19}F NMR (377 MHz, d -DMSO) δ -147.8 (dd, $J = 23.5, 8.8$ Hz), -159.13 (dd, $J = 23.2, 8.6$ Hz). HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{35}\text{H}_{31}\text{F}_4\text{N}_5\text{O}_6$: 694.2283; found: 694.2298.

***N*-((2*S*)-3-(3-fluorophenyl)-1-oxo-1-(((2*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenoxy)butan-2-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (5):** *tert*-Butyl ((2*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenoxy)butan-2-yl)carbamate (600

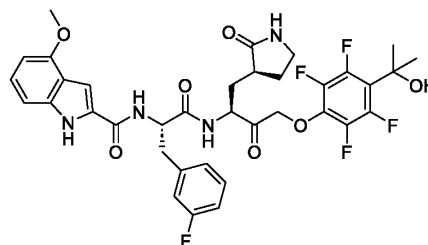


mg, 1.25 mmol) was deprotected according to GP2. Subsequently, GP3 was followed using (*S*)-3-(3-fluorophenyl)-2-(4-methoxy-1*H*-indole-2-carboxamido)propanoic acid (450 mg, 1.26 mmol) and (*S*)-3-((2*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt) (620 mg, 1.26 mmol) and the crude product was purified by flash column chromatography and RP-HPLC to yield **5** (155 mg, 17%) as off-white solid. ^1H NMR (400 MHz, d -DMSO) δ 11.53 (d, $J = 2.3$ Hz, 1H), 8.67 (dd, $J = 15.0, 8.1$ Hz, 2H), 7.63 (s, 1H), 7.34–7.18 (m, 4H), 7.08 (t, $J = 7.9$ Hz, 1H), 7.00–6.92 (m, 2H), 6.50 (d, $J = 7.7$ Hz, 1H), 5.60 (d, $J = 4.4$ Hz, 1H), 5.21–5.01 (m, 3H), 4.75–4.67 (m, 1H), 4.48 (ddd, $J = 11.7, 8.0, 3.9$ Hz, 1H), 3.89 (s, 3H), 3.17–2.99 (m, 4H), 2.31–2.24 (m, 1H), 2.11–1.94 (m, 2H), 1.69–1.56 (m, 2H), 1.46 (d, $J = 6.7$ Hz, 3H). ^{19}F (377 MHz, d -DMSO): δ -113.85 (m), -145.82 (d, $J =$

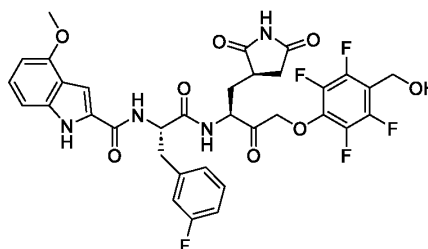
40

22.9, 8.2 Hz), -157.66 (dd, $J = 22.6, 8.4$ Hz). HRMS (m/z): $[M+H]^+$ calcd for $C_{35}H_{33}F_5N_4O_7$: 717.2342; found: 717.2324.

***N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenoxy)butan-2-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (6):** Following GP1, **6** was obtained using ***N*-((*S*)-1-(((*S*)-4-chloro-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide** (3.4 g, 6.26 mmol) and **2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenol** (1.4 g, 6.26 mmol). The crude product was purified by flash column chromatography and RP-HPLC to yield **6** (350 mg, 8%) as off-white solid. 1H NMR (400 MHz, d -DMSO) δ 11.54–11.49 (m, 1H), 8.75–8.62 (m, 2H), 7.63 (bs, 1H), 7.34–7.17 (m, 4H), 7.08 (t, $J = 7.97$ Hz, 1H), 7.03–6.91 (m, 2H), 6.50 (d, $J = 7.71$ Hz, 1H), 5.21–5.01 (ABq, 2H), 4.77–4.66 (m, 1H), 4.51–4.44 (m, 1H), 3.89 (s, 3H), 3.19–2.99 (m, 4H), 2.32–2.24 (m, 1H), 2.12–2.03 (m, 1H), 2.03–1.94 (m, 1H), 1.70–1.60 (m, 2H), 1.59–1.53 (m, 6H). ^{19}F (377 MHz, d -DMSO): δ -113.85 (m), -141.1 (d, $J = 22.5, 7.0$ Hz), -158.0 (d, $J = 22.2, 7.0$ Hz). HRMS (m/z): calcd for $C_{36}H_{35}F_5N_4O_7$: 731.2499; found: 731.2474.

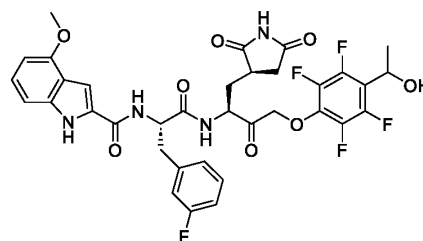


***N*-((*S*)-1-(((*S*)-1-((*R*)-2,5-dioxopyrrolidin-3-yl)-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (7):** Following GP1, **7** was obtained using ***N*-((*S*)-1-(((*S*)-4-chloro-1-((*R*)-2,5-dioxopyrrolidin-3-yl)-3-oxobutan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide** (300 mg, 0.539 mmol) and **2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenol** (105 mg, 0.539 mmol). The crude product was purified by flash column chromatography and RP-HPLC to yield **7** (13 mg, 4%) as an off-white solid. 1H NMR (400 MHz, d -DMSO) δ 11.52–11.49 (m, 1H), 11.13 (s, 1H), 8.71–8.61 (m, 2H), 7.33–7.18 (m, 4H), 7.08 (t, $J = 8.0$ Hz, 1H), 7.00–6.92 (m, 2H), 6.50 (d, $J = 7.7$ Hz, 1H), 5.20–5.01 (m, 3H), 4.73–4.65 (m, 1H), 4.53–4.44 (m, 3H), 3.89 (s, 3H), 3.15 (dd, $J = 14.0, 4.6$ Hz, 1H), 3.09–3.01 (m, 1H), 2.83–2.73 (m, 1H), 2.44–2.39 (m, 1H), 2.08–1.99 (m, 1H), 1.97–1.89 (m, 1H). ^{19}F (377 MHz, d -DMSO): δ -113.8 (m), -146.3 (dd, $J = 23.1, 8.5$ Hz), -157.4 (dd, $J = 23.0, 8.8$ Hz). HRMS (m/z): $[M+H]^+$ calcd for $C_{34}H_{29}F_5N_4O_8$: 717.1978; found: 717.1962.

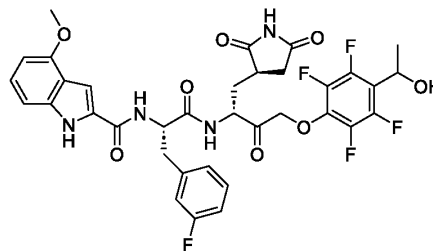


81

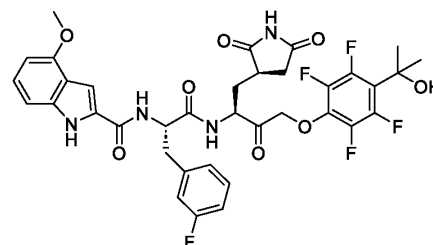
***N*-((2*S*)-1-(((2*S*)-1-((*R*)-2,5-dioxopyrrolidin-3-yl)-3-oxo-4-(2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenoxy)butan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (8):** Following GP1, **8** was obtained using ***N*-((*S*)-1-(((*S*)-4-chloro-1-((*R*)-2,5-dioxopyrrolidin-3-yl)-3-oxobutan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide** (300 mg, 0.539 mmol) and **2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenol** (113 mg, 0.539 mmol). The crude product was purified by flash column chromatography, RP-HPLC and chiral SFC to yield **8** (36 mg, 4%) as an off-white solid. ¹H NMR (400 MHz, *d*-DMSO) δ 11.51 (d, *J* = 2.4 Hz, 1H), 11.13 (bs, 1H), 8.66 (t, *J* = 7.7 Hz, 2H), 7.34–7.18 (m, 4H), 7.08 (t, *J* = 8.0 Hz, 1H), 7.00–6.91 (m, 2H), 6.53–6.46 (m, 1H), 5.60 (d, *J* = 4.4 Hz, 1H), 5.19–5.01 (m, 3H), 4.74–4.64 (m, 1H), 4.53–4.45 (m, 1H), 3.89 (s, 3H), 3.15 (dd, *J* = 13.8, 4.8 Hz, 1H), 3.09–3.01 (m, 1H), 2.85–2.75 (m, 1H), 2.46–2.40 (m, 1H), 2.02–1.98 (m, 1H), 1.97–1.86 (m, 1H), 1.45 (d, *J* = 6.7 Hz, 3H). ¹⁹F (377 MHz, *d*-DMSO): δ -113.8 (q, *J* = 9.2 Hz), -145.8 (dd, *J* = 22.9, 8.3 Hz), -157.5 (dd, *J* = 23.1, 8.2 Hz). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₅H₃₁F₅N₄O₈: 731.2135; found: 731.2112.



