target gene, wherein the polynucleotide cassettes comprise the aptamers disclosed herein. Further provided are methods of using said aptamers, riboswitches, and/or polynucleotide cassettes for the regulation of target genes, including therapeutic genes. Also provided herein are small molecules that are modulators of target gene expression where the target gene contains a riboswitch comprising an aptamer described herein.

[0006] In one example, the disclosure provides an aptamer comprising the aptamer encoding sequence disclosed herein. In examples, the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCC$X_7X_8X_9X_{10}X_{11}X_{12}CCTX_{13}X_{14}X_{15}$CCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGG$X_{22}X_{23}$C AGGGAG (SEQ ID NO:2); wherein:

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is G or T;

X_5 is A, G, or T;

X_6 is A or G;

X_7 is A or T;

X_8 is A, C, or T;

X_9 is A, C, or T;

X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide;

X_{12} is A;

X_{13} is A, C, or G;

X_{14} is any nucleotide;

X_{15} is C, G, or T;

X_{16} is G or T;

X_{17} is A or T;

X_{18} is any nucleotide;

X_{19} is A or G;

X_{20} is A, G, T;

X_{21} is C, G, T;

X_{22} is T; and

X_{23} is A, G, or T (taken together SEQ ID NO:681).

[0006a] In one aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCC
ATTGCACCTGATCCGGX16X17X18X19X20X21CCGGCGCAGGGAG (SEQ ID NO:6); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:6 is any nucleotide or no nucleotide; X16 is any nucleotide; X17 is any nucleotide; X18 is any nucleotide; X19 is any nucleotide; X20 is any nucleotide; and X21 is C, G, T.

[0006b] In another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX1X2AX3X4X5X6CCATC
GACCCX7X8X9X10X11X12CCTX13X14X15CCGGX16X17X18X19X20X21CCGGX22
X23CAGGGAG (SEQ ID NO:2); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:2 is any nucleotide or no nucleotide; X1 is C or T; X2 is any nucleotide; X3 is any nucleotide; X4 is G or T; X5 is A, G, or T; X6 is A or G; X7 is A or T; X8 is A, C, or T; X9 is A, C, or T; X10 is any nucleotide; X11 is any nucleotide or no nucleotide; X12 is A; X13 is A, C, or G; X14 is any nucleotide; X15 is C, G, or T; X16 is G or T; X17 is A or T; X18 is any nucleotide; X19 is A or G; X20 is A, G, T; X21 is C, G, T; X22 is T; and X23 is A, G, or T.**[0006c]** In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCC
X7X8X9X10X11X12CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:4 is any nucleotide or no nucleotide; X7 is A, G, or T; X8 is any nucleotide; X9 is any nucleotide; X10 is any nucleotide; X11 is any nucleotide or no nucleotide; X12 is A, C, or T.

[0006d] In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX1X2AX3X4X5X6CCATC GACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:3 is any nucleotide or no nucleotide; X1 is C, G, or T; X2 is any nucleotide; X3 is any nucleotide; X4 is any nucleotide; X5 is any nucleotide; and X6 is any nucleotide.

[0006e] In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCC ATTGCACCTX13X14X15CCGGATCATGCCGGX22X23CAGGGAG (SEQ ID NO:5); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:5 is any nucleotide or no nucleotide; X13, X14, X15, X22, and X23 is any nucleotide.

[0007] In examples, the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCC$X_7X_8X_9X_{10}X_{11}X_{12}CCTX_{13}X_{14}X_{15}$CCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGG$X_{22}X_{23}$C AGGGAG (SEQ ID NO:2); wherein:

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is G or T;

X_5 is A, G, or T;

X_6 is A or G;

X_7 is A;

X_8 is A, C, or T;

X_9 is A, C, or T;

X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide;

X_{12} is A;

X_{13} is A, C, or G;

X_{14} is any nucleotide;

X_{15} is C, G, or T;

X_{16} is G or T;

X_{17} is A or T;

X_{18} is any nucleotide;

X_{19} is A or G;

X_{20} is A, G, T;

X_{21} is C, G, T;

X_{22} is T; and

X_{23} is A, G, or T (taken together SEQ ID NO:682).

[0008] In examples, the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCC$X_7X_8X_9X_{10}X_{11}X_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:

X_7 is A, G, or T;

X_8 is any nucleotide;

X_9 is any nucleotide;

X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide;

X_{12} is A, C, or T (taken together SEQ ID NO:683).

[0009] In examples, X_7-X_{12} are not simultaneously A, T, T, G, C, and A, respectively.

[0010] In examples, the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:

X_7 is A or T;

X_8 is A, C, or T;

X_9 is A, C, or T;

X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide; and

X_{12} is A (taken together SEQ ID NO:684).

[0011] In examples, X_7-X_{12} are not simultaneously A, T, T, G, C, and A, respectively.

[0012] In examples, the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:

X_7 is A;

X_8 is A, C, or T;

X_9 is A, C, or T;

X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide; and

X_{12} is A (taken together SEQ ID NO:685).

[0013] In examples, X_7-X_{12} are not simultaneously A, T, T, G, C, and A, respectively.

[0014] In examples, the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX$_1$X$_2$AX$_3$X$_4$X$_5$X$_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:

X_1 is C, G, or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is any nucleotide;

X_5 is any nucleotide; and

X_6 is any nucleotide (taken together SEQ ID NO:686).

[0015] In examples, X_1-X_6 are not simultaneously C, A, T, C, G, and A, respectively.

[0016] In examples, the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is any nucleotide;

X_5 is A, G, or T; and

X_6 is any nucleotide (taken together SEQ ID NO:687).

[0017] In examples, X_1-X_6 are not simultaneously C, A, T, C, G, and A, respectively.

[0018] In examples, the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is G or T;

X_5 is A, G, or T; and

X_6 is A or G (taken together SEQ ID NO:688).

[0019] In examples, the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCATTGCACCT$X_{13}X_{14}X_{15}$CCGGATCATGCCGG$X_{22}X_{23}$CAGGGAG (SEQ ID NO:5); wherein:

X_{13}, X_{14}, X_{15}, X_{22}, and X_{23} is any nucleotide.

[0020] In examples, X_{13}, X_{14}, X_{15}, X_{22}, and X_{23} are not simultaneously G, A, T, C, and G, respectively.

[0021] In examples, the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCATTGCACCT$X_{13}X_{14}X_{15}$CCGGATCATGCCGG$X_{22}X_{23}$CAGGGAG (SEQ ID NO:5); wherein:

X_{13} is A, C, or G;

X_{14} is any nucleotide;

X_{15} is C, G, or T;

X_{22} is T; and

X_{23} is A, G, or T (taken together SEQ ID NO:689).

[0022] In examples, the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCATTGCACCTGATCCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGGCGCAGGGAG (SEQ ID NO:6); wherein:

X_{16} is any nucleotide;

X_{17} is any nucleotide;

X_{18} is any nucleotide;

X_{19} is any nucleotide;

X_{20} is any nucleotide; and

X_{21} is C, G, T (taken together SEQ ID NO:690).

[0023] In examples, X_{16}-X_{21}, are not simultaneously A, T, C, A, T, and G, respectively.

[0024] In examples, the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCATTGCACCTGATCCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGGCGCAGGGAG (SEQ ID NO:6); wherein:

X_{16} is G or T;

X_{17} is A or T;

X_{18} is any nucleotide;

X_{19} is A or G;

X_{20} is A, G, T; and

X_{21} is C, G, T (taken together SEQ ID NO:691).

[0025] In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 1 and 7-558. In examples, the aptamer encoding sequence comprises a sequence that is at least 95% identical, or at least 99% identical to a sequence selected from the group consisting of SEQ ID NOs: 1 and 7-558.

In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 1 and 7-558. In examples, the aptamer encoding sequence comprises a sequence that is selected from the group consisting of SEQ ID NOs: 1 and 7-558.

[0026] In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583. In examples, the aptamer encoding sequence comprises a sequence that is at least 95% identical, or at least 99% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583. In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding is sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583. In examples, the aptamer encoding

sequence comprises a sequence that is selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583.

[0027] In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447. In examples, the aptamer encoding sequence comprises a sequence that is at least 95% identical, or at least 99% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447. In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447. In examples, the aptamer encoding sequence comprises a sequence that is selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447.

[0028] In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378. In examples, the aptamer encoding sequence comprises a sequence that is at least 95% identical, or at least 99% identical to a sequence selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378. In yet another aspect, the present disclosure provides for a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378. In examples, the aptamer encoding sequence comprises a sequence that is selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378.

[0029] In examples, the aptamer sequence disclosed herein, further comprises additional sequence at the 5′ and 3′ ends that is complementary and capable of forming part of the aptamer P1 stem. In examples, this P1 stem of the aptamer is, comprises, or overlaps with the effector region of the riboswitches disclosed herein. In examples, the aptamer P1 stem comprises a 5′ splice site sequence of a 3′ intron

and sequence complementary thereto. For example, the P1 stem may comprise A G G G T G A G T; A A A G T A A G C; G C A G TA A G T; G A G G T G T G G; A/C A G G T A/G A G T; N A G G T A/G A G T; N A G G T A A G T; A/C A/T G G T A N G T; or N A G/A G T A A G T (where N can be A, G, C, or T).

[0030] In examples, the aptamers disclosed herein bind to one or more of the small molecules of Formula I to XXII, including those listed in Table A.

[0031] In one example, the disclosure provides the RNA aptamer encoded by the aptamer encoding sequences disclosed herein.

[0032] In one example, the disclosure provides nucleic acid sequence encoding a recombinant riboswitch for the regulation of target gene expression in response to a small molecule, wherein the riboswitch comprises an aptamer disclosed herein.

[0032a] In another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCC ATTGCACCTGATCCGGX16X17X18X19X20X21CCGGCGCAGGGAG (SEQ ID NO:6); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:6 is any nucleotide or no nucleotide; X16 is any nucleotide; X17 is any nucleotide; X18 is any nucleotide; X19 is any nucleotide; X20 is any nucleotide; and X21 is C, G, T; wherein X16-X21, are not simultaneously A, T, C, A, T, and G, respectively.

[0032b] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX1X2AX3X4X5X6CCATC GACCCX7X8X9X10X11X12CCTX13X14X15CCGGX16X17X18X19X20X21CCGGX22 X23CAGGGAG (SEQ ID NO:2); wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:2 is any nucleotide or no nucleotide; X1 is C or T; X2 is any nucleotide; X3 is any nucleotide; X4 is G or T; X5 is A, G, or T; X6 is A or G; X7 is A; X8 is A, C, or T; X9 is A, C, or T; X10 is any nucleotide; X11 is any nucleotide or no nucleotide; X12 is A; X13 is A, C, or G; X14 is any nucleotide; X15 is C, G, or T; X16 is G or T; X17 is A or T; X18 is any nucleotide; X19 is A or G; X20 is A, G, T; X21 is C, G, T; X22 is T; and X23 is A, G, or T.

[0032c] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCC X7X8X9X10X11X12CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4);

wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:4 is any nucleotide or no nucleotide; X7 is A, G, or T; X8 is any nucleotide; X9 is any nucleotide; X10 is any nucleotide; X11 is any nucleotide or no nucleotide; X12 is A, C, or T, wherein X7-X12 are not simultaneously A, T, T, G, C, and A, respectively.

[0032d] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX1X2AX3X4X5X6CCATC GACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3);

wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:3 is any nucleotide or no nucleotide; X1 is C G, or T; X2 is any nucleotide; X3 is any nucleotide; X4 is any nucleotide; X5 is any nucleotide; and X6 is any nucleotide; wherein X1-X6 are not simultaneously C, A, T, C, G, and A, respectively.

[0032e] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCC ATTGCACCTX13X14X15CCGGATCATGCCGGX22X23CAGGGAG (SEQ ID NO:5);

wherein: the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:5 is any nucleotide or no nucleotide; X13, X14, X15, X22, and X23 is any nucleotide, wherein X13, X14, X15, X22, and X23 are not simultaneously G, A, T, C, and G, respectively.

[0032f] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 1 and 7-558. In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 1 and 7-558.

[0032g] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583. In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an

aptamer that binds to a small molecule, wherein the aptamer encoding is sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583.

[0032h] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447. In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447.

[0032i] In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378. In yet another aspect, the present disclosure provides for a nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378.

[0033] In another example, the disclosure provides a polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises an aptamer encoding sequence disclosed herein.

[0034] In a further aspect, the present disclosure provides for a nucleic acid sequence encoding a recombinant riboswitch for the regulation of target gene expression in response to a small molecule, wherein the riboswitch comprises an aptamer encoded by SEQ ID NOs: 1 and 7-558 or a sequence that is at least 95% identical to SEQ ID NOs: 1 and 7-558. In yet another aspect, the present disclosure provides for a polynucleotide cassette for the regulation of the expression of a target gene in response to a small molecule, the polynucleotide cassette comprising: (a) a riboswitch; and (b) an alternatively-spliced exon, flanked by a 5' intron and a 3' intron, wherein the riboswitch comprises (i) an effector region comprising a stem that includes the 5' splice site sequence of the 3' intron, and (ii) the aptamer comprises a sequence of SEQ ID NOs: 1 and 7-558 or a sequence that is at least 95% identical to SEQ ID NOs: 1 and 7-558; and wherein the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene when the alternatively-spliced exon is spliced into the target gene mRNA. In examples, the polynucleotide cassette comprises sequence encoding:

(a) a riboswitch; and

(b) an alternatively-spliced exon, flanked by a 5′ intron and a 3′ intron,

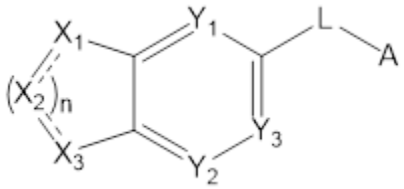
wherein the riboswitch comprises (i) an effector region comprising a stem forming sequence that includes the 5′ splice site sequence of the 3′ intron (and sequence complementary to the 5′ splice site sequence of the 3′ intron), and (ii) the aptamer comprises an aptamer sequence disclosed herein; and

wherein the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene when the alternatively-spliced exon is spliced into the target gene mRNA.

[0035] In examples, the effector stem is, or comprises, a P1 stem of the aptamers disclosed herein. In other words, the effector stem comprises a first sequence that is linked to the 5′ end of the aptamers disclosed herein and a second sequence that is linked to the 3′ end of the aptamers disclosed herein.

[0036] In examples, the polynucleotide cassette is located in the protein coding sequence of the target gene. In examples, the polynucleotide cassette is located in an untranslated region of the target gene or in an intron of the target gene.

[0037] In examples, the small molecule has the structure according to Formula I:



I

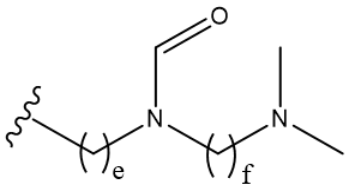
wherein

X_1, X_2, and X_3 are, in each instance, independently selected from CR_1, CHR_1, N, NH, O and S, wherein adjacent X_1, X_2, and X_3 are not simultaneously selected to be O or S;

the dashed lines represent optional double bonds;

Y_1, Y_2, and Y_3 are, in each instance, independently selected from CR_2 and N;

n is 1 or 2, wherein when n is 1, only one of the dashed lines is a double bond;

L-A is 

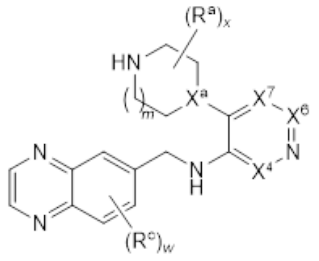
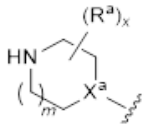
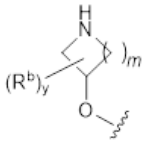
, or

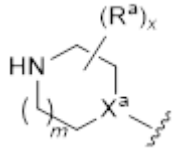
L is selected from

[0038] In examples, the small molecule has a structure according to Formula II-XXII, including, e.g., a structure provided in Table A.

[0039] In one aspect, the present disclosure provides for a vector comprising the polynucleotide cassette as disclosed herein, or the nucleic acid sequence as disclosed herein. In one example the disclosure provides a vector comprising a polynucleotide cassette, an aptamer encoding sequence/aptamer sequence, or riboswitch disclosed herein. In examples, the vector is a viral vector or a non-viral vector. In examples, the viral vector is an adenoviral vector, an adeno-associated virus vector, and a lentiviral vector.

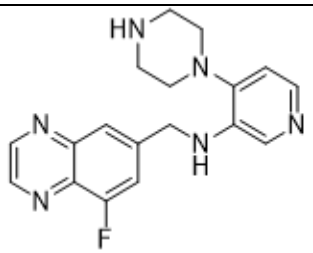
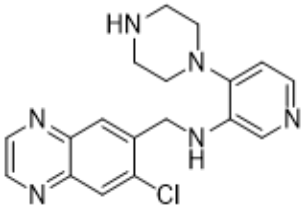
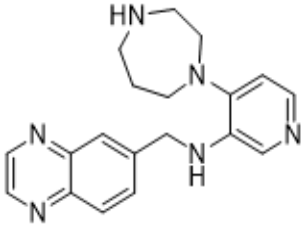
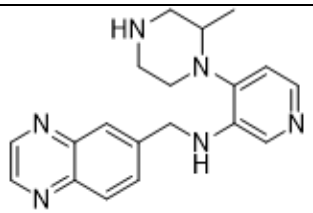
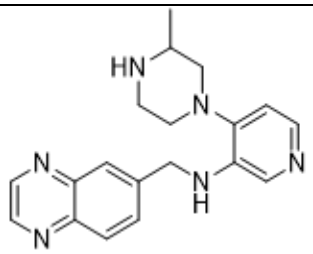
[0040] In one aspect, the present disclosure provides for a cell comprising the vector, the polynucleotide cassette, or the nucleic acid sequence as disclosed herein. In one example, the disclosure provides a cell comprising a vector, a polynucleotide cassette, an aptamer encoding sequence/aptamer sequence, or riboswitch disclosed herein.