9: During the synthesis and purification of ***N*-((2*S*)-1-(((2*S*)-1-((*R*)-2,5-dioxopyrrolidin-3-yl)-3-oxo-4-(2,3,5,6-tetrafluoro-4-(1-hydroxyethyl)phenoxy)butan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide**, **9** was isolated as minor isomer (7.5 mg, 1%). ¹H NMR (400 MHz, *d*-DMSO) δ 11.50–11.47 (m, 1H), 8.68 (dd, *J* = 23.3, 7.8 Hz, 2H), 7.33–7.17 (m, 4H), 7.08 (t, *J* = 7.9 Hz, 1H), 7.00–6.94 (m, 2H), 6.53–6.48 (m, 1H), 5.58 (d, *J* = 4.5 Hz, 1H), 5.22 (ABq, *J* = 34.6 Hz, 2H), 5.08–4.99 (m, 1H), 4.72–4.65 (m, 1H), 4.64–4.57 (m, 1H), 3.89 (s, 3H), 3.15–3.00 (m, 2H), 2.40–2.35 (m, 1H), 2.30–2.20 (m, 1H), 1.58–1.50 (m, 2H), 1.45 (d, *J* = 6.7 Hz, 3H). ¹⁹F (377 MHz, *d*-DMSO): δ -113.8 (m), -145.8 (dd, *J* = 22.2, 7.8 Hz), -157.6 (dd, *J* = 22.7, 8.0 Hz). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₅H₃₁F₅N₄O₈: 731.2135; found: 731.2143.



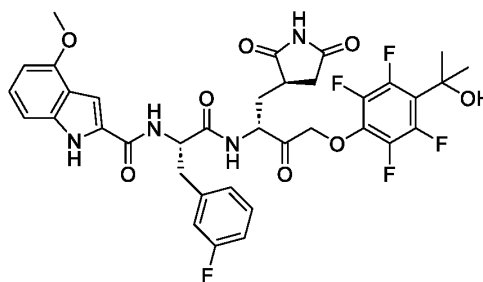
***N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenoxy)butan-2-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (10):** Following GP1, **10** was obtained using ***N*-((*S*)-1-(((*S*)-4-chloro-1-((*R*)-2,5-dioxopyrrolidin-3-yl)-3-oxobutan-2-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide** (300



mg, 0.539 mmol) and **2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenol** (120 mg, 0.539 mmol). The crude product was purified by flash column chromatography, RP-HPLC and chiral SFC to yield **10** (14 mg, 3%) as an off-white solid. ¹H NMR (400 MHz, *d*-DMSO) δ 11.52 (d, *J* = 2.3 Hz, 1H), 11.11 (bs, 1H), 8.66 (t, *J* = 7.5 Hz, 2H), 7.34–7.17 (m, 4H), 7.08 (d, *J* = 7.9 Hz, 1H), 7.00–6.90 (m, 2H), 6.50 (d, *J* = 7.7 Hz, 1H), 5.51 (s, 1H), 5.10 (ABq, *J* = 39.6 Hz, 2H), 4.74–4.64 (m, 1H), 4.54–4.44 (m, 1H), 3.89 (s, 3H), 3.15 (dd, *J* = 13.8, 4.8 Hz, 1H), 3.09–3.00 (m, 1H), 2.85–2.74 (m, 1H), 2.10–1.98 (m, 1H), 1.97–1.87 (m, 1H), 1.56 (bs, 4H), 1.35 (s, 3H). ¹⁹F (377 MHz, *d*-DMSO): δ -113.8 (d, *J* = 7.3 Hz), -141.1 (d, *J* = 22.7 Hz), -157.8 (dd, *J* = 21.9, 7.1 Hz). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₆H₃₃F₅N₄O₈: 745.2291; found: 745.2273.

10

11: During the synthesis and purification of *N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenoxy)butan-2-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide, **11** was isolated as minor isomer (4 mg, 1%). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₆H₃₃F₅N₄O₈: 745.2291; found: 745.2303.



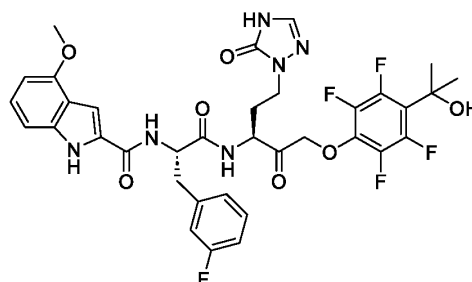
20

N-((*S*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-2-oxo-5-(5-oxo-1,5-dihydro-4*H*-1,2,4-triazol-4-yl)-1-(2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenoxy)pentan-3-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (**12**): Following GP1, **12** was obtained using *N*-((*S*)-1-(((*S*)-1-chloro-2-oxo-5-(5-oxo-4,5-dihydro-1*H*-1,2,4-triazol-1-yl)pentan-3-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (450 mg, 0.809 mmol)

and **2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenol** (109 mg, 0.485 mmol). The crude product was triturated with diethyl ether, purified by RP-HPLC and chiral SFC to yield **12** (42 mg, 12%) as off-white solid. *R*_f = 0.6 (EtOAc/pet-ether 3:2). ¹H NMR (400 MHz, *d*-DMSO) δ 11.52 (d, *J* = 2.3 Hz, 1H), 8.70 (dd, *J* = 18.0, 7.9 Hz, 2H), 7.82 (s, 1H), 7.34–7.24 (m, 2H), 7.24–7.18 (m, 2H), 7.08 (t, *J* = 8.0 Hz, 1H), 6.99–6.92 (m, 2H), 6.50 (d, *J* = 7.7 Hz, 1H), 5.51 (s, 1H), 5.10 (ABq, *J* = 37.9 Hz, 2H), 4.76–4.68 (m, 1H), 4.44–4.36 (m, 1H), 3.88 (s, 3H), 3.67 (t, *J* = 7.3 Hz, 2H), 3.15 (dd, *J* = 13.6, 4.5 Hz, 1H), 3.07–2.99 (m, 1H), 2.23–2.13 (m, 1H), 1.94–1.82 (m, 1H), 1.56 (t, *J* = 2.0 Hz, 6H). ¹⁹F (377 MHz, *d*-DMSO): δ -113.8 (m), -141.1 (d, *J* = 20.8 Hz), -157.9 (dd, *J* = 22.2, 7.6 Hz). HRMS (*m/z*): [*M*-*H*]⁻ calcd for C₃₅H₃₃F₅N₆O₇: 743.2258; found: 743.2272.

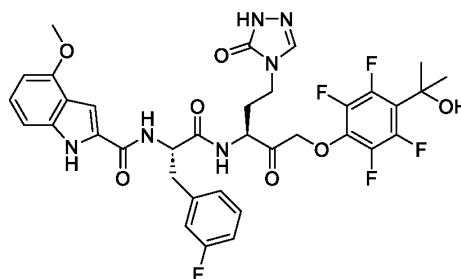
30

35

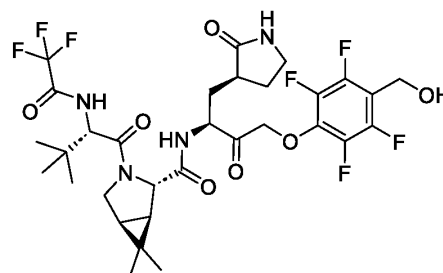


83

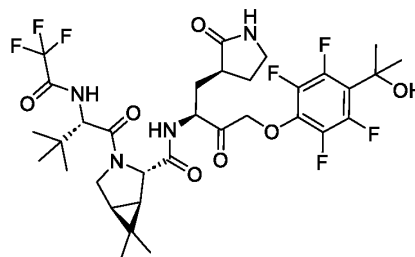
- N*-((*S*)-3-(3-fluorophenyl)-1-oxo-1-(((*S*)-2-oxo-5-(5-oxo-1,5-dihydro-4*H*-1,2,4-triazol-4-yl)-1-(2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenoxy)pentan-3-yl)amino)propan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide (13):**
- Following GP1, **13** was obtained using ***N*-((*S*)-1-(((*S*)-1-chloro-2-oxo-5-(5-oxo-1,5-dihydro-4*H*-1,2,4-triazol-4-yl)pentan-3-yl)amino)-3-(3-fluorophenyl)-1-oxopropan-2-yl)-4-methoxy-1*H*-indole-2-carboxamide** (290 mg, 0.521 mmol) and **2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenol** (234 mg, 1.043 mmol). The crude product was purified by RP-HPLC and chiral SFC to yield **13** (8 mg, 2%). ¹H NMR (400 MHz, *d*-DMSO) δ 11.64 (bs, 1H), 11.52 (d, *J* = 2.4 Hz, 1H), 8.74 (dd, *J* = 12.8, 7.8 Hz, 2H), 7.73 (d, *J* = 1.4 Hz, 1H), 7.35–7.18 (m, 4H), 7.12–7.06 (m, 1H), 7.00–6.93 (m, 2H), 6.50 (d, *J* = 7.7 Hz, 1H), 5.50 (bs, 1H), 5.06 (q, *J* = 17.9 Hz, 2H), 4.72–4.64 (m, 1H), 4.41–4.32 (m, 1H), 3.89 (s, 3H), 3.65–3.48 (m, 2H), 3.16 (dd, *J* = 13.8, 5.0 Hz, 1H), 3.12–3.03 (m, 1H), 2.25–2.14 (m, 1H), 1.88–1.77 (m, 1H), 1.56 (s, 6H). ¹⁹F (377 MHz, *d*-DMSO): δ -113.8 (d, *J* = 9.6 Hz), -141.1 (d, *J* = 21.9 Hz), 157.8 (dd, *J* = 22.0, 7.1 Hz). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₅H₃₃F₅N₆O₇: 745.2404; found: 745.2391.