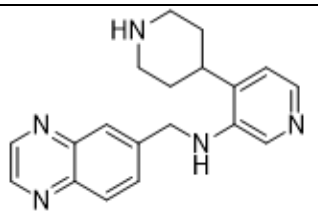
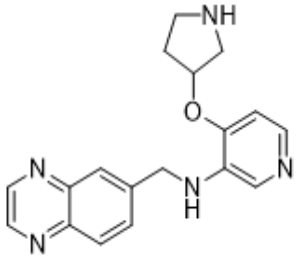
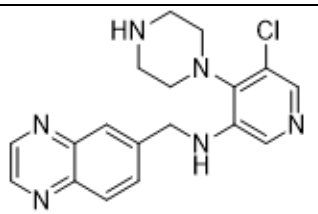
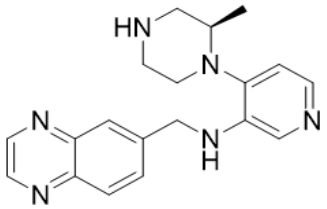
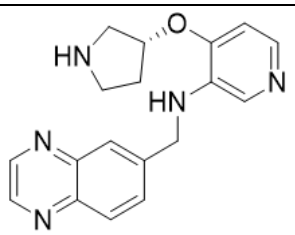
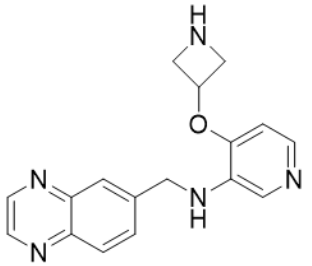
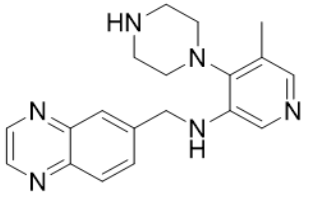
[0040a] In another aspect, the present disclosure provides for a compound having the structure according to formula XIV  or a pharmaceutically acceptable salt thereof, wherein A is selected from the group consisting of:  and ; X^a is selected from N and CH; each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH; each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH; m is 1 or 2; x is 0, 1, 2 or 3; y is 0, 1, 2 or 3; each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino; each R^d is independently selected from halo,

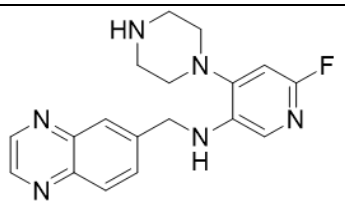
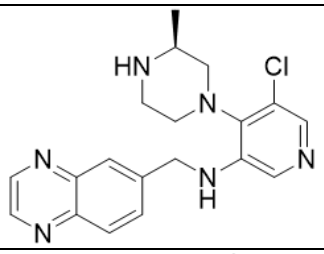
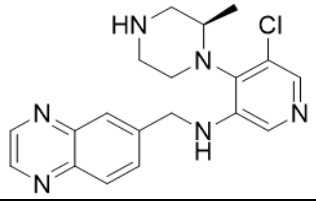
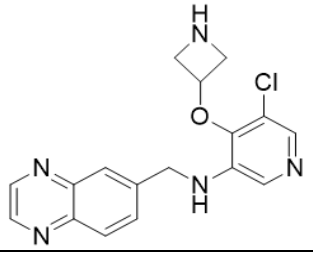
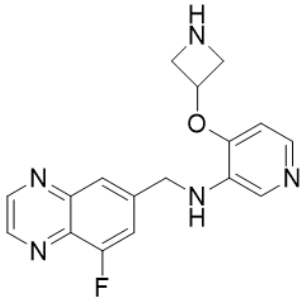
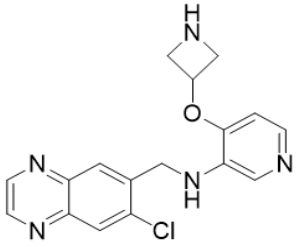
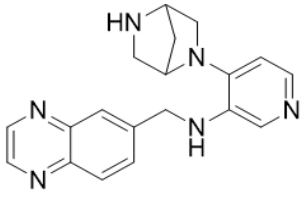
C_1 to C_3 alkyl, -OCH_3, -CF_3, -CH_2F, -CHF_2, -CN, hydroxyl and amino; alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH; w is 0, 1 or 2; and z is 0, 1 or 2; wherein when A is selected to be 

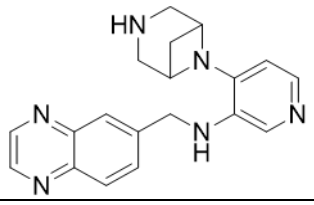
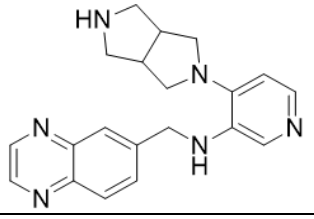
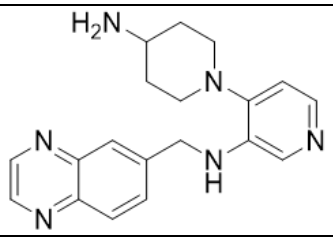
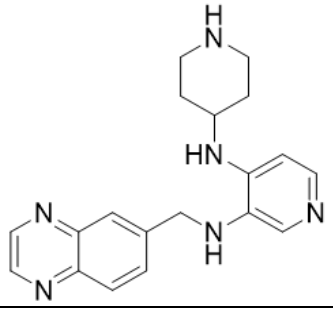
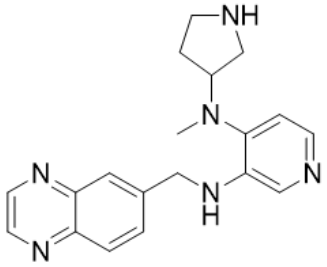
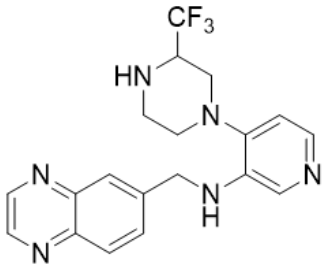
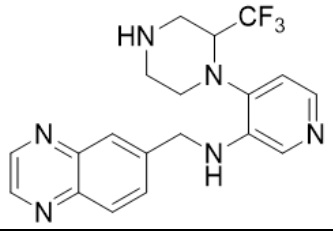
, x is 1, 2 or 3; and/or two R^d on adjacent ring positions are taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH.

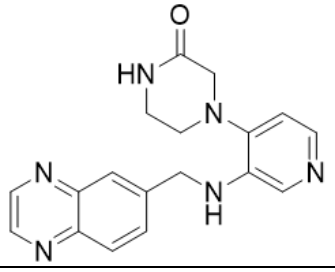
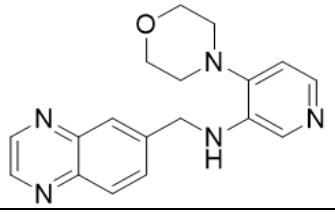
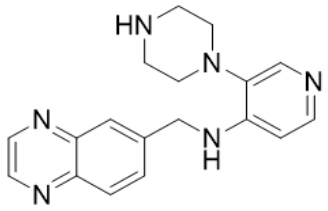
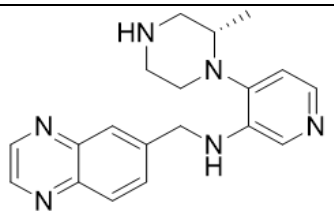
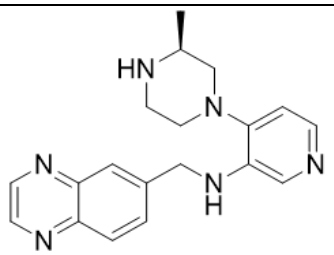
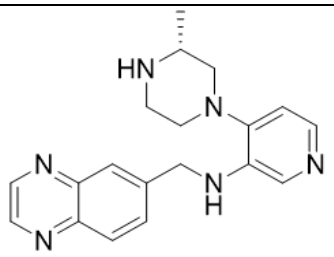
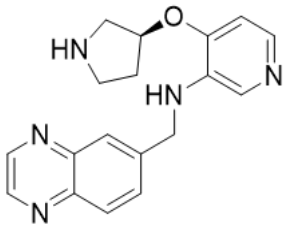
[0040b] In another aspect, the present disclosure provides for a compound having the structure according to one of:

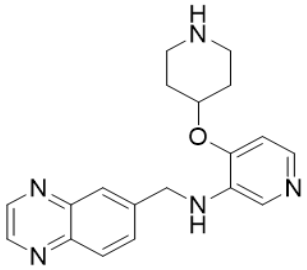
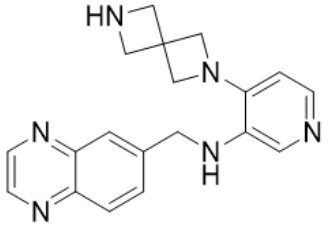
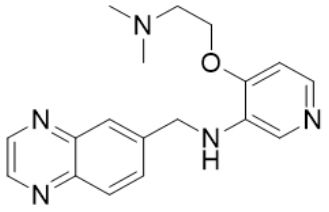
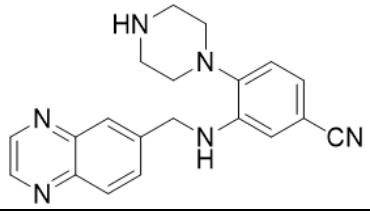
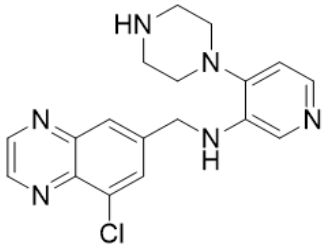
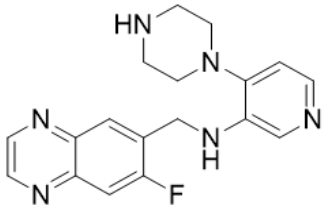
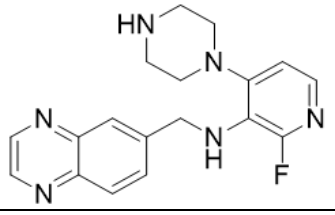
Ref. No.	Structure
012	
013	
014	
015	
016	

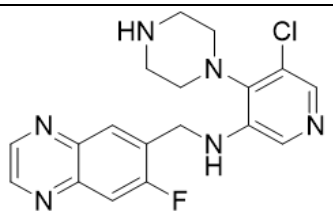
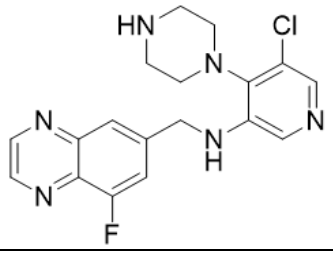
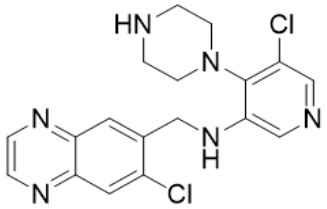
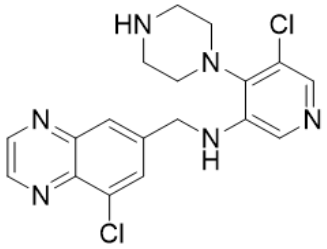
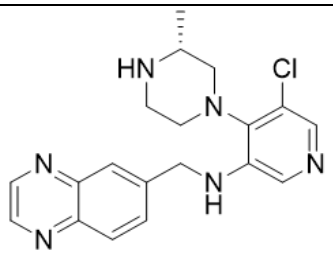
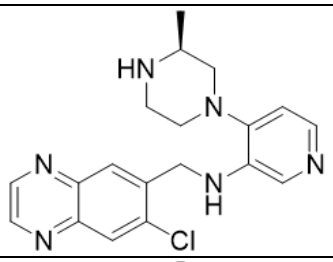
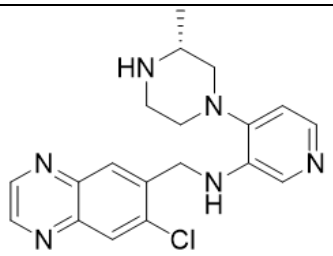
017	
018	
019	
020	
021	
022	
023	

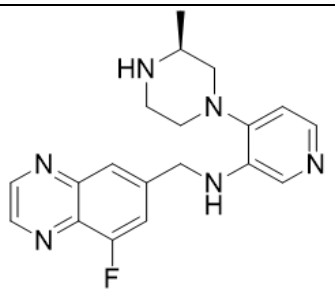
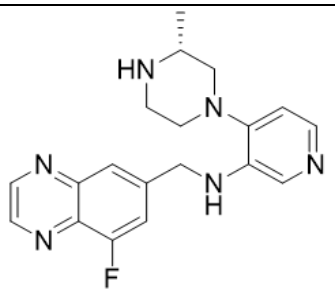
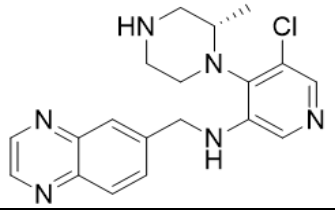
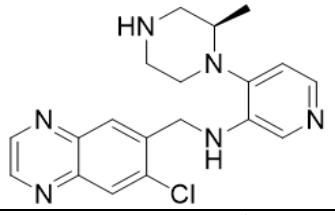
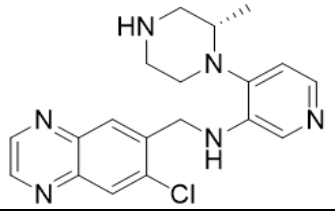
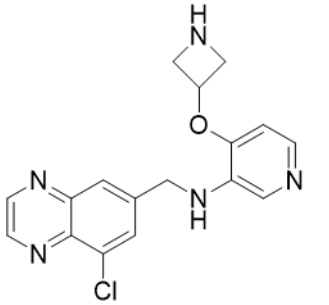
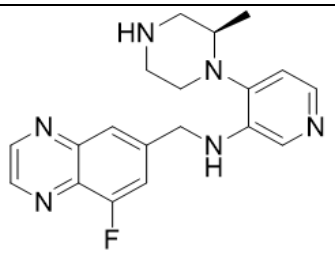
024	
025	
026	
027	
028	
029	
030	

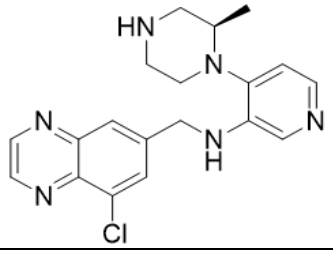
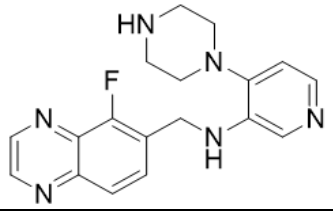
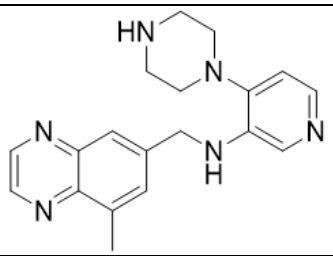
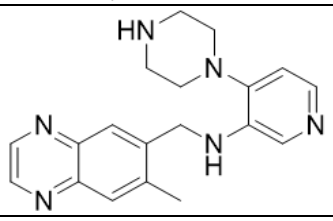
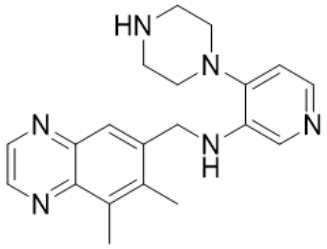
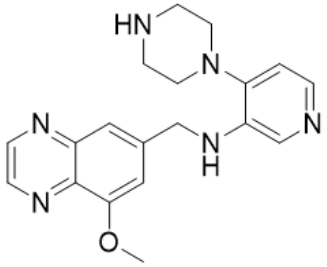
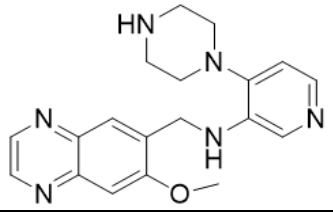
031	
032	
033	
034	
035	
036	
037	

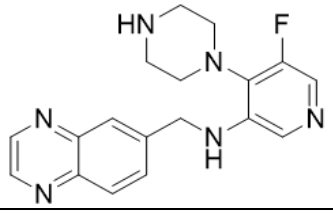
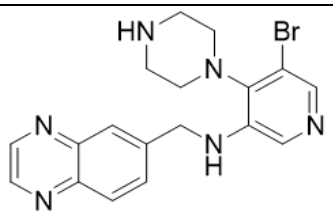
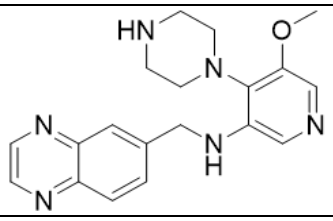
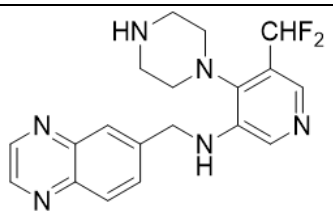
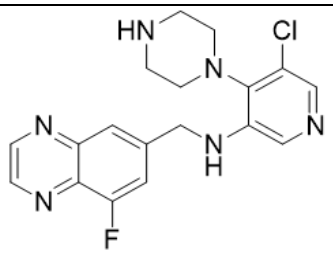
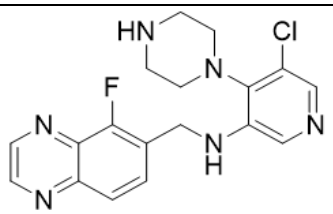
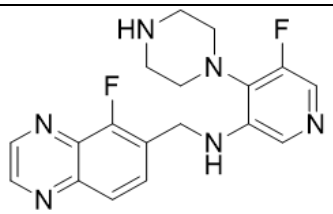
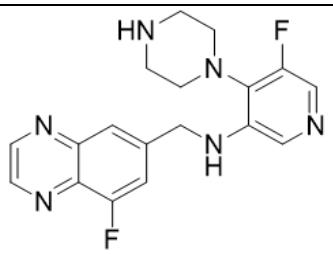
038	
039	
040	
041	
042	
043	
044	

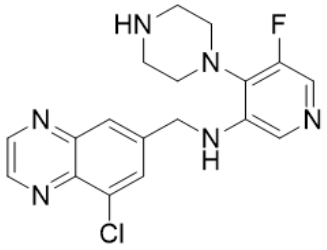
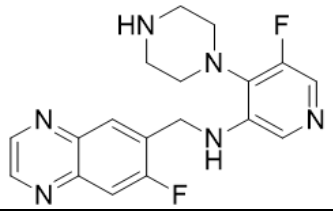
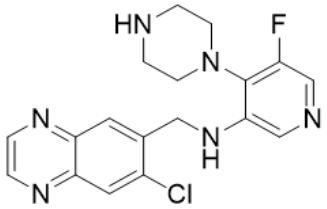
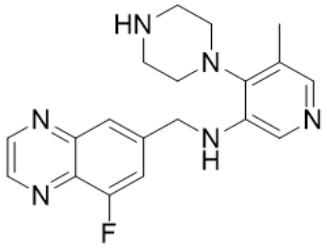
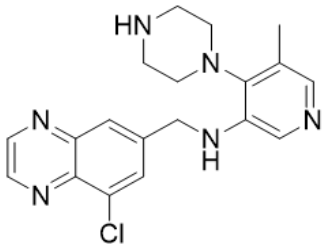
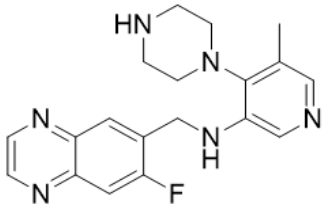
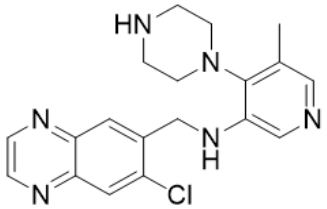
045	
046	
047	
048	
049	
050	
051	

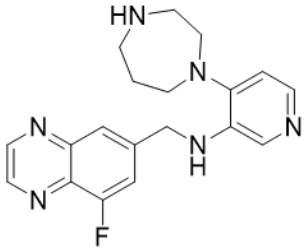
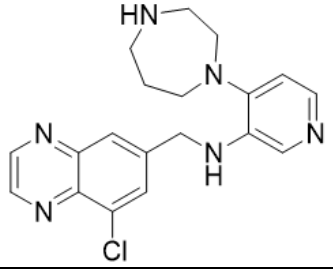
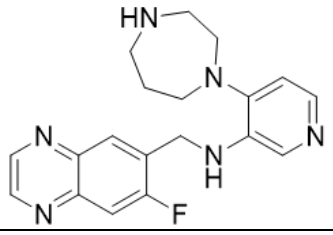
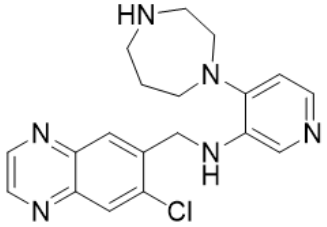
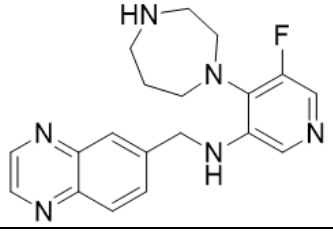
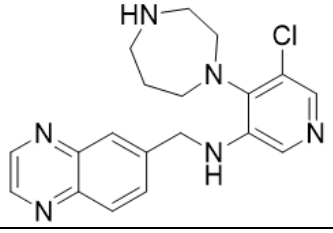
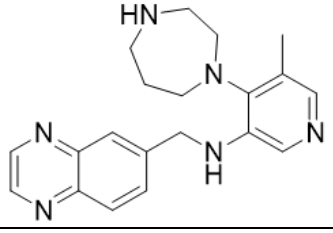
052	
053	
054	
055	
056	
057	
058	

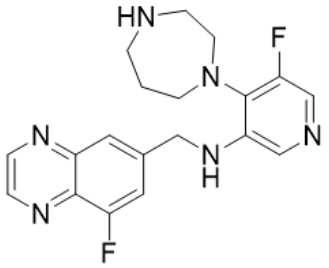
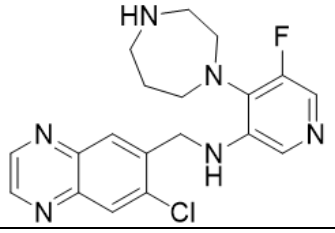
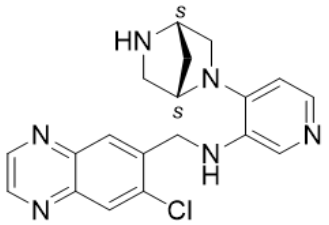
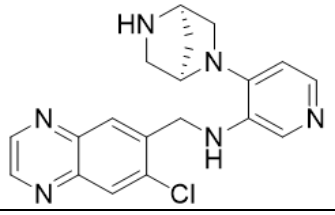
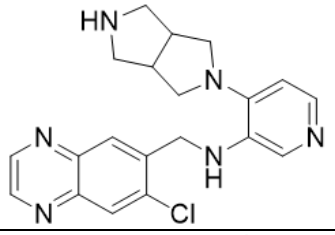
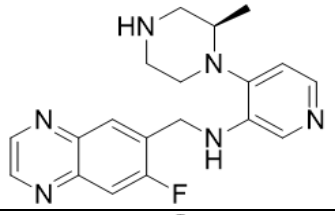
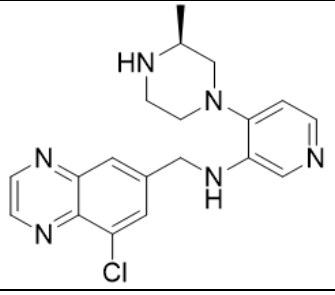
059	
060	
061	
062	
063	
064	
065	

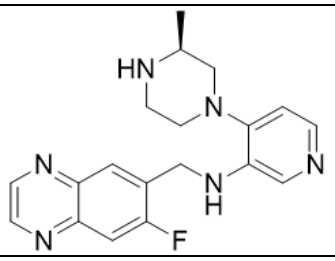
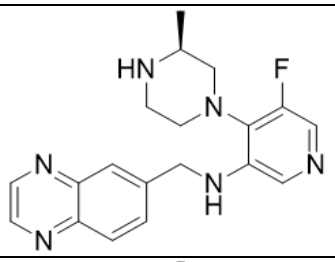
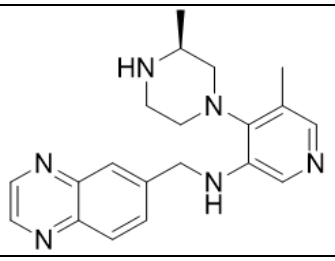
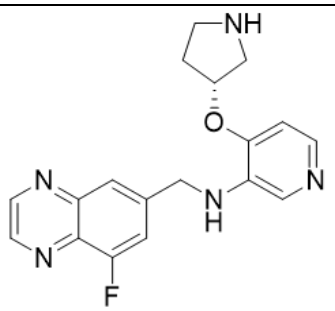
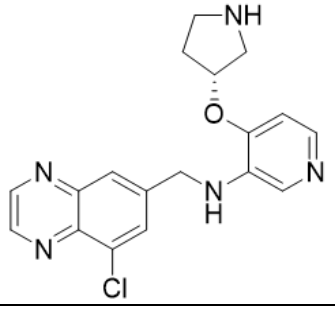
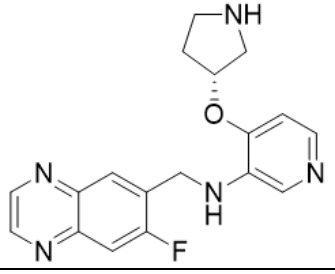
066	
067	
068	
069	
070	
071	
072	

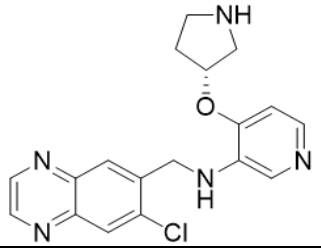
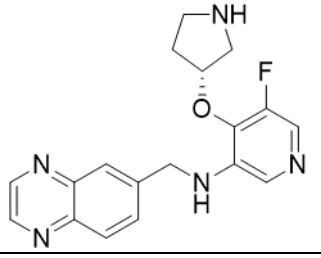
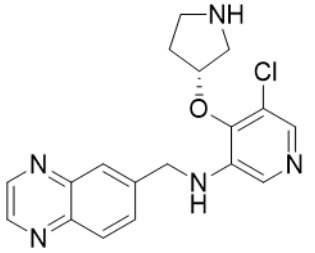
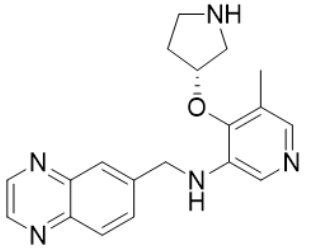
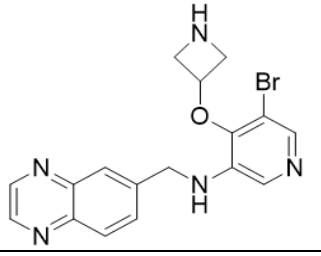
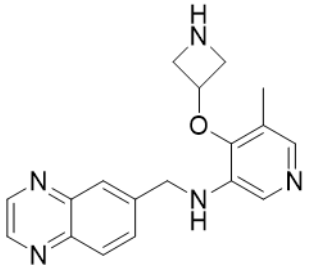
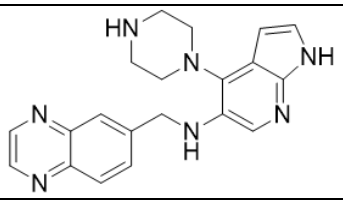
073	
074	
075	
076	
077	
078	
079	
080	

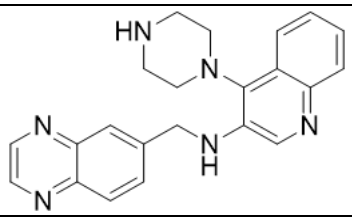
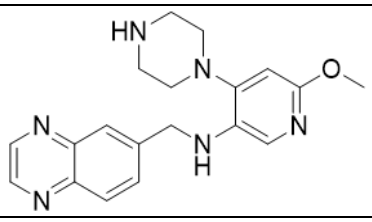
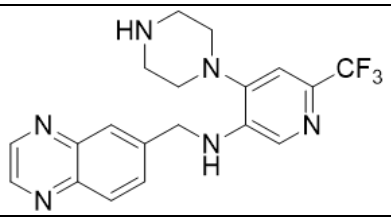
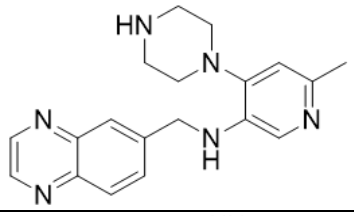
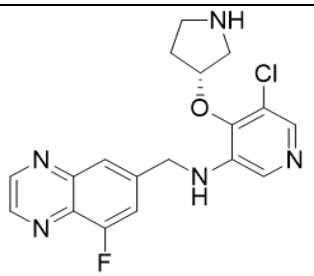
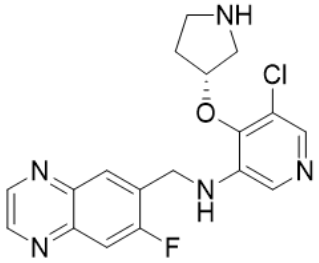
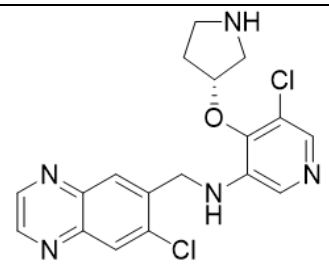
081	
082	
083	
084	
085	
086	
087	

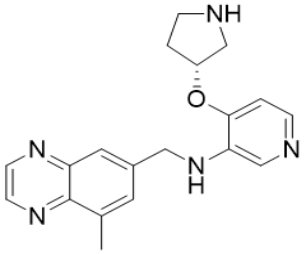
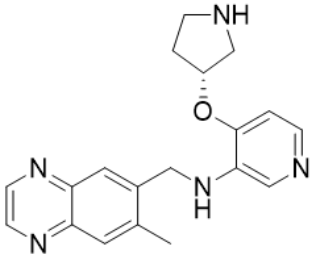
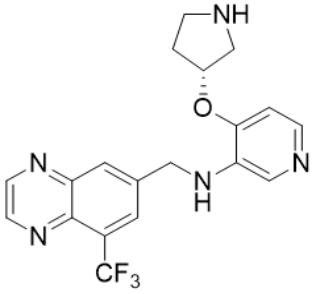
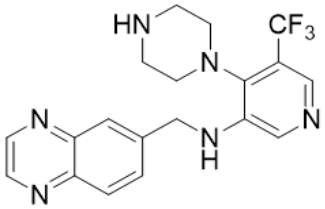
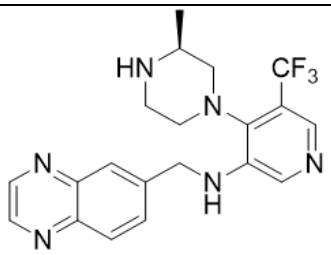
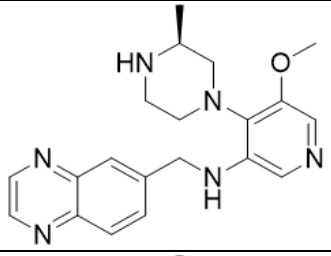
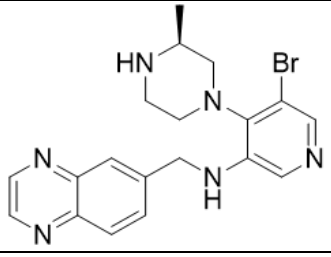
088	
089	
090	
091	
092	
093	
094	

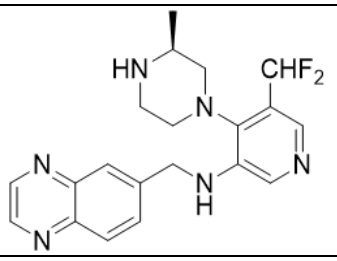
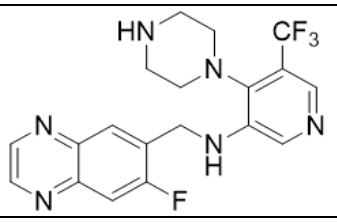
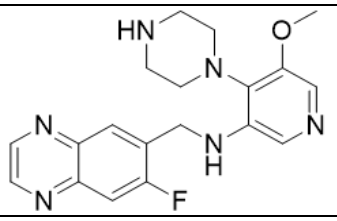
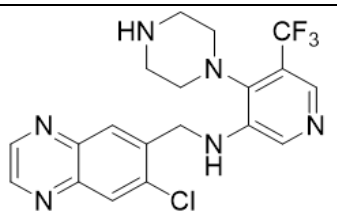
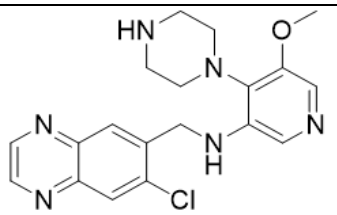
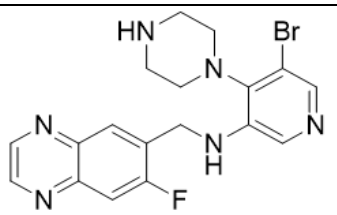
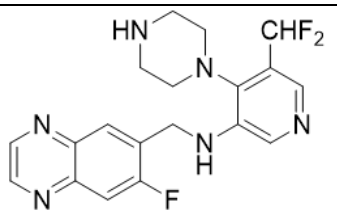
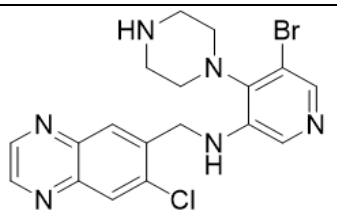
095	
096	
097	
098	
099	
100	
101	

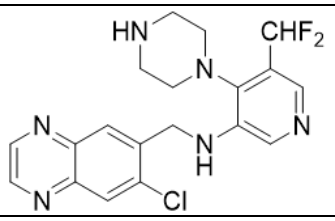
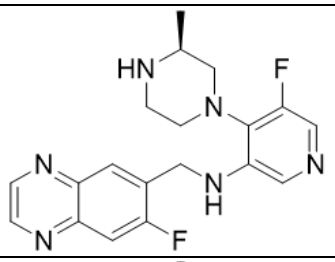
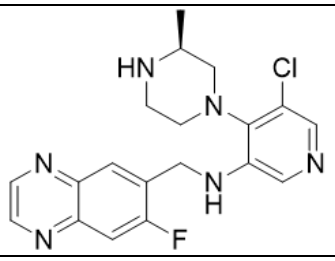
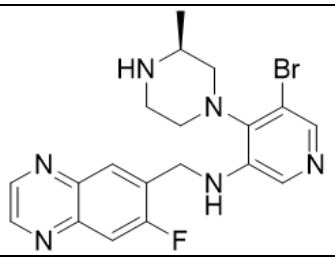
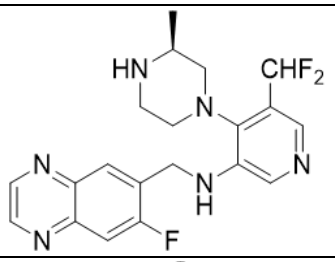
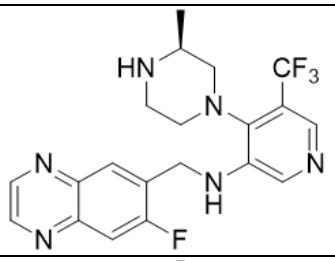
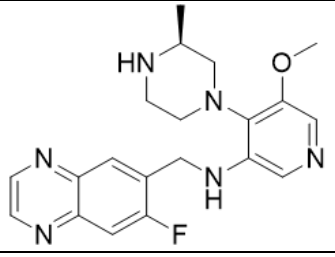
102	
103	
104	
105	
106	
107	