- (1*R*,2*S*,5*S*)-3-((*S*)-3,3-dimethyl-2-(2,2,2-trifluoroacetamido)butanoyl)-6,6-dimethyl-*N*-((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)-3-azabicyclo[3.1.0]hexane-2-carboxamide (14):** *tert*-Butyl ((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butan-2-yl)carbamate (3.0 g, 6.46 mmol) was deprotected following GP2. Subsequently, starting from **(*S*)-3-((*S*)-2-amino-3-oxo-4-(2,3,5,6-tetrafluoro-4-(hydroxymethyl)phenoxy)butyl)pyrrolidin-2-one (TFA salt)** (1.3 g, 2.64 mmol) and **(1*R*,2*S*,5*S*)-3-((*S*)-3,3-dimethyl-2-(2,2,2-trifluoroacetamido)butanoyl)-6,6-dimethyl-3-azabicyclo[3.1.0]hexane-2-carboxylic acid** (0.800 g, 2.20 mmol) GP3 was followed and the crude product was purified by flash column chromatography and RP-HPLC to yield **14** (300 mg, 20%). ¹H NMR (400 MHz, *d*-DMSO) δ 9.40 (bs, 1H), 8.69 (dd, *J* = 8.4 Hz, 1H), 7.61 (s, 1H), 5.48 (t, *J* = 5.8 Hz, 1H), 5.26 (bs, 2H), 4.54–4.45 (m, 3H), 4.42 (bs, 1H), 4.23 (s, 1H), 3.91 (dd, *J* = 10.3, 5.5 Hz, 1H), 3.69 (d, *J* = 10.4 Hz, 1H), 3.14 (t, *J* = 9.2 Hz, 1H), 3.04 (q, *J* = 9.1 Hz, 1H), 2.43–2.34 (m, 1H), 2.15–2.05 (m, 1H), 2.01–1.90 (m, 1H), 1.69–1.57 (m, 2H), 1.55 (dd, *J* = 7.7, 5.3 Hz, 1H), 1.35 (d, *J* = 7.6 Hz, 1H), 1.02 (s, 3H), 0.97 (s, 9H), 0.85 (s, 3H). ¹⁹F (377 MHz, *d*-DMSO): δ -72.93 (s), -146.34 (dd, *J* = 23.2, 8.7 Hz), -157.71 (dd, *J* = 23.3, 8.7 Hz). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₁H₃₇F₇N₄O₇: 711.2623; found: 711.2610.



- (1*R*,2*S*,5*S*)-3-((*S*)-3,3-dimethyl-2-(2,2,2-trifluoroacetamido)butanoyl)-6,6-dimethyl-*N*-((*S*)-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)-4-(2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenoxy)butan-2-yl)-3-azabicyclo[3.1.0]hexane-2-carboxamide (**15**): Following GP1, **15** was obtained using (1*R*,2*S*,5*S*)-*N*-((*S*)-4-chloro-3-oxo-1-((*S*)-2-oxopyrrolidin-3-yl)butan-2-yl)-3-((*S*)-3,3-dimethyl-2-(2,2,2-trifluoroacetamido)butanoyl)-6,6-dimethyl-3-azabicyclo[3.1.0]hexane-2-carboxamide (600 mg, 1.09 mmol) and 2,3,5,6-tetrafluoro-4-(2-hydroxypropan-2-yl)phenol (244 mg, 1.09 mmol). The crude product was purified by flash column chromatography, RP-HPLC and chiral SFC to yield **15** (80 mg, 10%). ¹H NMR (400 MHz, *d*-DMSO) δ 9.42 (bs, 1H), 8.69 (d, *J* = 8.4 Hz, 1H), 7.61 (s, 1H), 5.51 (s, 1H), 5.24 (bs, 2H), 4.53–4.46 (m, 1H), 4.42 (d, *J* = 5.8 Hz, 1H), 4.23 (s, 1H), 3.91 (dd, *J* = 10.4, 5.4 Hz, 1H), 3.69 (d, *J* = 10.4 Hz, 1H), 3.14 (t, *J* = 9.0 Hz, 1H), 3.04 (q, *J* = 8.6 Hz, 1H), 2.44–2.32 (m, 1H), 2.15–2.08 (m, 1H), 2.00–1.91 (m, 1H), 1.68–1.52 (m, 8H), 1.37–1.32 (m, 2H), 1.03 (s, 3H), 0.97 (s, 9H), 0.85 (s, 3H). ¹⁹F (377 MHz, *d*-DMSO): δ -72.93 (s), -141.10 (d, *J* = 19.3 Hz), -158.17 (dd, *J* = 22.1, 7.0 Hz). HRMS (*m/z*): [*M*+*H*]⁺ calcd for C₃₃H₄₁F₇N₄O₇: 739.2936; found: 739.2930.



Example 2 - properties of inhibitors

20

Some reference compounds were tested, being the following:

17	
18	
19	

2.1.1 SARS-CoV-2 M^{Pro} biochemical assay

M^{Pro} enzymatic reactions were performed in reaction buffer containing 20 mM TRIS HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM TCEP, 0.25 % glycerol, 0.02 % BSA and 0.003 % Tween-20. Recombinant SARS-CoV-2 3C-like main protease according to PDB-ID 7BQY (6.6 nM final assay concentration) was added to Greiner small volume 384-well black plates containing 30 μ M Dabcyl-KTSAVLQSGFRKM-E(Edans)-amide substrate solution (Biosynthan GmbH, dissolved at 20 mM in 100% DMSO). 100 μ M ML188 was included to generate no activity control samples. 11-point full concentration response curves were generated directly in assay plates using a Tecan D300, providing concentrations between 0.0063 nM and 250 μ M. DMSO was normalized to 2.5% across the entire plate. Plates were immediately transferred to a BMG CLARIOstar plate reader and the change in fluorescence intensity was monitored for 30 min at 15 s intervals.

2.1.2 Cathepsin biochemical assays

Cathepsin enzymatic reactions were conducted in assay buffer with a final DMSO concentration of 1% per well. For hCTSL, assay buffer was prepared using 50 mM MES, 5 mM DTT, 1 mM EDTA, 0.005% (w/v) Brij-35, pH 6.0, mCTSL assays were conducted in assay buffer containing 25 mM MES, 5 mM DTT, pH 6.0, and hCTSB and mCTSB assays were performed in assay buffer with 25 mM MES, pH 5.0. Compounds (10 mM stocks in 100 % DMSO) and controls (100% DMSO) were prediluted and 10 μ L was added in duplicate to a Thermo Scientific™ Nunc™ F96 MicroWell™ black 96-well F-bottom PS micro plate (assay plate). Dose response curves consisted of 8 datapoints prepared using 3-fold serial dilutions. DMSO concentration was kept constant at 10 %. Highest assay plate final well compound concentration was set to 50 μ M. Ten μ M E-64 (10 mM stock in 100% DMSO (Alfa Aesar™)) final assay plate well concentration was used to generate no fluorescence signal increase and an 8-point dose response curve (0.457 – 1000 nM) was used as an internal compound control. Substrate pre-dilution to 37.5 μ M (10 mM stock Z-Leu-Arg-AMC x HCl (Bachem Biochemica) in 100% DMSO) was prepared in assay buffer and transferred to the wells of a separate micro plate. As a positive fluorescence control 37.5 μ M AMC (50 mM stock in 100% DMSO (Alfa Aesar™)) in assay buffer was used in duplicate. The substrate pre-dilution plate was pre-incubated at 25°C in a micro plate fluorescence spectrophotometer (Molecular Devices, Gemini XPS) just before the enzyme-compound assay plate incubation. Assay plate final well concentration for substrate and fluorescence control was 15 μ M. Recombinant hCTSL enzyme (R&D Systems®) was activated and pre-diluted to 0.030 ng/ μ L according to the manufacturer's protocol. Fifty μ L was added to the wells of the assay plate containing the compound pre-dilutions, and 50 μ L assay buffer was used as a negative control. The assay plate was then incubated for 10 minutes at 25°C in a micro plate fluorescence spectrophotometer. Assay plate final well concentration was 0.015 ng/ μ L. Finally, 40 μ L per well was transferred from the substrate pre-dilution plate to the assay plate and an optical adhesive seal was applied. Final well volume was 100 μ L. Fluorescence signal was measured for 10 minutes at 1 minute intervals in a micro plate fluorescence spectrophotometer at 25°C (λ_{ex} = 346 nm, λ_{em} = 440). Reaction speeds were determined using the linear part of the fluorescence signal

curves. The DMSO control (DC) reaction speed was used to calculate the activity and inhibitory potency of the samples. CDD Vault was used to calculate the IC₅₀ values using the Levenberg–Marquardt algorithm to fit the Hill equation for dose-response data to the generated inhibition curves.