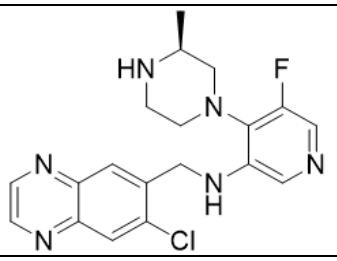
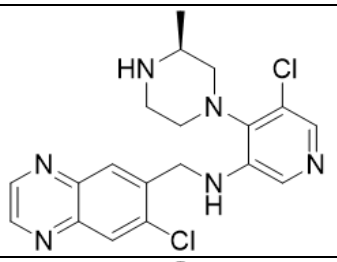
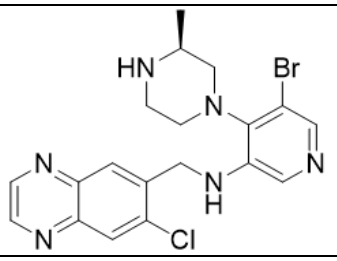
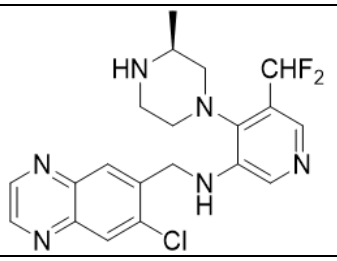
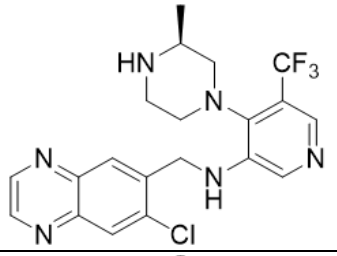
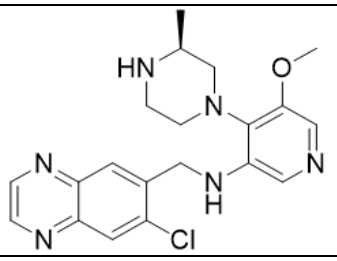
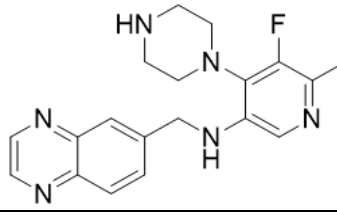
108	
109	
110	
111	
112	
113	
114	

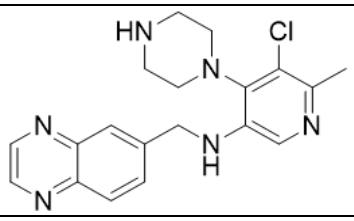
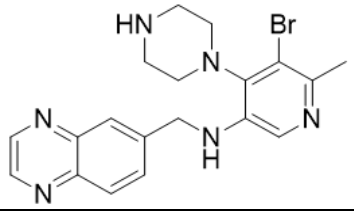
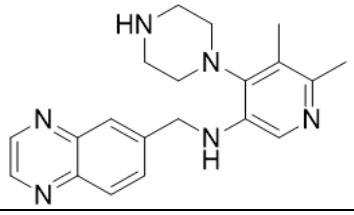
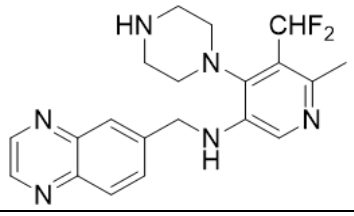
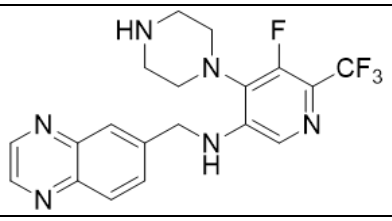
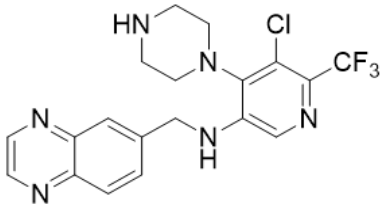
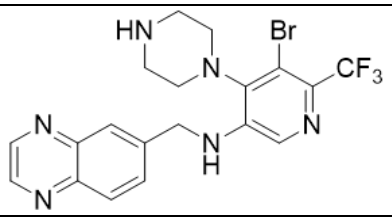
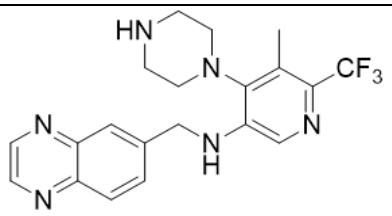
115	
116	
117	
118	
119	
120	
121	

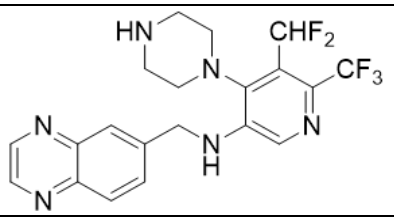
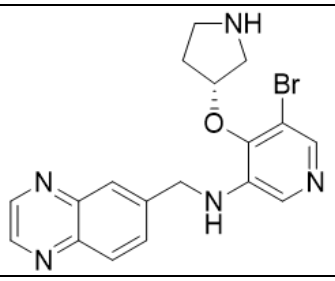
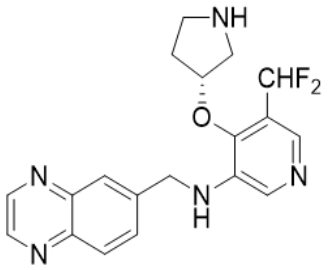
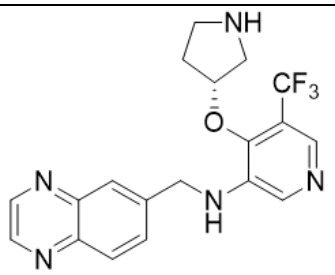
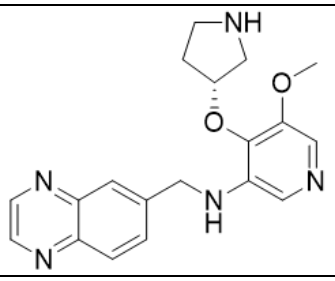
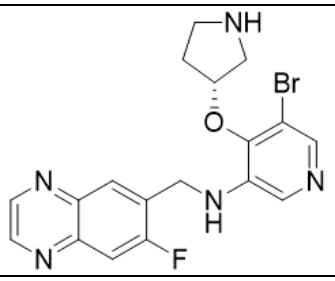
122	
123	
124	
125	
126	
127	
128	

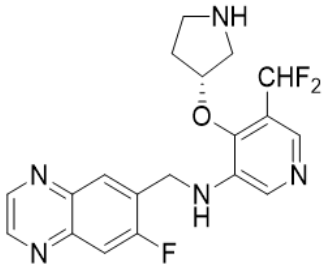
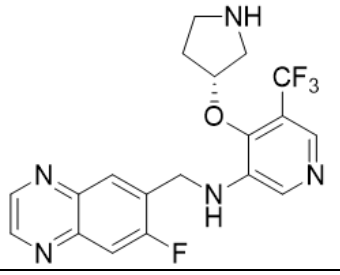
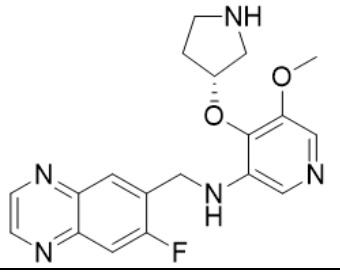
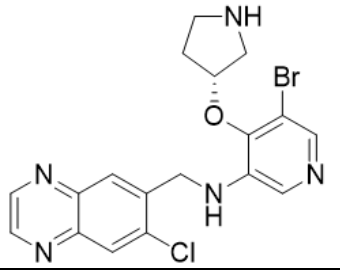
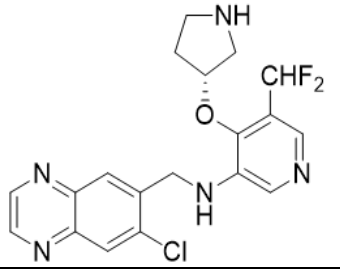
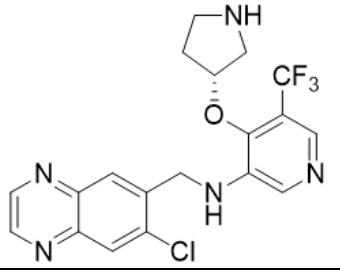
129	
130	
131	
132	
133	
134	
135	
136	

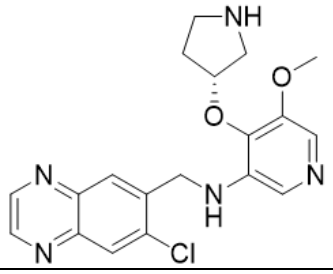
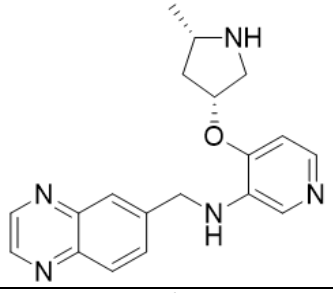
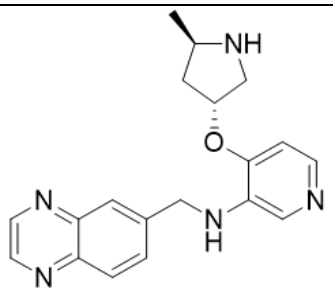
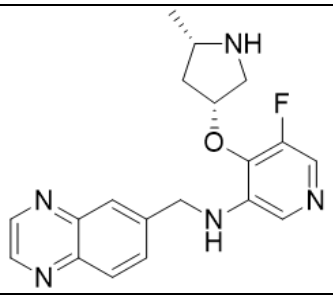
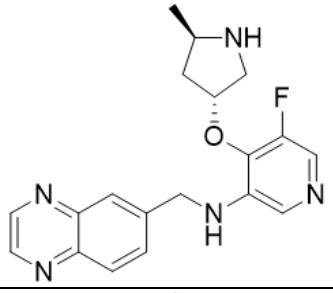
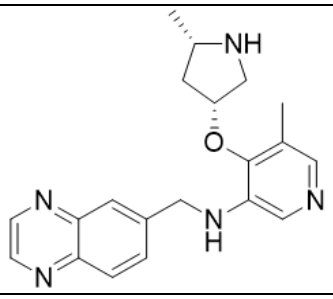
137	
138	
139	
140	
141	
142	
143	

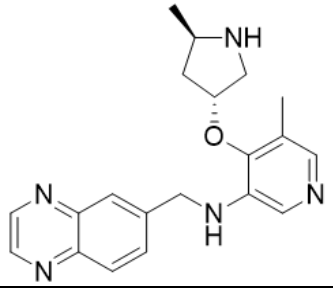
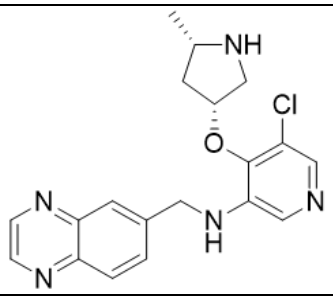
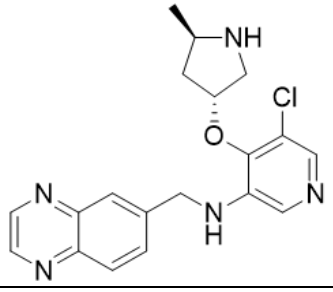
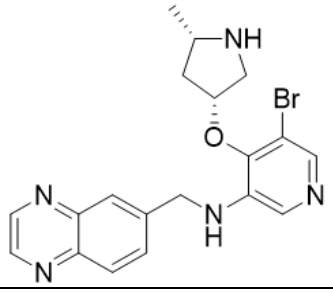
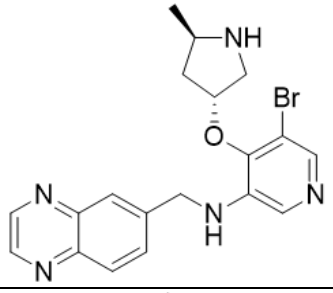
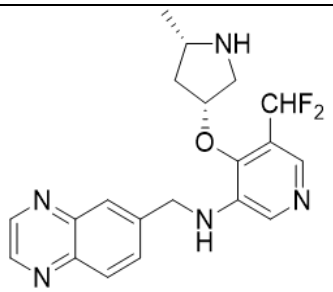
144	
145	
146	
147	
148	
149	
150	

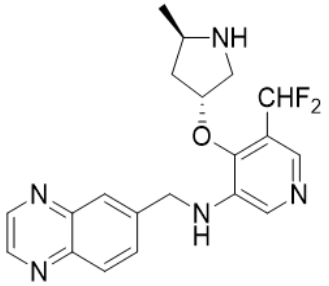
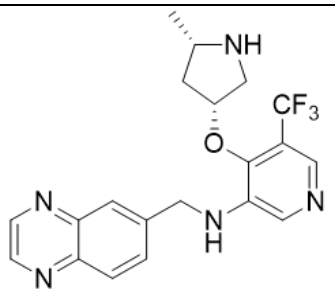
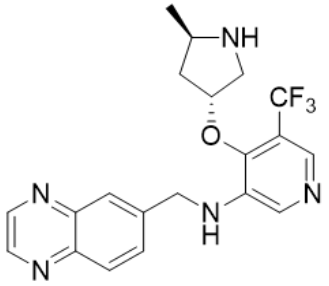
151	
152	
153	
154	
155	
156	
157	
158	

159	
160	
161	
162	
163	
164	

165	
166	
167	
168	
169	
170	

171	
172	
173	
174	
175	
176	

177	
178	
179	
180	
181	
182	

183	
184	
185	

[0041] The disclosure also provides methods for modulating the expression of a target gene using a polynucleotide cassette, an aptamer encoding sequence/aptamer sequence, or riboswitch disclosed herein, by provided to a cell or tissue a small molecule of Formula I-XXII, including, e.g., a small molecule provided in Table A.

BRIEF DESCRIPTION OF THE FIGURES

[0042] **Figures 1a-1d**. TPP aptamer homologous sequence regulates gene expression in response to TPP and fursultiamine. **Fig. 1a**: schematics of the synthetic riboswitch cassette containing intron-alternative exon-aptamer-intron. **Fig. 1b**: In the presence of aptamer ligand, aptamer ligand binding facilitates for the formation of hairpin stem that sequesters the accessibility of the 5' splice site (5' ss) of the alternative exon, resulting in the exclusion of the stop codon containing alternative exon and target gene expression. Riboswitch 12C6-1 regulates luciferase gene expression in response to TPP (**Fig. 1c**) and fursultiamine (**Fig. 1d**) treatment. HEK 293 cells were transfected with luciferase construct containing 12C6-1 riboswitch. The transfected cells were treated with TPP or fursultiamine at the indicated concentrations. The fold induction was calculated as the quotient of the luciferase activity

obtained from cells treated with compounds divided by the luciferase activity obtained from cells without compound treatment.

[0043] **Figures 2a-2d**. Comp. 004 activates TPP aptamer riboswitches in regulating luciferase expression in HEK 293 cells. Cells were transfected with the indicated riboswitch constructs and treated with various concentration of Comp. 004. Luciferase activity was measured 20 hours post compound treatment, and fold induction was calculated as the quotient of the luciferase activity obtained from cells treated with compounds divided by the luciferase activity obtained from cells without treatment. Riboswitch containing E.Coli thiM or Alishewanella tabrizica thiC aptamer (**Fig. 2a, 2b**) or 12C6-1 aptamer (**Fig. 2c, 2d**) regulates luciferase expression in response to Comp. 004 treatment.

We claim:

1. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCCATTGCACCTGATCCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGGCGCAGGGAG (SEQ ID NO:6); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:6 is any nucleotide or no nucleotide;
X_{16} is any nucleotide;
X_{17} is any nucleotide;
X_{18} is any nucleotide;
X_{19} is any nucleotide;
X_{20} is any nucleotide; and
X_{21} is C, G, T.

2. The polynucleotide cassette of claim 1, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:6 are C and G, respectively.

3. The polynucleotide cassette of claim 1 or 2, wherein the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCCATTGCACCTGATCCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGGCGCAGGGAG (SEQ ID NO:6); wherein:
X_{16} is G or T;
X_{17} is A or T;
X_{18} is any nucleotide;
X_{19} is A or G;
X_{20} is A, G, T; and
X_{21} is C, G, T.

4. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCC$X_7X_8X_9X_{10}X_{11}X_{12}CCTX_{13}X_{14}X_{15}$CCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGG$X_{22}X_{23}$C AGGGAG (SEQ ID NO:2); wherein:

the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:2 is any nucleotide or no nucleotide;

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is G or T;

X_5 is A, G, or T;

X_6 is A or G;

X_7 is A or T;

X_8 is A, C, or T;

X_9 is A, C, or T;

X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide;

X_{12} is A;

X_{13} is A, C, or G;

X_{14} is any nucleotide;

X_{15} is C, G, or T;

X_{16} is G or T;

X_{17} is A or T;

X_{18} is any nucleotide;

X_{19} is A or G;

X_{20} is A, G, T;

X_{21} is C, G, T;

X_{22} is T; and

X_{23} is A, G, or T.

5. The polynucleotide cassette of claim 4, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:2 are C and G, respectively.

6. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:4 is any nucleotide or no nucleotide;
X$_7$ is A, G, or T;
X$_8$ is any nucleotide;
X$_9$ is any nucleotide;
X$_{10}$ is any nucleotide;
X$_{11}$ is any nucleotide or no nucleotide;
X$_{12}$ is A, C, or T.

7. The polynucleotide cassette of claim 6, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:4 are C and G, respectively.

8. The polynucleotide cassette of claim 6 or 7, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:
X$_7$ is A or T;
X$_8$ is A, C, or T;
X$_9$ is A, C, or T;
X$_{10}$ is any nucleotide;
X$_{11}$ is any nucleotide or no nucleotide; and
X$_{12}$ is A.

9. The polynucleotide cassette of claim 6 or 7, wherein the aptamer encoding sequence comprises

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:

X$_7$ is A;

X$_8$ is A, C, or T;

X$_9$ is A, C, or T;

X$_{10}$ is any nucleotide;

X$_{11}$ is any nucleotide or no nucleotide; and

X$_{12}$ is A.

10. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises: CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX$_1$X$_2$AX$_3$X$_4$X$_5$X$_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:

the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:3 is any nucleotide or no nucleotide;

X$_1$ is C, G, or T;

X$_2$ is any nucleotide;

X$_3$ is any nucleotide;

X$_4$ is any nucleotide;

X$_5$ is any nucleotide; and

X$_6$ is any nucleotide.

11. The polynucleotide cassette of claim 10, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:3 are C and G, respectively.

12. The polynucleotide cassette of claim 10 or 11, wherein the aptamer encoding sequence comprises CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX$_1$X$_2$AX$_3$X$_4$X$_5$X$_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:

X$_1$ is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is any nucleotide;

X_5 is A, G, or T; and

X_6 is any nucleotide.

13. The polynucleotide cassette of claim 10 or 11, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCATCGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:
X_1 is C or T;
X_2 is any nucleotide;
X_3 is any nucleotide;
X_4 is G or T;
X_5 is A, G, or T; and
X_6 is A or G.

14. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCATTGCACCT$X_{13}X_{14}X_{15}$CCGGATCATGCCGG$X_{22}X_{23}$CAGGGAG (SEQ ID NO:5); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:5 is any nucleotide or no nucleotide;
X_{13}, X_{14}, X_{15}, X_{22}, and X_{23} is any nucleotide.

15. The polynucleotide cassette of claim 14, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:5 are C and G, respectively.

16. The polynucleotide cassette of claim 14 or 15, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCATTGCACCTX$_{13}$X$_{14}$X$_{15}$CCGGATCATGCCGGX$_{22}$X$_{23}$CAGGGAG (SEQ ID NO:5); wherein:

X$_{13}$ is A, C, or G;

X$_{14}$ is any nucleotide;

X$_{15}$ is C, G, or T;

X$_{22}$ is T; and

X$_{23}$ is A, G, or T.

17. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 1 and 7-558.

18. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 1 and 7-558.

19. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583.

20. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding is sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583.

21. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least

95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447.

22. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447.

23. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378.

24. A polynucleotide cassette for regulating the expression of a target gene, wherein the polynucleotide cassette comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378.

25. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCATTGCACCTGATCCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGGCGCAGGGAG (SEQ ID NO:6); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:6 is any nucleotide or no nucleotide;
X_{16} is any nucleotide;
X_{17} is any nucleotide;
X_{18} is any nucleotide;
X_{19} is any nucleotide;
X_{20} is any nucleotide; and
X_{21} is C, G, T;
wherein X_{16}-X_{21}, are not simultaneously A, T, C, A, T, and G, respectively.

26. The nucleic acid sequence of claim 25, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:6 are C and G, respectively.

27. The nucleic acid sequence of claim 25 or 26, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCCATTGCACCTGATCCGG $X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGGCGCAGGGAG (SEQ ID NO:6); wherein:
X_{16} is G or T;
X_{17} is A or T;
X_{18} is any nucleotide;
X_{19} is A or G;
X_{20} is A, G, T; and
X_{21} is C, G, T.

28. A nucleic acid sequence encoding an aptamer, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCATCGACCC$X_7X_8X_9X_{10}X_{11}X_{12}CCTX_{13}X_{14}X_{15}$CCGG$X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}$CCGG$X_{22}X_{23}$CAGGGAG (SEQ ID NO:2); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:2 is any nucleotide or no nucleotide;
X_1 is C or T;
X_2 is any nucleotide;
X_3 is any nucleotide;
X_4 is G or T;
X_5 is A, G, or T;
X_6 is A or G;
X_7 is A;
X_8 is A, C, or T;
X_9 is A, C, or T;
X_{10} is any nucleotide;

X_{11} is any nucleotide or no nucleotide;

X_{12} is A;

X_{13} is A, C, or G;

X_{14} is any nucleotide;

X_{15} is C, G, or T;

X_{16} is G or T;

X_{17} is A or T;

X_{18} is any nucleotide;

X_{19} is A or G;

X_{20} is A, G, T;

X_{21} is C, G, T;

X_{22} is T; and

X_{23} is A, G, or T.

29. The nucleic acid sequence of claim 28, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:2 are C and G, respectively.

30. A nucleic acid sequence encoding an aptamer, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCC$X_7X_8X_9X_{10}X_{11}X_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:4 is any nucleotide or no nucleotide;
X_7 is A, G, or T;
X_8 is any nucleotide;
X_9 is any nucleotide;
X_{10} is any nucleotide;
X_{11} is any nucleotide or no nucleotide;
X_{12} is A, C, or T;
wherein X_7-X_{12} are not simultaneously A, T, T, G, C, and A, respectively.

31. The nucleic acid sequence of claim 30, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:4 are C and G, respectively.

32. The nucleic acid sequence of claim 30 or 31, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:
X$_7$ is A or T;
X$_8$ is A, C, or T;
X$_9$ is A, C, or T;
X$_{10}$ is any nucleotide;
X$_{11}$ is any nucleotide or no nucleotide; and
X$_{12}$ is A;
wherein X$_7$-X$_{12}$ are not simultaneously A, T, T, G, C, and A, respectively.

33. The nucleic acid sequence of claim 30 or 31, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGA CCCX$_7$X$_8$X$_9$X$_{10}$X$_{11}$X$_{12}$CCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:4); wherein:
X$_7$ is A;
X$_8$ is A, C, or T;
X$_9$ is A, C, or T;
X$_{10}$ is any nucleotide;
X$_{11}$ is any nucleotide or no nucleotide; and
X$_{12}$ is A;
wherein X$_7$-X$_{12}$ are not simultaneously A, T, T, G, C, and A, respectively.

34. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGX$_1$X$_2$AX$_3$X$_4$X$_5$X$_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:3 is any nucleotide or no nucleotide;

X_1 is C G, or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is any nucleotide;

X_5 is any nucleotide; and

X_6 is any nucleotide;

wherein X_1-X_6 are not simultaneously C, A, T, C, G, and A, respectively.

35. The nucleic acid sequence of claim 34, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:3 are C and G, respectively.

36. The nucleic acid sequence of claim 34 or 35, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3);

wherein:

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is any nucleotide;

X_5 is A, G, or T; and

X_6 is any nucleotide;

wherein X_1-X_6 are not simultaneously C, A, T, C, G, and A, respectively.

37. The nucleic acid sequence of claim 34 or 35, wherein the aptamer encoding sequence comprises:

CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAG$X_1X_2AX_3X_4X_5X_6$CCAT CGACCCATTGCACCTGATCCGGATCATGCCGGCGCAGGGAG (SEQ ID NO:3);

wherein:

X_1 is C or T;

X_2 is any nucleotide;

X_3 is any nucleotide;

X_4 is G or T;

X_5 is A, G, or T;

X_6 is A or G.

38. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCATTGCACCTX$_{13}$X$_{14}$X$_{15}$CCGGATCATGCCGGX$_{22}$X$_{23}$CAGGGAG (SEQ ID NO:5); wherein:
the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:5 is any nucleotide or no nucleotide;
X$_{13}$, X$_{14}$, X$_{15}$, X$_{22}$, and X$_{23}$ is any nucleotide,
wherein X$_{13}$, X$_{14}$, X$_{15}$, X$_{22}$, and X$_{23}$ are not simultaneously G, A, T, C, and G, respectively.

39. The nucleic acid sequence of claim 38, wherein the first and the last nucleotide of the aptamer encoding sequence of SEQ ID NO:5 are C and G, respectively.

40. The nucleic acid sequence of claim 38 or 39, wherein the aptamer encoding sequence comprises:
CTGGGGAGTCCTTCATGCGGGGCTGAGAGGATGGAAGCAATCGACCATCGACCCATTGCACCTX$_{13}$X$_{14}$X$_{15}$CCGGATCATGCCGGX$_{22}$X$_{23}$CAGGGAG (SEQ ID NO:5); wherein:
X$_{13}$ is A, C, or G;
X$_{14}$ is any nucleotide;
X$_{15}$ is C, G, or T;
X$_{22}$ is T; and
X$_{23}$ is A, G, or T.

41. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 1 and 7-558.

42. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to

a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 1 and 7-558.

43. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583.

44. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding is sequence selected from the group consisting of SEQ ID NOs: 7-17, 89-96, 174-349, and 358-583.

45. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447.

46. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 7-11, 89-94, 174-349, and 358-447.

47. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence that is at least 95% identical to a sequence selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378.

48. A nucleic acid sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence comprises a sequence encoding an aptamer that binds to a small molecule, wherein the aptamer encoding sequence is selected from the group consisting of SEQ ID NOs: 174, 358, 363, and 378.

49. A nucleic acid sequence encoding a recombinant riboswitch for the regulation of target gene expression in response to a small molecule, wherein the riboswitch comprises an aptamer encoded by SEQ ID NOs: 1 and 7-558 or a sequence that is at least 95% identical to SEQ ID NOs: 1 and 7-558.

50. A polynucleotide cassette for the regulation of the expression of a target gene in response to a small molecule, the polynucleotide cassette comprising:

(a) a riboswitch; and

(b) an alternatively-spliced exon, flanked by a 5' intron and a 3' intron,

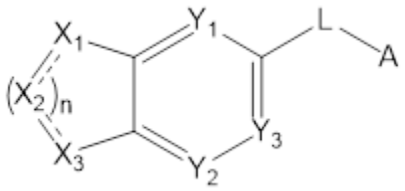
wherein the riboswitch comprises (i) an effector region comprising a stem that includes the 5' splice site sequence of the 3' intron, and (ii) the aptamer comprises a sequence of SEQ ID NOs: 1 and 7-558 or a sequence that is at least 95% identical to SEQ ID NOs: 1 and 7-558; and

wherein the alternatively-spliced exon comprises a stop codon that is in-frame with the target gene when the alternatively-spliced exon is spliced into the target gene mRNA.

51. The polynucleotide cassette of claim 50, wherein the polynucleotide cassette is located in the protein coding sequence of the target gene.

52. The polynucleotide cassette of claim 50, wherein the polynucleotide cassette is located in an untranslated region of the target gene or in an intron of the target gene.

53. The polynucleotide cassette of any one of claims 1-24 and 50-52 or the nucleic acid sequence of any one of claims 25-49, wherein the aptamer binds to, or otherwise responds to the presence of, a small molecule having the structure according to Formula I:



I

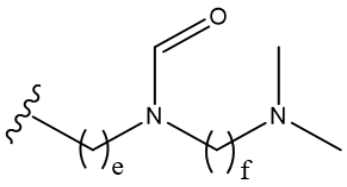
wherein

X_1, X_2, and X_3 are, in each instance, independently selected from CR_1, CHR_1, N, NH, O and S, wherein adjacent X_1, X_2, and X_3 are not simultaneously selected to be O or S;

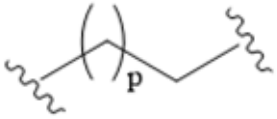
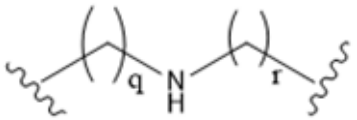
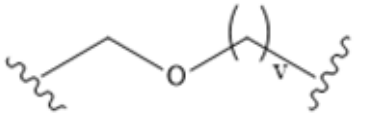
the dashed lines represent optional double bonds;

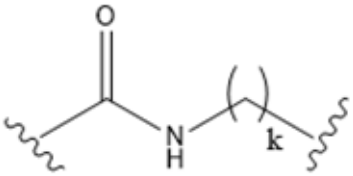
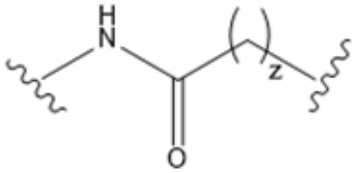
Y_1, Y_2, and Y_3 are, in each instance, independently selected from CR_2 and N;

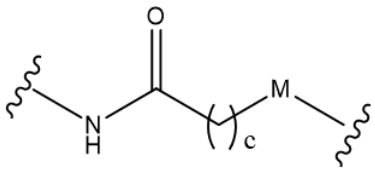
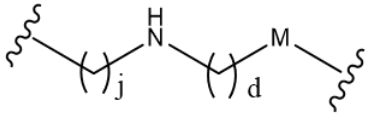
n is 1 or 2, wherein when n is 1, only one of the dashed lines is a double bond;

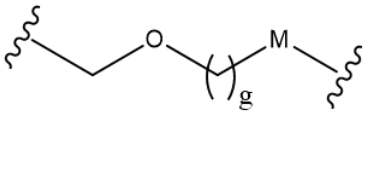
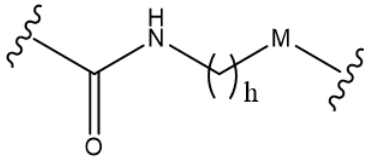
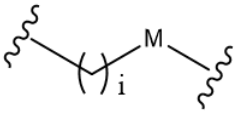
L-A is  , or

L is selected from

 ,  ,  ,

 ,  ,

 ,  ,

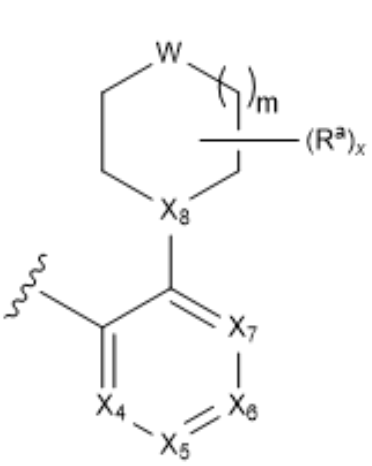
 ,  , and 

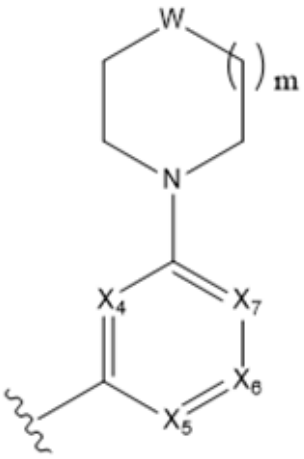
wherein k, p, q, r, and v are independently selected from integers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, z is selected from integers 1, 2, 3, 4, and 5;

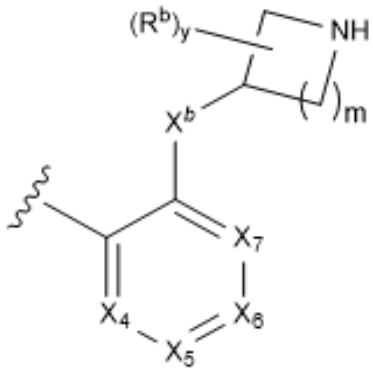
c, d, e, f, g, h and i are independently selected from integers 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10; j is selected from integers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

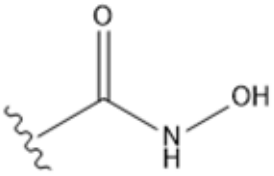
M is selected from –NH-, -O-, -NHC(=O)-, -C(=O)NH-, -S-, and -C(=O)-; and

A is selected from



,


,




and ,

wherein X_4, X_5, X_6, and X_7, are independently selected from CR_3 and N;

X_8 is N or CH;

X^b is selected from O, NH, and NCH_3;

wherein each of R_1, R_2, and R_3 are independently selected from -H, -Cl, -Br, -I, -F, $-CF_3$, $-CH_2F$, $-CHF_2$, -OH, -CN, $-NO_2$, $-NH_2$, $-NH(C_1\text{-}C_6$ alkyl), $-N(C_1\text{-}C_6$ alkyl)$_2$, -COOH, $-COO(C_1\text{-}C_6$ alkyl), $-CO(C_1\text{-}C_6$ alkyl), $-O(C_1\text{-}C_6$ alkyl), $-OCO(C_1\text{-}C_6$ alkyl), $-NCO(C_1\text{-}C_6$ alkyl), $-CONH(C_1\text{-}C_6$ alkyl), and substituted or unsubstituted $C_1\text{-}C_6$ alkyl;

additionally or alternatively, two R_3 on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

m is 1 or 2;

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group, or two R^a attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

x is 0, 1, 2 or 3;

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6-

membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

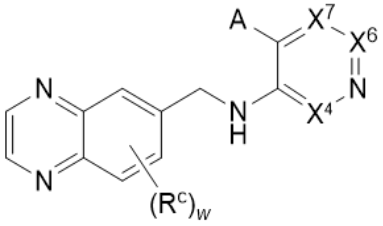
y is 0, 1, 2 or 3; and

W is O or NR_4, wherein R_4 is selected from selected from -H, -CO(C_1-C_6 alkyl), substituted or unsubstituted C_1-C_6 alkyl, substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, substituted or unsubstituted cycloalkyl, -CO(aryl), -CO(heteroaryl), and -CO(cycloalkyl);

provided that at least two of X_1, X_2, X_3, X_4, X_5, X_6, and X_7 are N;

or a pharmaceutically acceptable salt thereof.

54. The polynucleotide cassette of claim 53, wherein the small molecule has a structure according to formula XIII



XIII

or a pharmaceutically acceptable salt thereof, wherein

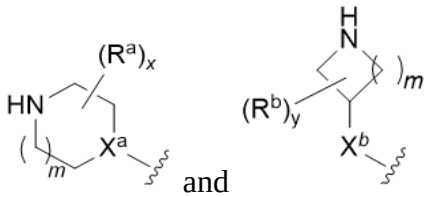
X^4 is selected from CH, CR^d and N;

X^6 is selected from CH, CR^d and N;

X^7 is selected from CH, CR^d and N;

wherein 0 or 1 of X^4, X^6 or X^7 is N;

A is selected from the group consisting of:



and

X^a is selected from N and CH;

X^b is selected from O, NH, and NCH_3;

each R^a is independently selected from C_1 to C_3 alkyl, -OCH_3, -CF_3, -CH_2F, -CHF_2, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

x is 0, 1, 2 or 3;

y is 0, 1, 2 or 3;

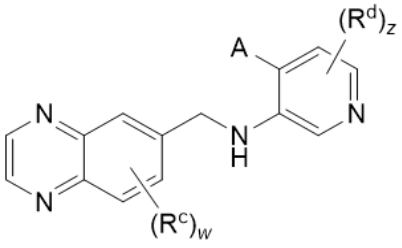
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH; and

w is 0, 1 or 2.

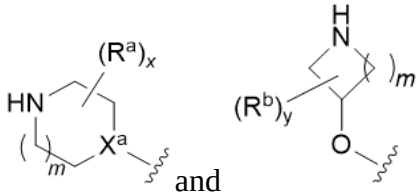
55. The polynucleotide cassette of claim 54, wherein the small molecule has a structure according to formula XIV



XIV

or a pharmaceutically acceptable salt thereof, wherein

A is selected from the group consisting of:



and

X^a is selected from N and CH;

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon

atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

x is 0, 1, 2 or 3;

y is 0, 1, 2 or 3;

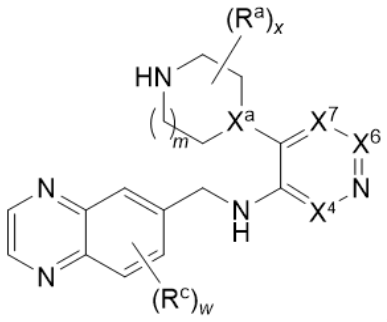
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino; alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

w is 0, 1 or 2; and

z is 0, 1 or 2.