5

2.1.3 CPE reduction assays

For SARS-CoV-1 and SARS-CoV-2 CPE (cytopathic effect) reduction assays, Vero E6 cells were seeded in 96-well plates at a density of 5×10^3 cells per well. For MERS-CoV cytopathic effect reduction assays were performed with Huh-7 cells (1.5×10^4 cells/well of a 96-well plate). Cells seeded on the day prior to infection, were incubated with 100 µL volumes of 2-fold serial dilutions of compounds in infection medium, followed by infection with 300 PFU of SARS-CoV-1, or SARS-CoV-2, or 225 PFU of MERS-CoV in 50 µL, yielding a total assay volume of 150 µL. Non-infected cells were treated in parallel with the same dilution series of compounds to determine their cytotoxicity. After incubation for 96 h (SARS-CoV-2), 72 h (SARS-CoV-1) or 42 h (MERS-CoV) at 37 °C, cell viability was quantified with the CellTiter-96 Aqueous Non-radioactive Cell Proliferation Kit (Promega, Madison, WI, USA) and absorption at 495 nm was measured with an EnVision multilabel plate reader (PerkinElmer). After normalization to uninfected (and untreated cells), EC₅₀ and CC₅₀ values were determined by nonlinear regression using GraphPad Prism v8.0 (San Diego, CA, USA).

20

2.1.4 Cellular M^{Pro} assay

96-well plates were coated with 1/20 Poly-L-Lysine/MQ for 5 min and subsequently washed twice with PBS. 293T cells were seeded at 15.000 cells/well in 100 µL and incubated overnight at 37 °C. Compound dilutions were prepared in 10% DMEM at 10 and 50 µM for screening or concentration ranges. Medium was removed from the wells and 100 µL of compound was added to the cells. 10 µL transfection mix containing pG5luc mix (50 ng pBIND-Mpro-VP16 (SARS2) plasmid (WT or C145A Mpro), 50 ng pG5luc plasmid, up to 5 µL OptiMEM) and lipofectamine mix (0.5 µL lipofectamine 2000, 4.5 µL OptiMEM) was added to each well and mixed by pipetting up and down once, followed by incubation for 20-24 hours. For dual luciferase readout, medium was removed from the wells and 25 µL 5x diluted lysis buffer in MQ was added to each well and incubated for 30 min at rt. 10 µL lysate was added to a white 96-well plate and firefly luciferase was measured with a luminometer, plotting FLuc/RLuc ratio. In parallel, cell viability was determined by an MTS assay. Herein, CellTiter 96® AQueous One Solution was applied on top of the existing medium. After 1-2 h incubation at 37 °C, OD490nm was measured.

35

2.1.5 Pseudovirus neutralization assay

VeroE6 cells were seeded in 96-well plates at 25.000 cells/well in 100 µL and incubated overnight at 37 °C. Compound dilutions were prepared in 10% DMEM at 10 µM for screening or concentration ranges. SARS2pp virus dilution was prepared 1:50 in 1% DMEM. Medium was removed from the wells, 50 µL of compound was added to the cells and plates were incubated for

40

1 h at 37 °C. After that, 50 µL virus dilution was added to the wells, followed by incubation overnight at 37 °C. For firefly luciferase readout, medium was removed from the wells and 60 µL 5x diluted lysis buffer in MQ was added to each well and incubated for 30 min at rt. 100 µL homemade firefly substrate (UU-FLAR) was prepared per well. Subsequently, 25 µL lysate was added to a white 96-well plate and firefly luciferase output was measured with a luminometer. Cytotoxicity was determined as described previously.

2.1.6 LogD

In vitro LogD determination at pH 7.4 was carried out using the octanol buffer partitioning method. Equal volumes of sodium phosphate buffer (10 mM, pH 7.4) and n-octanol were added to a separation funnel and mixed thoroughly by shaking and inverting the funnel several times. The two layers were allowed to separate for 2 days and then dispensed into separate glass bottles. 500 µL of organic phase (1-Octanol) was added to each well of a 2 mL deep well plate, followed by 500 µL of buffer and 15 µL of test substance was added (dissolved at 10 mM in 100% DMSO). The plate was vortexed for 1 h on a plate shaker at 1200 rpm. After incubation, the samples were allowed to equilibrate for 20 min and then centrifuged at 4000 rpm for 30 min for complete phase separation and analyzed by LC-UV (Shimadzu UFLC, Waters xBridge C18 column, 50*4.6 mm, 3.5 µM; A: 10 mM ammonium acetate/0.1% formic acid in MQ; B: 100% acetonitrile).

2.1.7 Kinetic Solubility

4 µL of 10 mM DMSO stock from the stock plate was added to the deep well plate containing 396 µL of PBS buffer pH 7.4 (100 µM final concentration, 1.0% DMSO content). The well-sealed sample plate was vortexed at 800 rpm for 24 h on a thermomixer at room temperature. At the end of the incubation period, the sample plate was centrifuged at 4000 rpm for 10 mins and analyzed in LC-UV (Shimadzu UFLC, Waters xBridge C18 column, 50*4.6 mm, 3.5 µM; A: 10 mM ammonium acetate/0.1% formic acid in MQ; B: 100% acetonitrile) against calibration curve (CC). Solubility (µM) was obtained against the CC curve.

2.1.8 Plasma Stability

In vitro plasma stability determination was carried out in mouse CD1 plasma. The frozen plasma was thawed at room temperature and centrifuged at 1400x RCF 4°C, for 15 min. Approximately 90% of the clear supernatant fraction was transferred to a separate tube and was used for the assay. Test compound stocks of 2 µM were prepared by diluting in plasma. 500 µL of plasma containing the test compound was incubated for 120 min at 37 °C in a shaker water bath with gentle shaking. 50 µL aliquots of sample at 0, 15, 30, 60 and 120 min were precipitated immediately with 300 µL acetonitrile containing internal standard and were centrifuged at 4000x RCF, 4°C for 10 min. 150 µL of supernatant was diluted with 150 µL of water and analyzed by LC-MS/MS (EXION LC, QTRAP 4500, KINETEX C18 column, 100 Å, 5 µM; A: 10 mM ammonium acetate/0.1% formic acid in MQ; B: MeOH/MeCN 1:1).

2.1.9 Plasma Protein Binding

In vitro plasma protein binding studies were carried out by ultra-centrifugation using mouse plasma. 2.4 mL of plasma (for hamster: 596 μ L) containing internal QC (Warfarin, 10 mM) was preincubated for 20 min at 37 °C. 6 μ L of test compound stock solution (800 μ M) was added and mixed properly. 25 μ L of plasma was separated and crashed with 300 μ L of acetonitrile containing internal standard. 400 μ L of blank plasma was transferred to a 0.5 mL Beckman tubes (part no. 344625) and centrifuged for 3 h, 627000 x g, 37°C (for hamster: 5 h, 223000 x g). After 45 min and after 3 h, 25 μ L of supernatant samples was separated and crashed with 300 μ L of acetonitrile containing internal standard. 25 μ L of blank plasma was added to crashed centrifuged samples for matrix matching. All stability and centrifuged samples were vortexed at 1000 rpm for 5 min and centrifuged for 10 min, 4000 rpm. The supernatant was separated, diluted 2-fold with MQ and analyzed by LC-MS/MS. Results were also obtained with hamster plasma instead of mouse plasma, as indicated.