56. The polynucleotide cassette of claim 54, wherein the small molecule has a structure according to formula XVI



XVI

or a pharmaceutically acceptable salt thereof, wherein

X^4 is selected from CH, CR^d and N;

X^6 is selected from CH, CR^d and N;

X^7 is selected from CH, CR^d and N;

wherein 0 or 1 of X^4, X^6 or X^7 is N;

X^a is selected from N and CH;

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

x is 0, 1, 2 or 3;

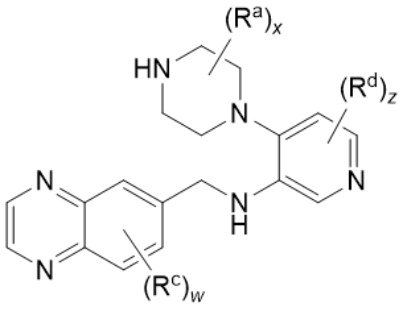
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH; and

w is 0, 1 or 2.

57. The polynucleotide cassette of claim 54, wherein the small molecule has a structure according to formula XVII



XVII

or a pharmaceutically acceptable salt thereof, wherein

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

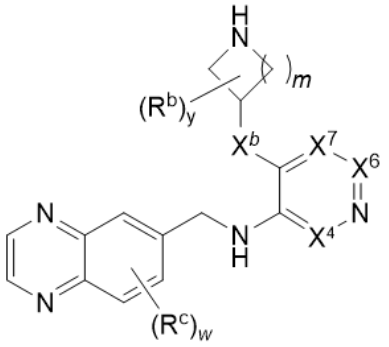
alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

w is 0, 1 or 2;

x is 0, 1, 2 or 3; and

z is 0, 1 or 2.

58. The polynucleotide cassette of claim 54, wherein the small molecule has a structure according to formula XX



XX

or a pharmaceutically acceptable salt thereof, wherein

X^4 is selected from CH, CR^d and N;

X^6 is selected from CH, CR^d and N;

X^7 is selected from CH, CR^d and N;

wherein 0 or 1 of X^4, X^6 or X^7 is N;

X^b is selected from O, NH, and NCH_3;

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

y is 0, 1, 2 or 3;

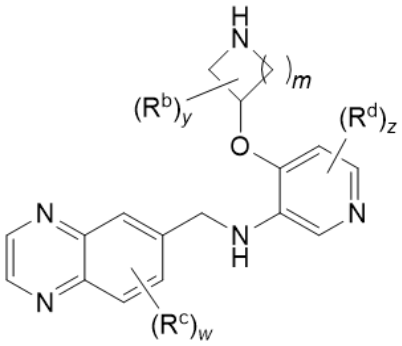
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH; and

w is 0, 1 or 2.

59. The polynucleotide cassette of claim 54, wherein the small molecule has a structure according to formula XXI



XXI

or a pharmaceutically acceptable salt thereof, wherein

each R^b is independently selected from C_1 to C_3 alkyl, -OCH_3, -CF_3, -CH_2F, -CHF_2, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^c is independently selected from halo, C_1 to C_3 alkyl, -OCH_3, -CF_3, -CH_2F, -CHF_2, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, -OCH_3, -CF_3, -CH_2F, -CHF_2, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

m is 1 or 2;

w is 0, 1 or 2

y is 0, 1 or 2; and

z is 0, 1 or 2.

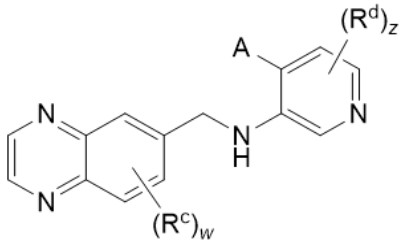
60. A vector comprising the polynucleotide cassette of any one of claims 1-24 and 50-59, or the nucleic acid sequence of any one of claims 25-49.

61. The vector of claim 60, wherein the vector is a viral vector.

62. The vector of claim 61, wherein the viral vector is selected from the group consisting of an adenoviral vector, an adeno-associated virus vector, and a lentiviral vector.

63. A cell comprising the vector of any one of claims 60-62, the polynucleotide cassette of any one of claims 1-24 and 50-59, or the nucleic acid sequence of any one of claims 25-49.

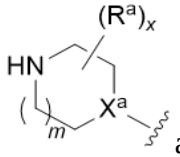
64. A compound having the structure according to formula XIV



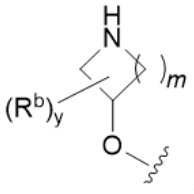
XIV

or a pharmaceutically acceptable salt thereof, wherein

A is selected from the group consisting of:



and



;

X^a is selected from N and CH;

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

x is 0, 1, 2 or 3;

y is 0, 1, 2 or 3;

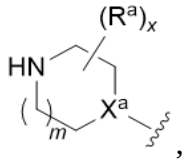
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino; alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

w is 0, 1 or 2; and

z is 0, 1 or 2;

wherein when A is selected to be

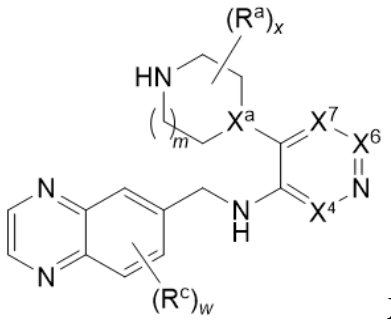


,

x is 1, 2 or 3; and/or

two R^d on adjacent ring positions are taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH.

65. The compound of claim 64, having the structure according to formula XVI



XVI

or a pharmaceutically acceptable salt thereof, wherein

X^4 is selected from CH, CR^d and N;

X^6 is selected from CH, CR^d and N;

X^7 is selected from CH, CR^d and N;

wherein 0 or 1 of X^4, X^6 or X^7 is N;

X^a is selected from N and CH;

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon

atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

x is 1, 2 or 3;

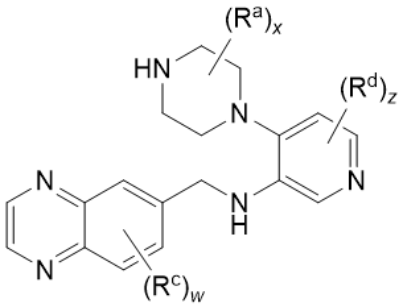
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH; and

w is 0, 1 or 2.

66. The compound of claim 64, having the structure according to formula XVII



XVII

or a pharmaceutically acceptable salt thereof, wherein

each R^a is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^a attached to the same carbon atom form an oxo group; or two R^a attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

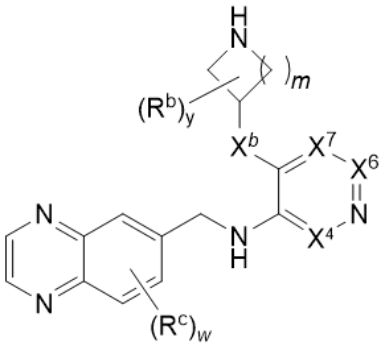
alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

w is 0, 1 or 2;

x is 1, 2 or 3; and

z is 0, 1 or 2.

67. The compound of claim 64, having the structure according to formula XX



XX

or a pharmaceutically acceptable salt thereof, wherein

X^4 is selected from CH, CR^d and N;

X^6 is selected from CH, CR^d and N;

X^7 is selected from CH, CR^d and N;

wherein 0 or 1 of X^4, X^6 or X^7 is N;

X^b is selected from O, NH, and NCH_3;

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6-membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

m is 1 or 2;

y is 0, 1, 2 or 3;

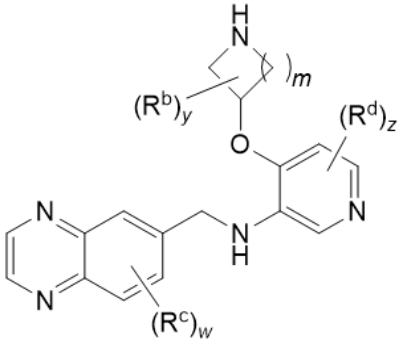
each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH; and

w is 0, 1 or 2.

68. The compound of claim 64, having the structure according to formula XXI



XXI

or a pharmaceutically acceptable salt thereof, wherein

each R^b is independently selected from C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, halo, hydroxyl and amino; or additionally or alternatively, two R^b attached to the same carbon atom form an oxo group; or two R^b attached to different carbon atoms form a 4- to 6- membered carbocyclic ring or a 4- to 6- membered heterocyclic ring having 1 or 2 heteroatoms selected from O and NH;

each R^c is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

each R^d is independently selected from halo, C_1 to C_3 alkyl, $-OCH_3$, $-CF_3$, $-CH_2F$, $-CHF_2$, -CN, hydroxyl and amino;

alternatively, two R^d on adjacent ring positions may be taken together to form a 5- or 6-membered aromatic ring having from 0 to 2 heteroatoms selected from O, S, N and NH;

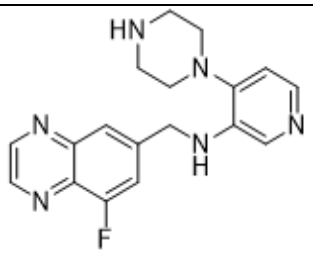
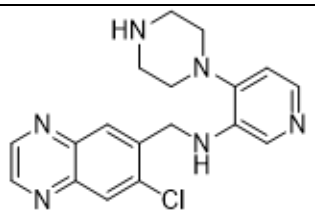
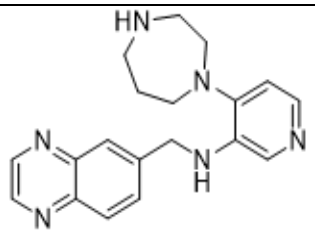
m is 1 or 2;

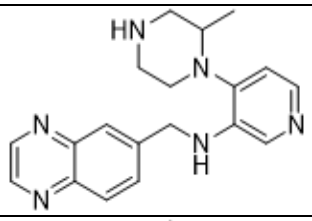
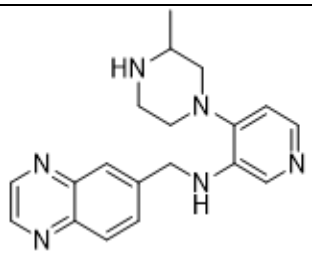
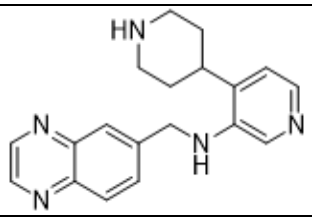
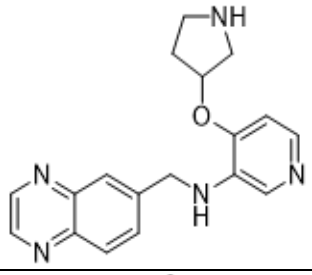
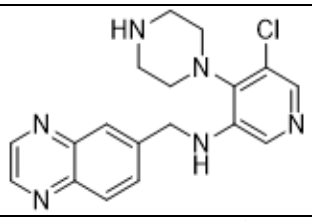
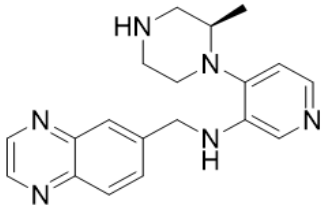
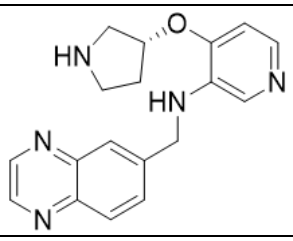
w is 0, 1 or 2

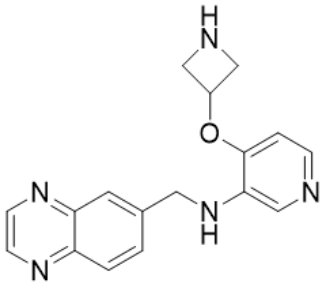
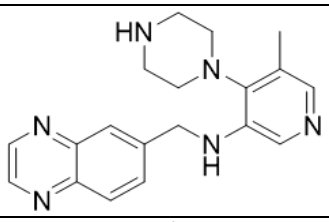
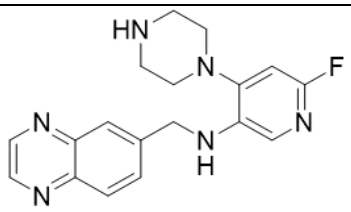
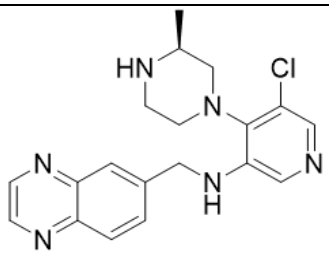
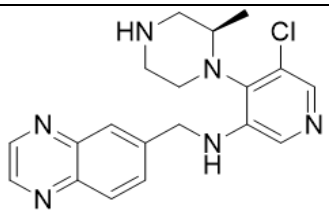
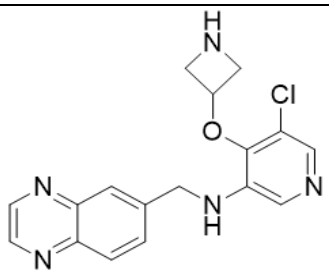
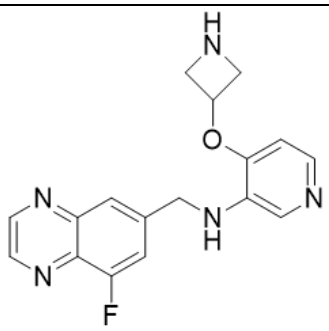
y is 0, 1 or 2; and

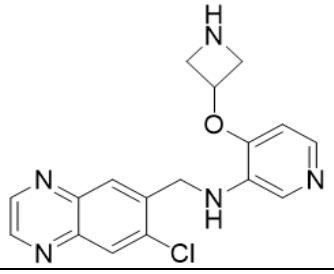
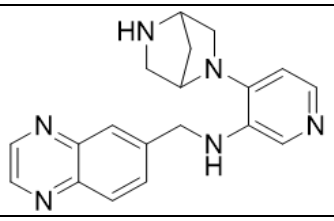
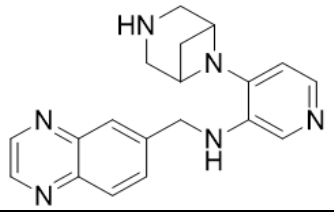
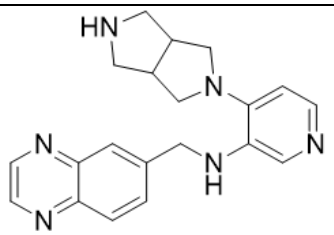
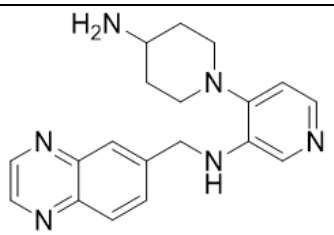
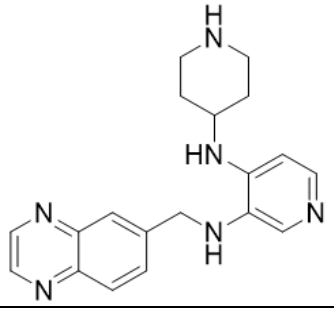
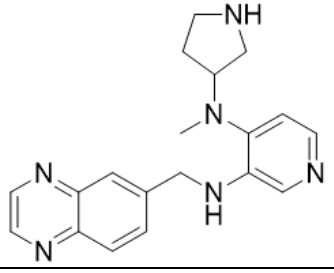
z is 0, 1 or 2.