2.1.10 Hepatocyte Stability

In vitro hepatocyte stability assays were carried out using cryopreserved mouse hepatocytes (ThermoFisher scientific) or hamster (Zenotech) or human (Lonza) hepatocytes. Hepatocytes were thawed, resuspended in prewarmed InvitroGRO HT medium and the cell count was taken using a hemacytometer. For mouse, 200 μ L of hepatocytes (0.4×10^6 cells/mL) were added to the wells (48 well plate), human 1×10^6 cells/mL, hamster 150 μ L 1×10^6 cells/mL. and pre-incubated for 30 min and then 200 μ L of compound stock (2 μ M, water/MeCN 1:1, 0.1% final DMSO concentration) was added to it and incubated in a CO₂ incubator for 120 min. The plate was vortexed at 500 rpm. At time points 0, 15, 30, 60, 90 and 120 min, 50 μ L of incubation mixture was precipitated with 200 μ L of acetonitrile containing internal standard (propranolol). At the end of the experiment assay samples were centrifuged and supernatants were submitted for LC-MS/MS analysis (EXION LC, TRIPLE-QUADE-4500; A: 10 mM ammonium acetate/0.1% formic acid in MQ; B: MeOH/MeCN 1:1). In case of Ritonavir co-administration, 75 μ L of buffer containing 1 μ M Ritonavir was added to 150 μ L cells and pre-incubated for 10 min. Subsequently, 75 μ L of 4 μ M test compound solution was added (final concentrations: 0.25 μ M Ritonavir, 1 μ M test compound). The plate was incubated, vortexed and analyzed as described above. Studies were also performed using human hepatocytes or hamster hepatocytes. Ritonavir concentrations were sometimes varied as indicated in the results section.

2.1.11 Microsomal Stability

In vitro microsomal stability assays were carried out using mouse liver microsomes with either a) 10 μ M test compound concentration, 1.0 mg/mL final protein concentration and 2.5 μ M Ritonavir or b) 1 μ M test compound concentration and 0.5 mg/mL final protein concentration. 75 μ L liver microsomes (3.33 mg/mL) were pre-incubated with 2.5 μ L test compound solution (100 μ M, water/MeCN 1:1) and 85 μ L 100 mM potassium phosphate buffer pH 7.4 for 10 min at 37 °C. 60 min samples without cofactor were obtained by mixing 32.5 μ L of pre-incubation solution and

17.5 μ L 100 mM potassium phosphate buffer pH 7.4, followed by incubation for 60 min at 37 °C. 0 min reference samples were obtained by mixing 16.25 μ L of pre-incubation solution with 200 μ L MeCN containing internal standard and 8.75 μ L cofactor (NADPH, final concentration 1 mM). The remaining pre-incubation solution was mixed with 62 μ L cofactor (NADPH, final concentration 1 mM), followed by incubation for 60 min at 37 °C. For analysis, 25 μ L aliquots of the incubated samples were mixed with 200 μ L MeCN containing internal standard. Subsequently, samples were vortexed for 5 min at 1200 rpm and centrifuged for 10 min at 4000 rpm. The supernatant was diluted 2-fold with MQ and injected on LC-MS/MS (EXION LC, QTRAP 4500, KINETEX C18 column, 100 Å, 5 μ M; A: 10 mM ammonium acetate/0.1% formic acid in MQ; B: MeOH/MeCN 1:1). In case of Ritonavir co-administration, 2.5 μ L Ritonavir (100 μ M, water/MeCN 1:1) was added to 75 μ L liver microsomes (1.666 mg/mL) and 85 μ L 100 mM potassium phosphate buffer pH 7.4, followed by pre-incubation for 15 min at 37 °C. Subsequently, 2.5 μ L of test compound solution was added and the plate was incubated, vortexed and analyzed as described above.

2.1.12 Caco-2 permeability assay

In vitro evaluation of apparent permeability was performed using 21 day cultured Caco-2 cells monolayer in 96-well format (American Type Culture Collection). 10 mM stock solutions in DMSO containing control compounds propranolol (high permeability), atenolol (low permeability) and digoxin (Pgp substrate) were diluted to 10 μ M solutions in HBSS buffer pH 7.4, and 10 mM test compound solutions in DMSO were diluted with HBSS buffer pH 7.4 to a final concentration of 2 μ M. 250 μ L DMEM was added to the basal compartment of 96-well multi-screen plates (polycarbonate high pore density, Millipore, 0.4 μ m pore size, 0.11 cm² active membrane area) and 12000 cells/well (0.16x 10⁶ cells/mL) were seeded in the apical wells, followed by CO₂ incubation at 37 °C for cell proliferation. Every alternate day, the utilized medium was replenished by fresh medium. On the day of the assay, medium was removed and washed with HBSS buffer. Plates were subsequently incubated with HBSS buffer for 30 min in an incubator and wells with TEER values greater than 230 ohm.cm² were selected for the assay. 75 μ L of test compound in HBSS was added to apical wells and 250 μ L of HBSS buffer and 1% DMSO was added to basal wells. Apical to basal permeability samples were collected at 120 min. 250 μ L of test compound in HBSS was added to basal wells and 75 μ L of HBSS buffer and 1% DMSO was added to apical wells. Basal to apical permeability samples were collected at 120 min. For sample processing, single point calibration curve in HBSS buffer was used. 25 μ L of assay samples were precipitated with 300 μ L MeCN containing internal standard (Telmisartan & Metoprolol), vortexed for 5 min at 850 rpm and centrifuged at 3500 rpm for 5 min. After centrifugation, 150 μ L of supernatant was diluted with 150 μ L MQ and submitted for LC-MS/MS analysis (EXION LC, QTRAP 6500+, Phenomenex C18 column, 50*4.6, 5 μ M; A: 10 mM ammonium acetate/0.1% formic acid in MQ; B: MeOH/MeCN 1:1).

2.1.13 Protease panel

Calpain-1, Caspase 2, Cathepsin D and Neutrophil Elastase 2 assays were performed by Eurofins Discovery using the following assay conditions:

Protease	Substrate	Source	Pre-incubation	Incubation	Incubation Buffer
Calpain-1	0.05% Casein-FITC	Human erythrocytes		0 min, 37 °C	50 mM Tris-HCl, pH 7.4
Caspase 2	25 µM CAS No. : 219138-08-6	Human recombinant <i>E. coli</i>	5 min, 37 °C	0 min, 37 °C	50 mM HEPES, pH 7.4, 100 mM NaCl, 0.1% CHAPS, 1 mM EDTA, 10% glycerol, 10 mM DTT
Cathepsin D	20 µM CAS No. : 839730-93-7	Human liver	5 min, 37 °C	0 min, 37 °C	50 mM CH ₃ COONa, pH 4.0
Neutrophil Elastase 2	200 µM CAS No. : 70967-90-7	Human neutrophils	5 min, 25 °C	0 min, 25 °C	62.5 mM HEPES, pH 7.8, 625 mM NaCl, 1.25 mg/ml BSA

- 5 Thrombin and trypsin assays were performed as usual. Continuous fluorometric assays were done in black 96-well V-bottom plates (Greiner Bio-One) using a BMG Labtech Fluostar OPTIMA microtiter fluorescence plate reader (excitation = 355 nm; emission = 460 nm). Inhibitors were preincubated with thrombin (10 nM, Sigma-Aldrich) or trypsin (1 nM, Sigma-Aldrich) in assay buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween 20) for 15 min. Enzymatic cleavage
- 10 was initiated by the addition of the substrate Cas. No: 70375-24-5 (final concentration: 50 µM, Bachem). Enzyme activities were monitored for 15 min and determined as a slope of relative fluorescence units per second (RFU/s). Camostat mesylate served as inhibition control. All experiments were performed in triplicate and percentage inhibition was calculated as the mean of the values with standard deviations. Values are reported in relation to an uninhibited control.

15

2.1.14 Pharmacokinetics in hamsters

- For in vivo pharmacokinetics (PK) studies, 6-8 weeks old female Syrian golden hamsters were used. Animals were fasted for 8-10 h and were fed 4 h post animal dosing in the case of PO administration. As vehicle for **1** and **5**, 10% DMSO, 20% PEG400, 65% PG and 5% PBS pH 7.4
- 20 was used. As vehicle for **14** and **15** 5% DMSO, 65% PG and 30% normal saline was used. The vehicle for ritonavir oral dosing was 5% DMSO, 65% PG and 30% normal saline. Animals were dosed either 1) intravenously through slow infusion during 30 min via cephalic vein, 2) intraperitoneally, 3) subcutaneously or 4) orally by gavage. All animals received ritonavir by oral administration 30 min prior to dosing. Post dose, serial blood samplings were collected (30-50 µL)
- 25 from lateral saphenous vein by using 25-gauge needle at time points 15 min, 30 min, 1 h, 2 h, 4 h,

- 6 h, 8 h, 12 h and 24 h. Blood was collected in 1.5 mL Eppendorf tubes containing 0.010 mL of 10% K₂EDTA, mixed gently and placed on ice, followed by centrifugation at 10000 rpm for 10 min. Plasma was harvested and stored at -80 °C. Compound concentrations were quantified in plasma by LCMS/MS using a fit for purpose bioanalytical method. PK data analysis was performed using noncompartmental methods in WinNonlin.

2.2 Results

2.2.1 SARS-CoV-2 M^{Pro} biochemical assay

- Compounds showed attractive IC₅₀ values for inhibiting M^{Pro} in cell-free biochemical assays. The following compounds showed an IC₅₀ under 5 µM: 1, 5, 6, 4, 3, 2. The following compounds showed an IC₅₀ under 50 nM: 1, 5, 6, 2. The following compounds showed an IC₅₀ under 10 nM: 1, 5, 6. Compound 1 was under 3 nM and also outperformed all three reference compounds. Reference compound 17 had an IC₅₀ of about 40 nM.

2.2.2 Cathepsin biochemical assays

- Compounds showed attractive IC₅₀ values for inhibiting cathepsin in cell-free biochemical assays. The following values were found.

No	IC ₅₀ (CTS)
1	29.1 nM (hCTSL) 435.7 nM (hCTSB) 85.4 nM (mCTSL) > 50 µM (mCTSB)
2	21.6 µM (hCTSL) 7.17 µM (hCTSB) 13.5 µM (mCTSL)
3	> 50 µM (hCTSL) 7.74 µM (hCTSB) 10.4 µM (mCTSL)
4	23.0 nM (hCTSL) 314 nM (hCTSB) 226 nM (mCTSL) 12.2 nM (mCTSB)
5	45.6 nM (hCTSL) 1.91 µM (hCTSB) 301 nM (mCTSL)
6	112.5 nM (hCTSL) 1.17 µM (hCTSB) 540.3 nM (mCTSL)
7	12.8 nM (hCTSL) 302 nM (hCTSB)

	96.7 nM (mCTSL)
8	22.2 nM (hCTSL) 821 nM (hCTSB) 160 nM (mCTSL)
9	85.2 nM (hCTSL) 4.18 μ M (hCTSB) 840 (mCTSL)
10	11.1 nM (hCTSL) 1.31 μ M (hCTSB) 316 nM (mCTSL)
11	626 nM (hCTSL) 6.16 μ M (hCTSB) 3.01 μ M (mCTSL)
12	0.33 nM (hCTSL) 0.222 μ M (hCTSB) 0.0219 μ M (mCTSL)
13	1.94 nM (hCTSL) 0.209 μ M (hCTSB) 0.0236 μ M (mCTSL)
14	> 50 μ M (hCTSL) 15.1 μ M (hCTSB) 35.6 μ M (mCTSL)
15	> 50 μ M (hCTSL) > 50 μ M (mCTSL)
17	3.44 μ M (hCTSB) 0.0405 μ M (hCTSL)
18	140 nM (hCTSL) 1.51 μ M (hCTSB) 0.549 μ M (mCTSL)
19	> 50 μ M (hCTSL) > 50 μ M (hCTSB) > 50 μ M (mCTSL) 1.33 μ M (mCTSB)

2.2.3 CPE reduction assays

Compounds showed attractive EC₅₀ values for reducing cytopathic effects. The following compounds showed an EC₅₀ under 1.5 μ M: 1, 15, 14, 6, 5, 10, 12, 8, 11, 3, 4, 9, 13, 7, 2. The following compounds showed an EC₅₀ under 1 μ M: 1, 15, 14, 6, 5, 10, 12, 8, 11, 3, 4, 9, 13. The following compounds showed an EC₅₀ under 750 nM: 1, 15, 14, 6, 5, 10, 12, 8, 11. The following compounds showed an EC₅₀ under 500 nM: 1, 15, 14, 6, 5, 10. The following compounds

showed an EC50 under 80 nM: 1, 15, 14, 6, 5. The following compounds showed an EC50 under 50 nM: 1, 15, 14. The following compounds showed an EC50 under 33 nM: 1, 15.

Reference compound 17 showed an EC50 of about 85 nM. Reference compound 18 showed an EC50 of about 9 nM. Reference compound 19 showed an EC50 of about 33.3 nM. It is of interest that compounds 1, 5, and 6 are all superior to reference compound 17, indicating that the invention offers an improvement over this compound that is independent of the choices for X and X'.

2.2.4 Cellular M^{Pro} assay

Compounds showed attractive EC50 values for inhibiting M^{Pro} in cellular assays. The following compounds showed an EC50 under 20 µM: 1, 2, 4, 5, 6. Compound 2 had an EC50 of over 18 µM. The following compounds showed an EC50 under 3 µM: 1, 4, 5, 6. The following compounds showed an EC50 under 35 nM: 1, 5, 6. The following compounds showed an EC50 under 10 nM: 5, 6. Reference compound 19 showed an EC50 of about 208 nM.

2.2.5 Pseudovirus neutralization assay

Compounds showed attractive EC50 values for pseudovirus neutralisation. The following compounds showed an EC50 under 6 µM: 1, 3, 4, 5, 6. The following compounds showed an EC50 under 3 µM: 1, 5, 6. Compound 1 had an EC50 of below 2 µM, about 1970 nM.

2.2.6 LogD

Compounds showed attractive lipophilicity values. The following compounds showed LogD under 5: 4, 1, 15, 14, 6, 5, 3, 20. The following compounds showed LogD under 3.5: 4, 1, 15, 14, 6, 5, 20. The following compounds showed LogD under 3: 4, 1, 15, 14, 20. The following compounds showed LogD under 2.6: 4, 1, 15, 20. The following compounds showed LogD under 2: 4, 1.

2.2.7 Kinetic Solubility

Compounds showed attractive solubility values. The following compounds showed solubility above 3.5 µM: 1, 5, 6, 12, 13, 14, 15, 20. The following compounds showed solubility above 5 µM: 1, 5, 12, 13, 14, 15, 20. The following compounds showed solubility above 10 µM: 1, 5, 14, 15, 20. The following compounds showed solubility over 50 µM: 14, 15, 20. It is of interest that compounds 14 and 15, which differ only in substituents X and X', were found to have solubility within 10% of each other, suggesting that the substituents X and X' do not have a strong impact on solubility. Reference compounds 17 and 18 had low solubility, with 17 having a solubility of 1.40 µM and 18 having a solubility of 4.09 µM.

2.2.8 Plasma Stability

Compounds showed attractive plasma stability values. Compounds 1, 4, 12, 13, 14, 15, and 16 had plasma stability well over 60%, and compounds 1, 4, 12, 14, 15, and 16 had plasma

stability well over 70%. For instance, compounds 1 and 4 both showed values in the 85-89% range, with compound 4 being more stable than compound 1. Compounds 14, 15, and 16 had stability over 90%, with compound 15 being most stable.

5 2.2.9 Plasma Protein Binding

Compounds showed attractive plasma protein binding values. All tested compounds showed values for bound fractions below 99.7%, with all but 12, 13, and 16 showing values below 99.5%. This demonstrates that an unbound amount of compound was always available. The following compounds showed over 95% plasma protein binding: 14, 1, 4, 6, 5, 12, 13, 15, 16. The following compounds showed over 98% plasma protein binding: 1, 4, 6, 5, 15, 12, 13, 16. The following compounds showed over 98.5% plasma protein binding: 4, 6, 5, 12, 13, 15, 16. The following compounds showed over 99% plasma protein binding: 5, 12, 13, 15, 16. These results were obtained using mouse plasma.

Using hamster plasma, compounds 1, 5, 6, 14, and 15 were tested. Again, all compounds showed values for bound fractions below 99.7%. Compounds 1, 4, and 6 had a bound fraction in the range of 91 to 94%. Compounds 14 and 15 had a much lower bound fraction, with compound 15 even being below 70%.

20 2.2.10 Hepatocyte Stability

Compounds showed attractive hepatocyte stability values, the below table shows results for a drug:ritonavir ratio of 4:1.

Compound	Values (mouse)	Values (human)
1	CL _{int} = 185.8 µL/min/10 ⁶ cells t _{1/2} = 9.328 min	CL _{int} = 12.5 µL/min/10 ⁶ cells t _{1/2} = 55.5 min (with ritonavir: CL _{int} = 6.9 µL/min/10 ⁶ cells t _{1/2} = 100.3 min)
3	CL _{int} = 9.308 µL/min/10 ⁶ cells t _{1/2} > 120 min	
4	CL CL _{int} int = 181.7 µL/min/10 ⁶ cells t _{1/2} = 9.539 min	
5	CL _{int} = 64.49 µL/min/10 ⁶ cells t _{1/2} = 26.87 min	
6	CL _{int} = 47.57 µL/min/10 ⁶ cells t _{1/2} = 36.43 min	
7	CL _{int} = 155.5 µL/min/10 ⁶ cells t _{1/2} = 11.14 min	
8	CL _{int} = 133.6 µL/min/10 ⁶ cells t _{1/2} = 12.97 min	
10	CL _{int} = 124.31 µL/min/10 ⁶ cells t _{1/2} = 13.94 min	

12	CL _{int} = 76.84 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 22.55 min (with ritonavir: CL _{int} = 37.09 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 46.72 min)	
13	CL _{int} = 78.77 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 22.00 min (with ritonavir: CL _{int} = 49.28 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 35.17 min)	
14	CL _{int} = 121.2 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 14.30 min (with ritonavir: CL _{int} = 21.80 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 79.49 min)	
15	CL _{int} = 112.1 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 15.46 min (with ritonavir: CL _{int} = 26.72 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 64.86 min)	CL _{int} = 6.6 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 105.3 min (with ritonavir: CL _{int} = 0.65 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = >120 min)
16	CL _{int} = 206.32 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 8.40 min (with ritonavir: CL _{int} = 210.32 $\mu\text{L}/\text{min}/10^6$ cells t _{1/2} = 8.24 min)	

Reference compounds, for instance compound 18, showed CL_{int} = 153.7 $\mu\text{L}/\text{min}/10^6$ cells and t_{1/2} = 11.27 min. Reference compound 18 showed CL_{int} = 240.27 $\mu\text{L}/\text{min}/10^6$ cells, t_{1/2} = 7.21 min without ritonavir, and CL_{int} = 117.13 $\mu\text{L}/\text{min}/10^6$ cells, t_{1/2} = 14.79 min with ritonavir. This illustrates how inhibitors where at least one of X or X' is not H show improved hepatocyte stability.