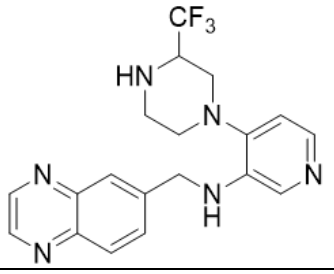
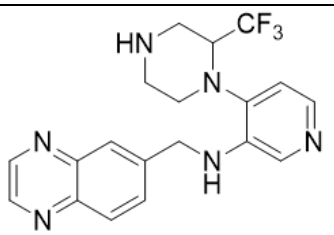
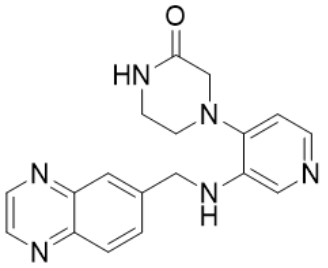
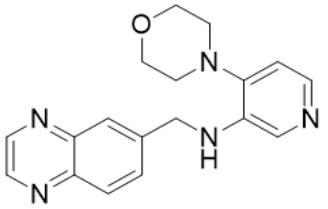
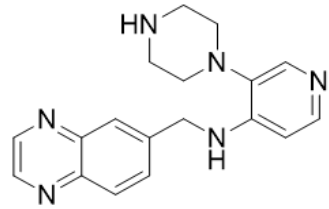
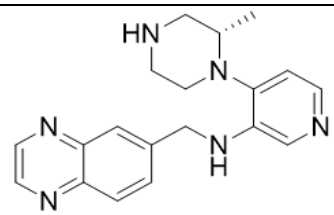
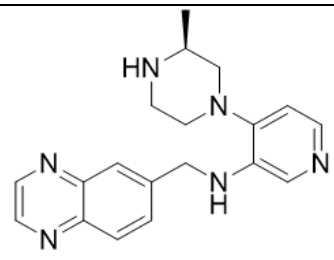
69. A compound having the structure according to one of:

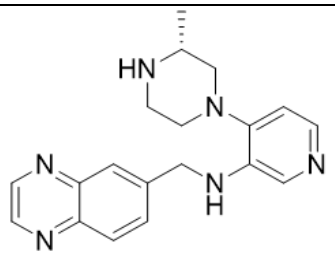
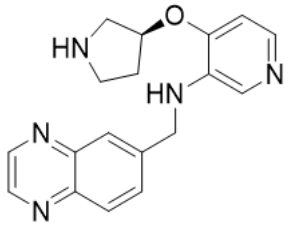
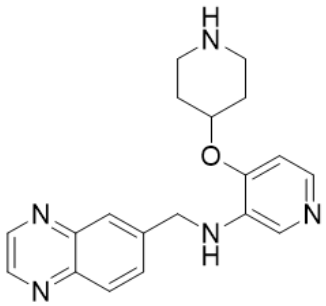
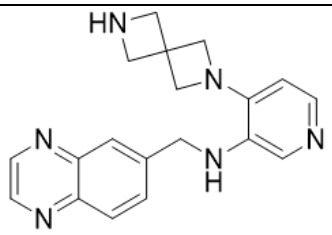
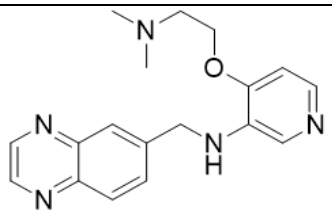
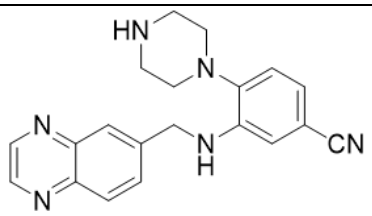
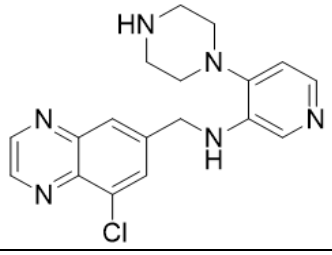
Ref. No.	**Structure**
012	
013	
014	

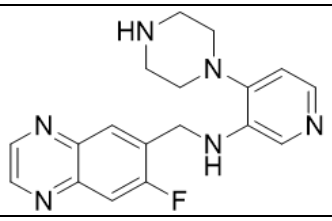
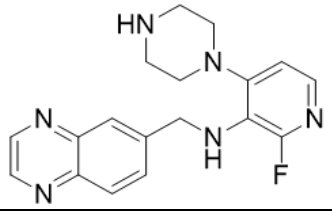
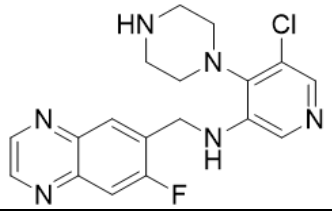
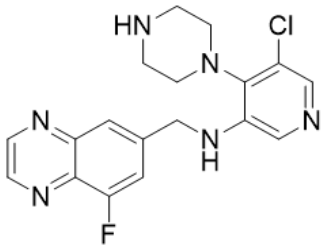
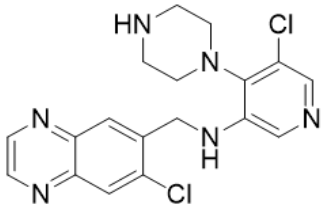
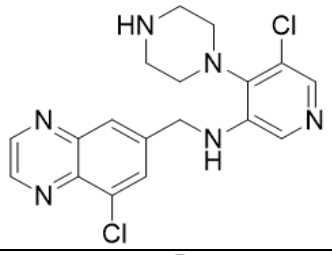
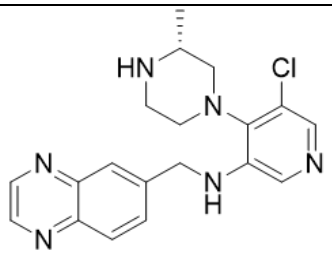
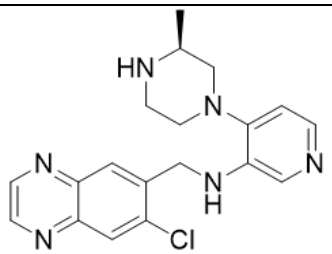
015	
016	
017	
018	
019	
020	
021	

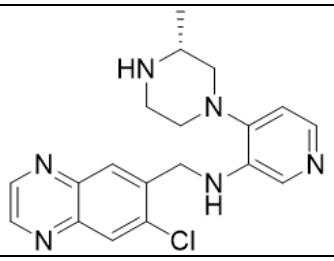
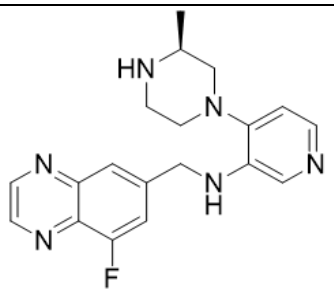
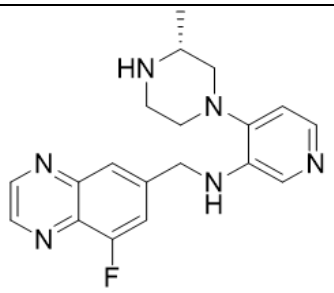
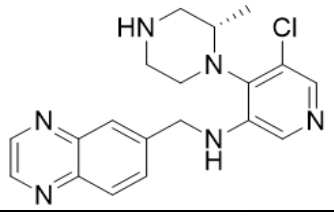
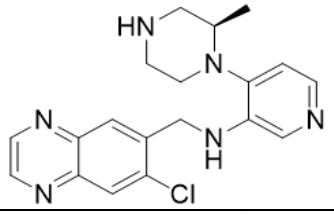
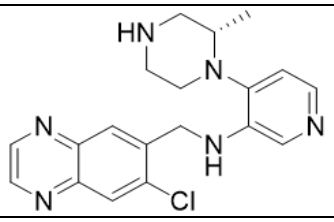
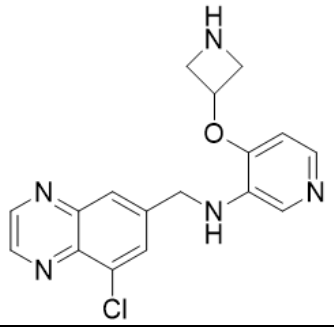
022	
023	
024	
025	
026	
027	
028	

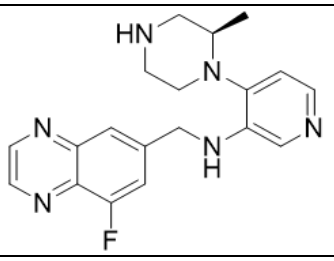
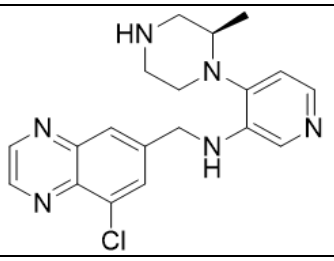
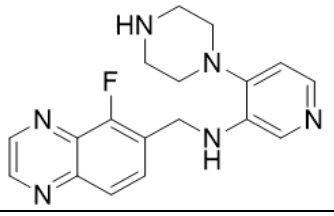
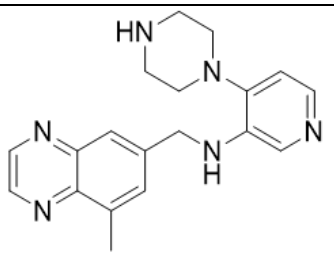
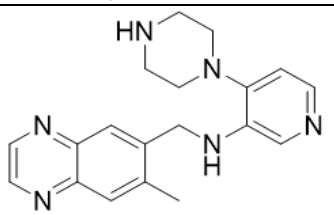
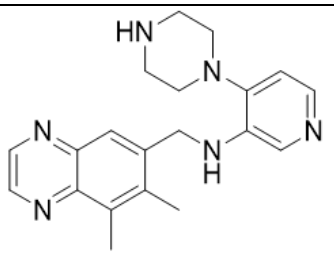
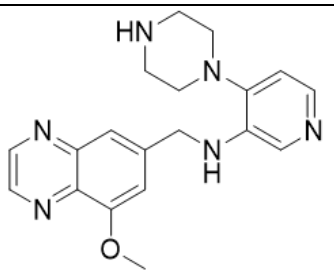
029	
030	
031	
032	
033	
034	
035	

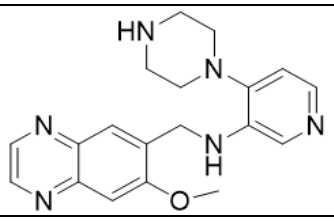
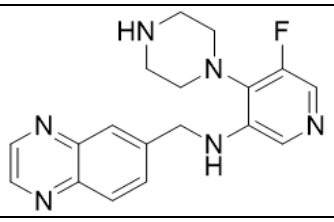
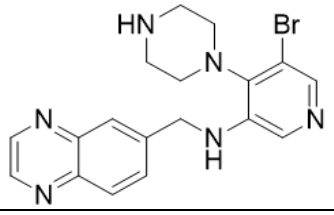
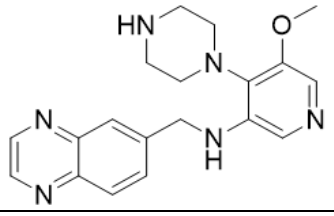
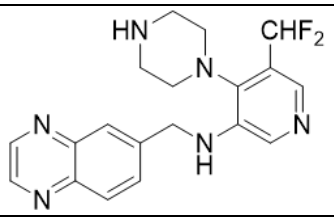
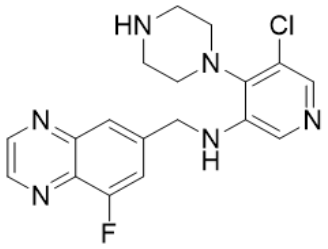
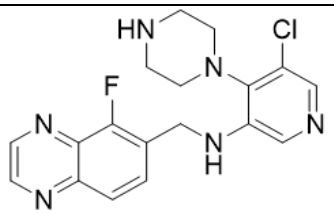
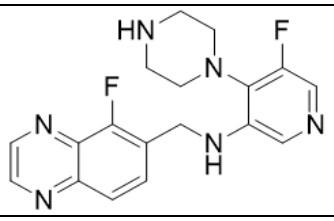
036	
037	
038	
039	
040	
041	
042	

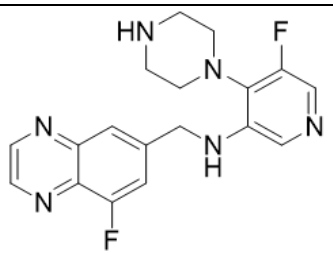
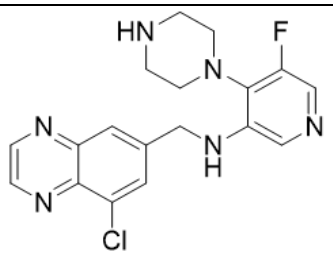
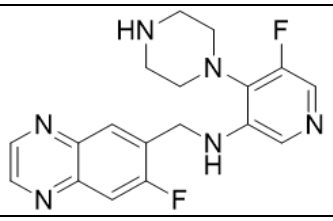
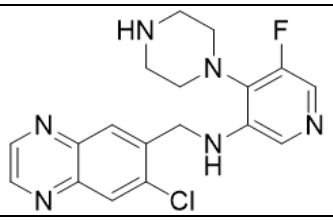
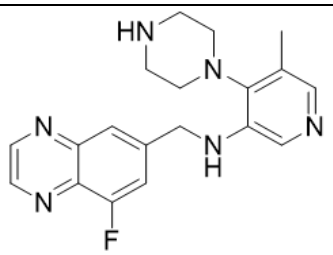
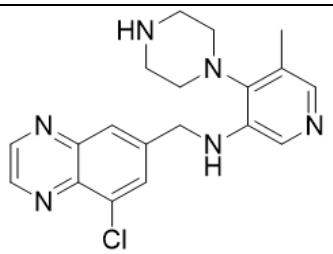
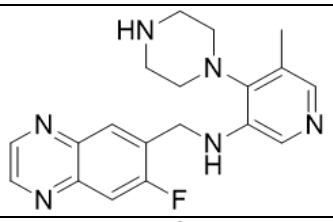
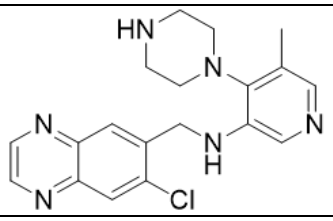
043	
044	
045	
046	
047	
048	
049	

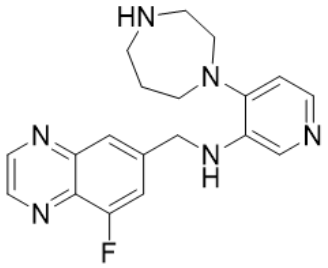
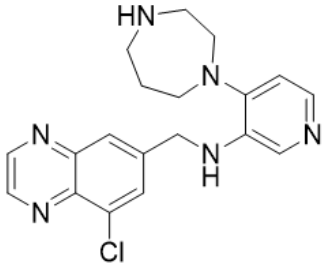
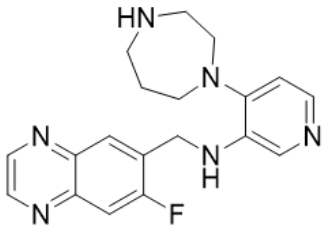
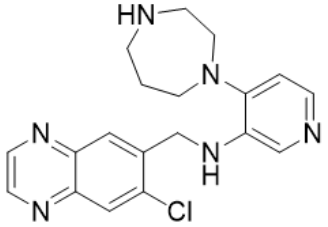
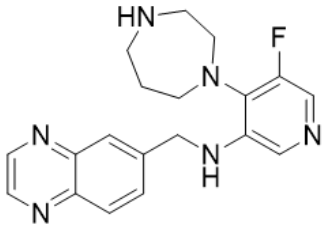
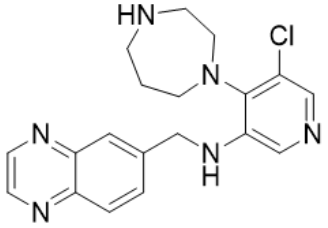
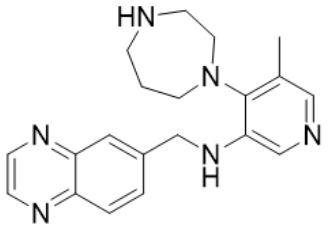
050	
051	
052	
053	
054	
055	
056	
057	

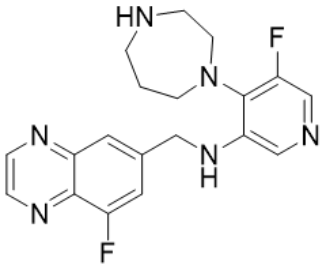
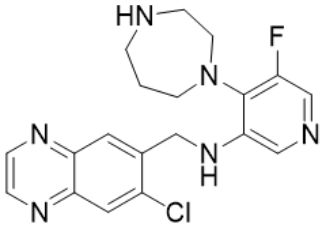
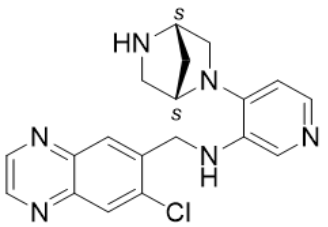
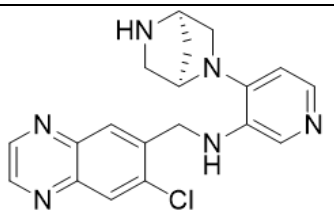
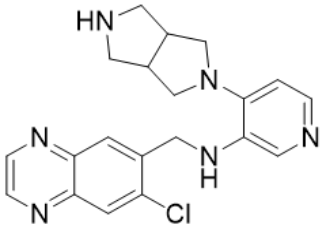
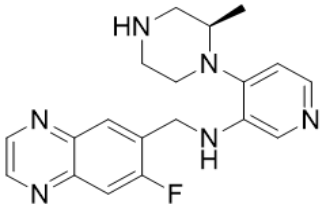
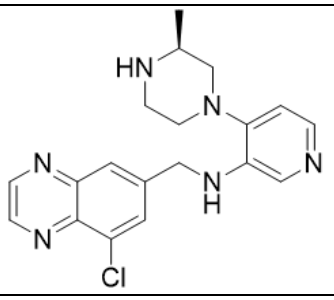
058	
059	
060	
061	
062	
063	
064	

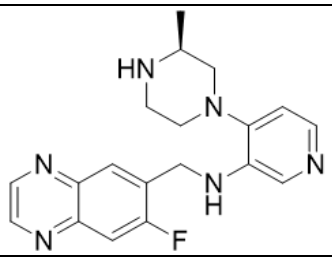
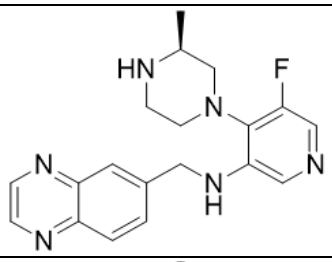
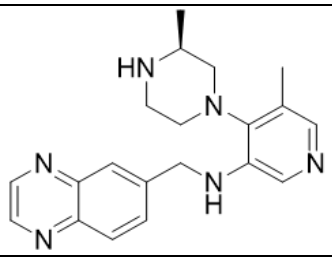
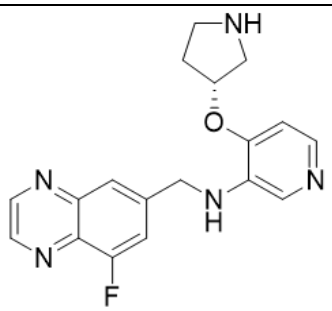
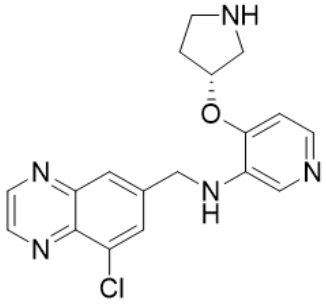
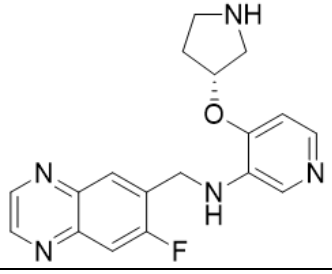
065	
066	
067	
068	
069	
070	
071	