Variations in Ritonavir ratio were tested in hamster hepatocyte stability tests. Results are shown below. The top value in each cell is CL_{int} in $\mu\text{L}/\text{min}/10^6$ cells, the bottom value is t_{1/2}.

Compound	No Ritonavir	Compound:ritonavir 4:1	2:1	1:1
1		77.1 9.0 min	48.13 14.40 min	31.40 22.08 min
5		29.7 23.3 min	CL _{int} = 31.70 21.87 min	22.57 30.71 min
6		19.4 35.7 min	30.73 22.56 min	16.42 42.21 min
14		41.9 16.6 min	27.36 25.33 min	12.48 55.55 min
15		44.7 5.5 min	25.44 27.24 min	9.30 74.50 min
20	48.19 14.38 min	51.13 13.56 min	43.60 15.90 min	26.01 26.65 min

2.2.11 Microsomal Stability

Compounds showed attractive microsomal stability values. For instance, compounds 1, 4, 5, and 6 showed good half-life and clearance values.

Compound	Values
5	CL = 1100 $\mu\text{L}/\text{min}/\text{mg}$ protein $t_{1/2}$ = 1.28 min
4	CL = 1615 $\mu\text{L}/\text{min}/\text{mg}$ protein $t_{1/2}$ = 0.86 min
1	CL = 586 $\mu\text{L}/\text{min}/\text{mg}$ protein $t_{1/2}$ = 1.18 min
6	CL = 780 $\mu\text{L}/\text{min}/\text{mg}$ protein $t_{1/2}$ = 1.90 min

- 5 Reference compounds, for instance compound 18, were immediately degraded, to such an extent that no clearance values or half-life values could be determined. In the presence of ritonavir, compound 1 had clearance of 4.3 $\mu\text{L}/\text{min}/\text{mg}$ protein, and a half-life of over 2 hours.

2.2.12 Caco-2 permeability assay

- 10 Compounds showed attractive cell uptake and permeability values. For instance, compounds 5, 6, 14, and 15 showed good results in Caco-2 assays. In the table, A to B is apical to basolateral, and vice versa. P_{app} is the permeability coefficient.

Compound	Values
5	P_{app} (A to B) = $0.21 \cdot 10^{-6}$ cm/sec P_{app} (B to A) = $13.30 \cdot 10^{-6}$ cm/sec Efflux Ratio = 65.14
6	P_{app} (A to B) = $0.16 \cdot 10^{-6}$ cm/sec P_{app} (B to A) = $12.78 \cdot 10^{-6}$ cm/sec Efflux Ratio = 82.28
14	P_{app} (A to B) = $0.24 \cdot 10^{-6}$ cm/sec P_{app} (B to A) = $16.57 \cdot 10^{-6}$ cm/sec Efflux Ratio = 78.75
15	P_{app} (A to B) = $0.00 \cdot 10^{-6}$ cm/sec P_{app} (B to A) = $24.53 \cdot 10^{-6}$ cm/sec Efflux Ratio > 24.53

2.2.13 Protease panel

Inhibition (%) at 10 μ M testing concentration of selected compounds in a panel of proteases.

Compound	Calpain-1	Caspase 2	Cathepsin D	Neutrophil Elastase 2	Thrombin	Trypsin
1	3	11	8	2	0.2 $\pm$ 1.1	-3.4 $\pm$ 9.0
14	N.D.	N.D.	N.D.	N.D.	0.4 $\pm$ 0.9	9.9 $\pm$ 6.6
5	5	20	9	2	-0.2 $\pm$ 0.6	7.5 $\pm$ 4.9
15	-7	7	-23	4	4.1 $\pm$ 1.5	10.2 $\pm$ 3.3

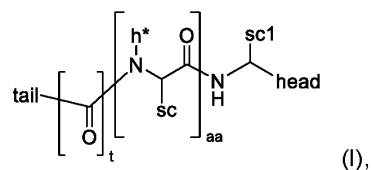
2.2.14 In vivo pharmacokinetics

- 5 The table below shows in vivo pharmacokinetics parameters of selected compounds pretreated with ritonavir (100 mg/kg) in female Syrian golden hamsters following intravenous (I.V., 5 mg/kg), intraperitoneal (I.P., 100 mg/kg), subcutaneous (S.C., 100 mg/kg) and oral (P.O., 200 mg/kg) administration (n=3).

	1			
	I.V.	I.P.	S.C.	P.O.
C _{max} [ng/mL]		1874 $\pm$ 291	1206 $\pm$ 196	650 $\pm$ 338
t _{max} [h]		4.00 $\pm$ 0.00	1.00 $\pm$ 0.00	1.00 $\pm$ 0.00
t _{1/2} [h]	2.77 $\pm$ 0.76	8.23 $\pm$ 1.68	11.1 $\pm$ 2.18	24.8 $\pm$ 4.21
AUC _{0-inf} [ng*h/mL]	1443 $\pm$ 203	19542 $\pm$ 2955	14205 $\pm$ 4940	2511 $\pm$ 1970
%F		62	40	3.0
	5			
	I.V.	I.P.	S.C.	P.O.
C _{max} [ng/mL]		951 $\pm$ 221	372 $\pm$ 31	29 $\pm$ 4.5
t _{max} [h]		1.00 $\pm$ 0.00	3.33 $\pm$ 1.15	1.50 $\pm$ 2.17
t _{1/2} [h]	2.47 $\pm$ 0.35	3.45 $\pm$ 1.18	4.11 $\pm$ 0.79	1.52
AUC _{0-inf} [ng*h/mL]	3420 $\pm$ 416	6109 $\pm$ 735	3003 $\pm$ 155	31
%F		8.2	4.3	0.0
	14			
	I.V.	I.P.	S.C.	P.O.
C _{max} [ng/mL]		4552 $\pm$ 303	2575 $\pm$ 500	597 $\pm$ 56
t _{max} [h]		0.42 $\pm$ 0.14	1.00 $\pm$ 0.00	0.33 $\pm$ 0.14
t _{1/2} [h]	3.54 $\pm$ 0.68	4.68 $\pm$ 0.74	3.17 $\pm$ 0.64	7.54 $\pm$ 6.40
AUC _{0-inf} [ng*h/mL]	6454 $\pm$ 778	9528 $\pm$ 79	12889 $\pm$ 2489	1633 $\pm$ 210
%F		7.3	10.0	0.6
	15			
	I.V.	I.P.	S.C.	P.O.
C _{max} [ng/mL]		11931 $\pm$ 1965	4096 $\pm$ 643	1581 $\pm$ 92
t _{max} [h]		3.00 $\pm$ 1.73	6.00 $\pm$ 0.00	2.67 $\pm$ 1.15
t _{1/2} [h]	3.82 $\pm$ 0.28	5.30 $\pm$ 1.68	3.45 $\pm$ 0.13	3.20 $\pm$ 1.60
AUC _{0-inf} [ng*h/mL]	8633 $\pm$ 129	141356 $\pm$ 21008	49595 $\pm$ 5012	20061 $\pm$ 2749
%F		78.2	28.6	5.4

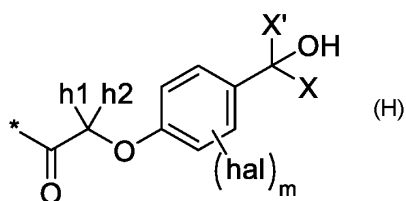
Claims

1. A compound of general formula (I), or a salt thereof:



wherein

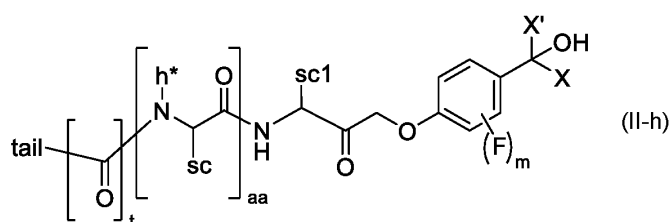
- 5 **sc1** and each instance of **sc** are independently an amino acid side chain;
 h* is H or -CH₃ or -CH₂CH₃ or -C(=O)CH₃, or **h*** together with its adjacent **sc** forms a cyclic amino acid side chain;
 aa is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;
 t is 0 or 1;
 10 **tail** is a linear, branched, or cyclic C1-22 alkyl, alkoxy, or alkylamine, wherein the alkyl, alkoxy, or alkylamine is optionally unsaturated and optionally substituted with halogen, alkylamine, alkylamide, aminoalkylamine, aminoalkylamide, alkoxy, or haloalkoxy, or is an independently chosen instance of **sc1**;
 head is a warhead of general formula (H)



- 15 wherein
 h1 and **h2** are each independently H, halogen, or C1-4 alkyl, or **h1** and **h2** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure;
 hal is a halogen;
 20 **m** is 0, 1, 2, 3, or 4;
 X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-6 membered cyclic structure.
- 25 2. The compound according to claim 1, wherein
 h1 and **h2** are H;
 hal is F;
 m is 3 or 4, preferably 4; and/or
 X and **X'** are each independently H or C1-2 alkyl, preferably H or -CH₃.

- 30 3. The compound according to claim 1 or 2, wherein it is of general formula (II-h)

99



wherein

m is 1, 2, 3, or 4;

X and **X'** are each independently H, halogen, or C1-4 alkyl, or **X** and **X'** together with the carbon atom to which they are attached form a 3-4 membered cyclic structure, preferably **X** and **X'** are each independently H or -CH₃.

4. The compound according to any one of claims 1-3, wherein each instance of **sc** is independently

a linear, branched, or cyclic C1-12 alkyl, wherein alkyl is optionally unsaturated and optionally substituted with halogen, or

-(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, or

sc forms a cyclic amino acid side chain together with its adjacent **h***, wherein the cyclic amino acid side chain together with the carbon and nitrogen atoms to which it is attached forms a 4-7 membered ring.

5. The compound according to any one of claims 1-4, wherein **tail** is

a linear or branched C1-20 alkyl or alkoxy, wherein the alkyl or alkoxy is optionally unsaturated and optionally substituted with halogen, alkoxy, or haloalkoxy, or

-(CH₂)₀₋₄-[C5-10 (hetero)aryl] that is optionally substituted with halogen, alkyl, alkoxy, or

-C(=O)O(CH₂)₀₋₄H.

6. The compound according to any one of claims 1-5, wherein

aa is 1 or 2, preferably 1;

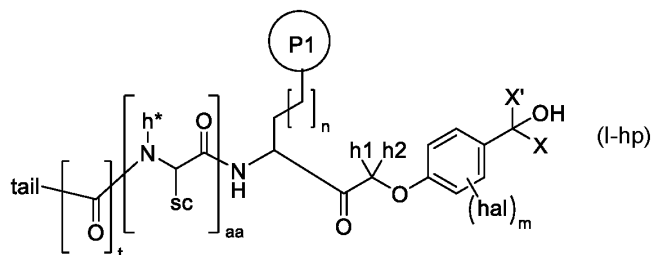
t is 0 or 1, preferably 1;

h* is H or together with its adjacent **sc** forms a cyclic amino acid side chain;

sc is butyl such as 2-methyl-n-propyl, is optionally substituted -CH₂-aryl such as -CH₂-fluorophenyl, or is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, wherein optional substitutions are preferably selected from halogen of alkoxy; and/or

tail is linear C12-18 alkyl such as -C₁₅H₃₁, is optionally substituted heteroaryl such as indolyl, preferably 1*H*-indol-3-yl, more preferably methoxyl-1*H*-indol-3-yl, or is optionally substituted C4-11 alkoxy such as -O-CH₂-phenyl, wherein optional substitutions are preferably selected from halogen or alkoxy, preferably alkoxy, more preferably methoxyl.

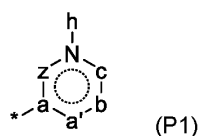
7. The compound according to any one of claims 1-6, wherein it is of general formula (I-hp)



wherein

n is 0 or 1;

- 5 **P1** is a cyclic structure of general formula (P1):



wherein

a is C, CH, or N;

a' is absent or is CH or -C(=O)- or N or NH;

- 10 **b** is CH or N or NH;

c is CH or -C(=O)- or N or NH;

h is H or is absent; and

z is -C(=O)- or N.

- 15 8. The compound according to claim 7, wherein

when **a** is CH, **n** is 0;

when **a'** is CH, **n** is 0;

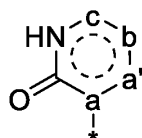
when **a'** is absent and **z** is N, **n** is 0;

when **a'** is absent, **n** is 1; or

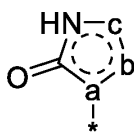
- 20 when **a'** is absent and **a** is N and **z** is -C(=O)-, **n** is 1;

or wherein **a** is N or C.

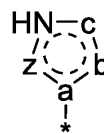
9. The compound according to claim 7 or 8, wherein **P1** is selected from



(P1a)

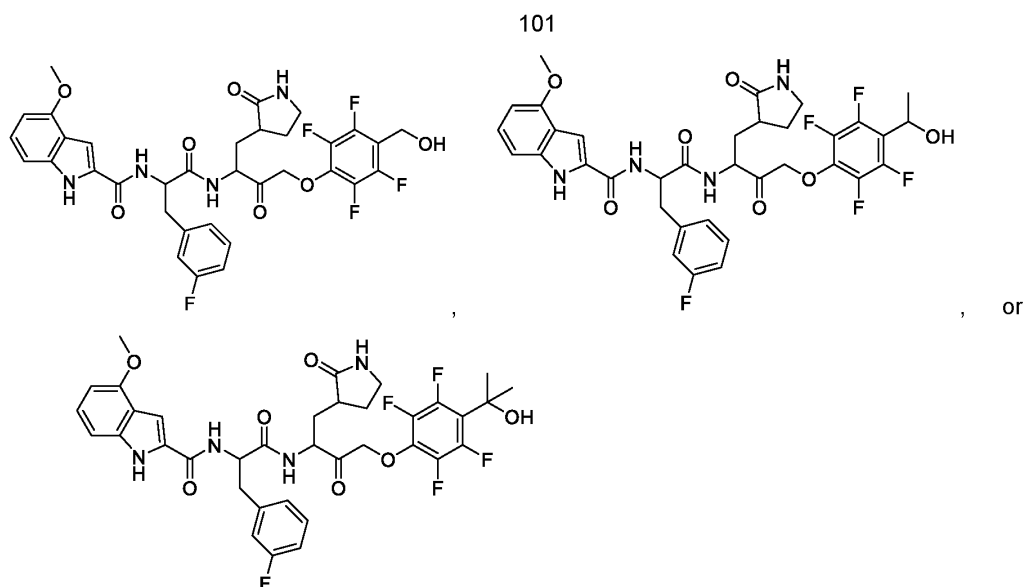


(P1b)

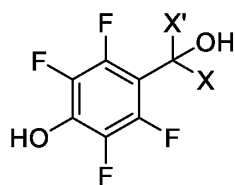


(P1c).

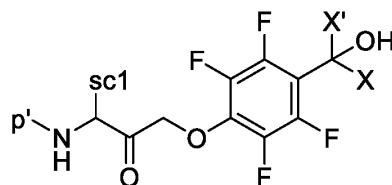
- 25 10. The compound according to any one of claims 1-9, wherein it is selected from



11. A composition comprising
 5 a compound as defined in any one of claims 1-10, and
 a pharmaceutically acceptable excipient,
 preferably wherein the composition is a pharmaceutical composition.
12. A compound as defined in any one of claims 1-10, or a composition as defined in claim 11,
 10 for use as a medicament.
13. A compound according to general formula (intermediate-1) or (intermediate-2):



(intermediate-1)



(intermediate-2)

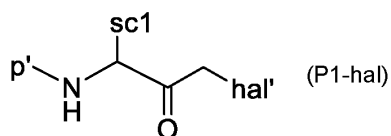
wherein

- p'** is H or a protecting group;
 15 **sc1** is an amino acid side chain;
X is H, halogen, or C1-4 alkyl;
X' is H, halogen, or C1-4 alkyl;
 or **X** and **X'** together with the carbon atom to which they are attached form a 3-6
 membered cyclic structure;
 20 wherein for the compound of general formula (intermediate-1) **X'** is halogen or
 C1-4 alkyl when X is H.

102

14. Method for providing a compound according to any one of claims 1-10, the method comprising the step of

- i) contacting a compound according to general formula (intermediate-1) according to claim 13 with a compound of general formula (P1-hal):



5

wherein

p' is a protecting group;

sc1 is an amino acid side chain;

hal' a halogen, preferably chloride;

10

to obtain a compound of general formula (intermediate-2) according to claim 13; and/or

- ii) removing the protecting group from a compound according to general formula (intermediate-2) according to claim 13, preferably by using a carboxylic acid such as trifluoroacetic acid.

15. A method for inhibiting a viral protease, the method comprising the steps of:

- i) providing a compound as defined in any one of claims 1-10, or a composition as defined in claim 11;
- ii) contacting the viral protease with the provided compound or composition.

20