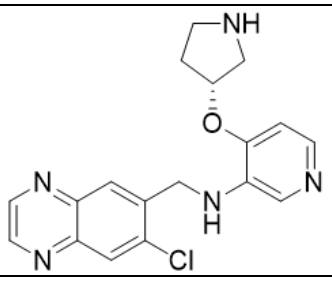
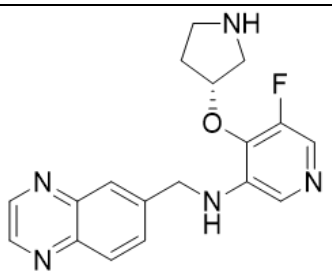
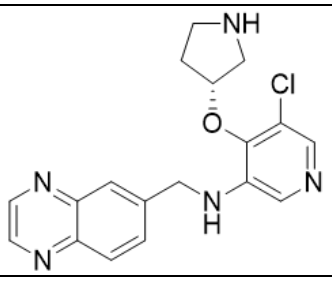
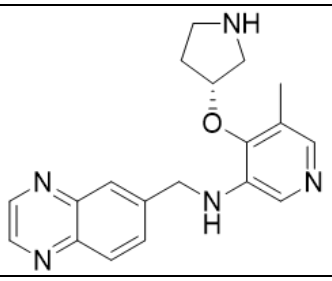
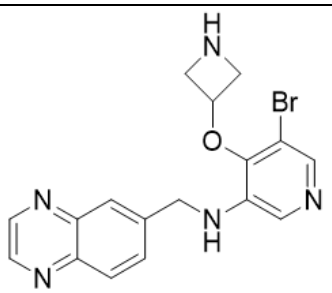
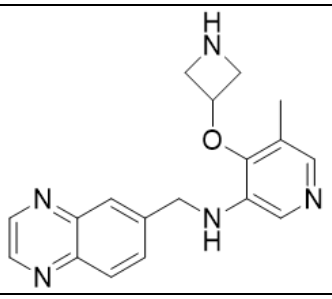
072	
073	
074	
075	
076	
077	
078	
079	

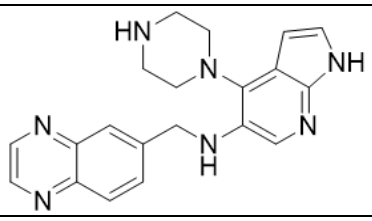
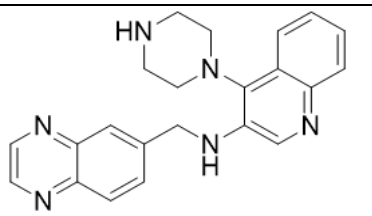
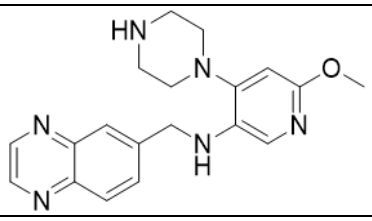
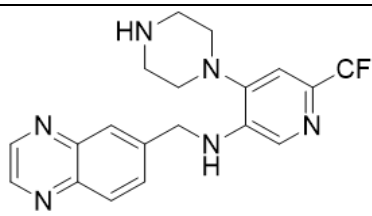
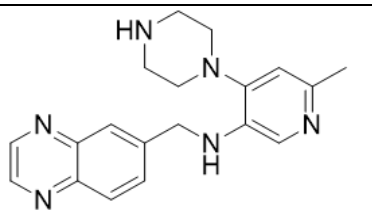
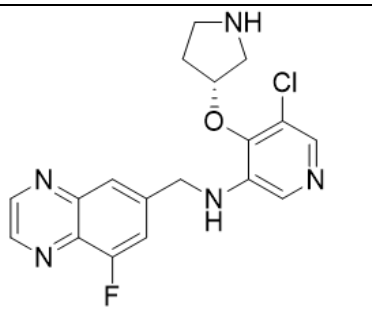
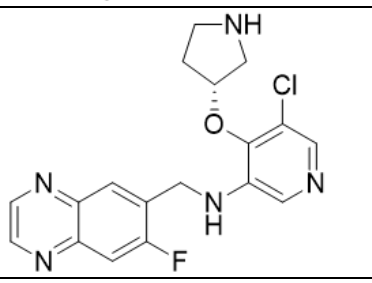
080	
081	
082	
083	
084	
085	
086	
087	

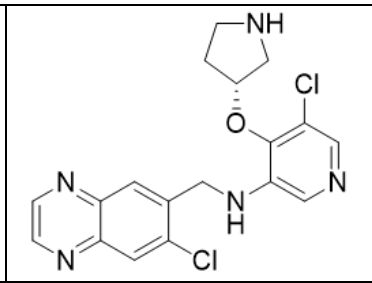
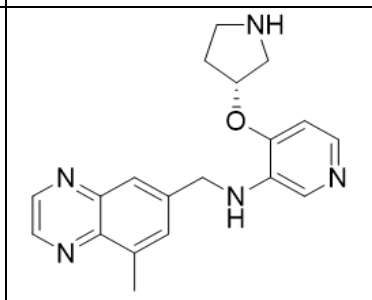
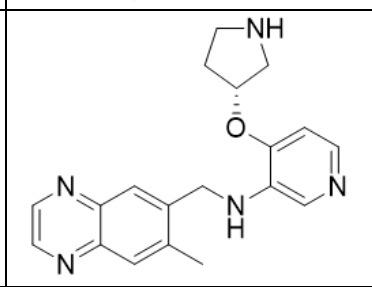
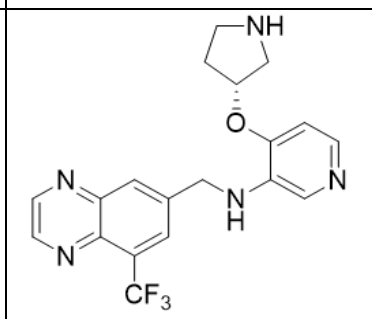
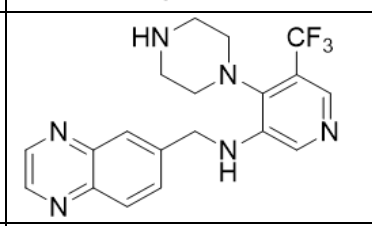
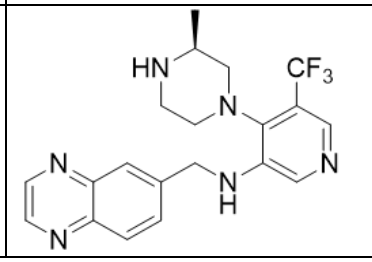
088	
089	
090	
091	
092	
093	
094	

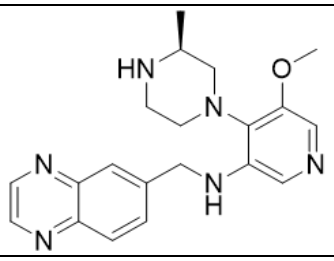
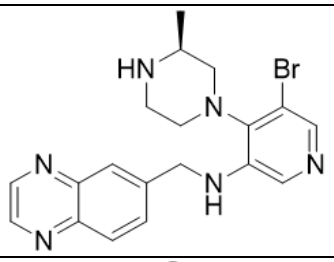
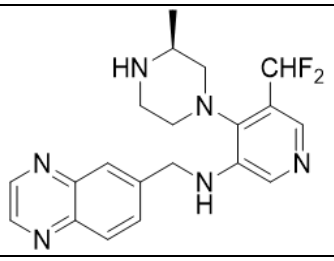
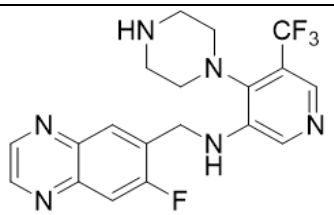
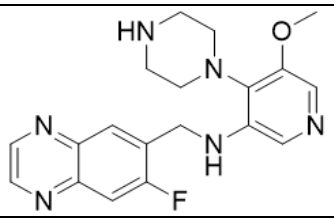
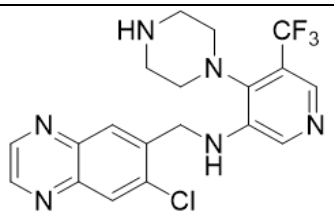
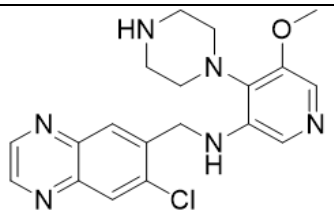
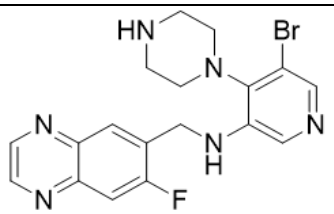
095	
096	
097	
098	
099	
100	
101	

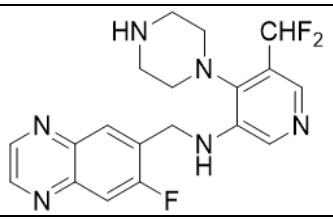
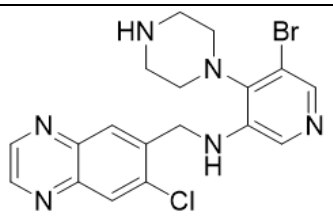
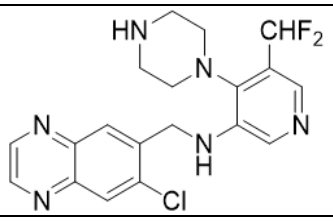
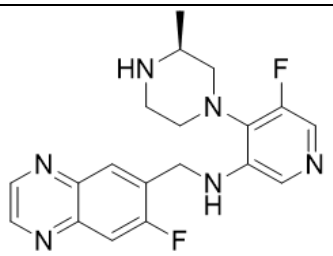
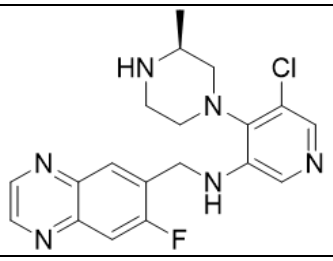
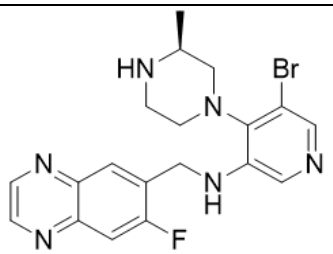
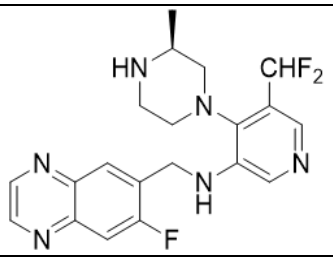
102	
103	
104	
105	
106	
107	

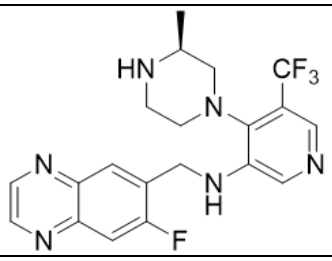
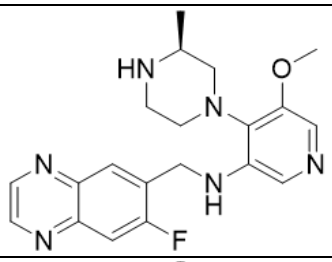
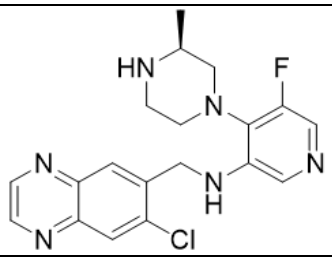
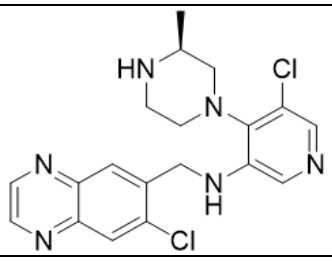
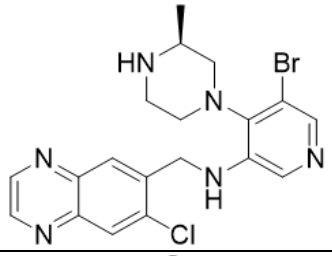
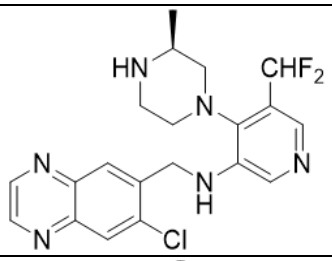
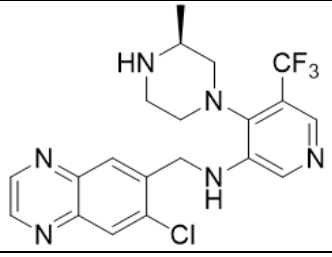
108	
109	
110	
111	
112	
113	

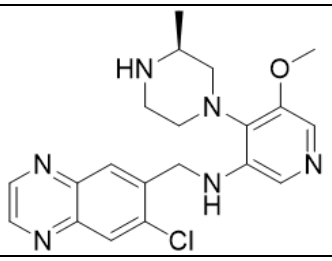
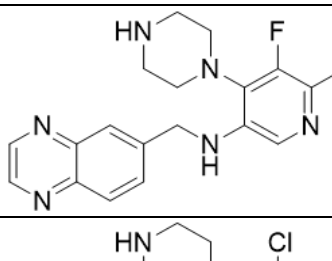
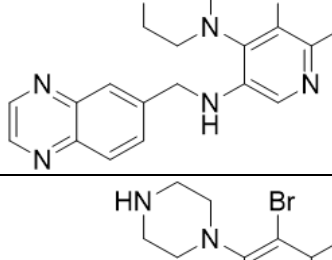
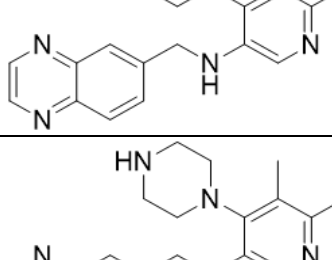
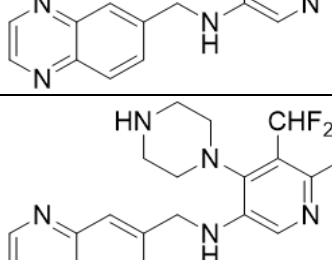
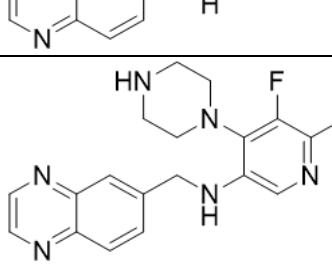
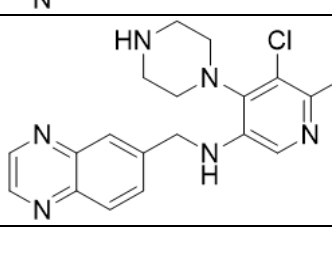
114	
115	
116	
117	
118	
119	
120	

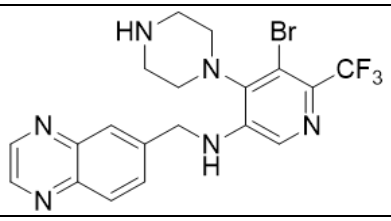
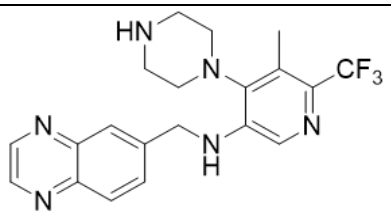
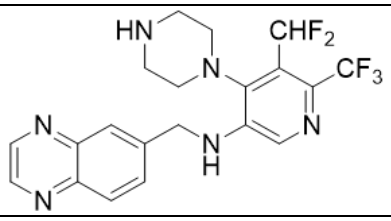
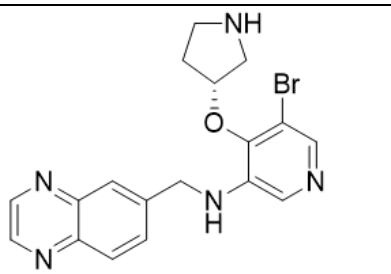
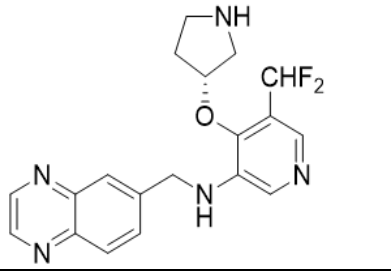
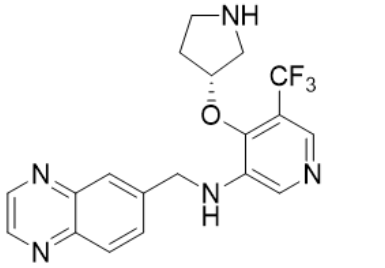
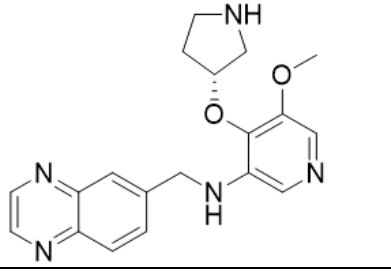
121	
122	
123	
124	
125	
126	

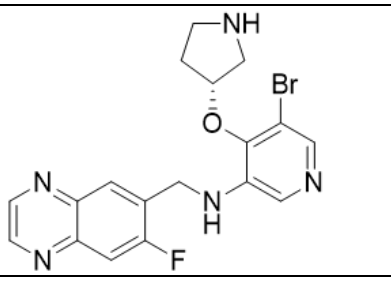
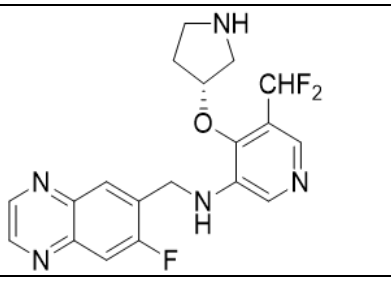
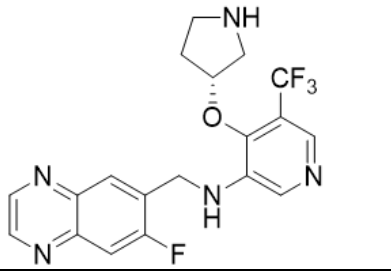
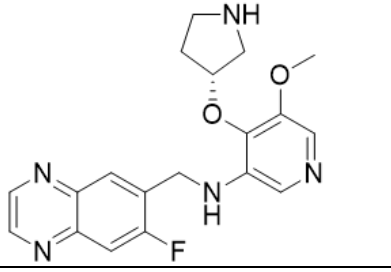
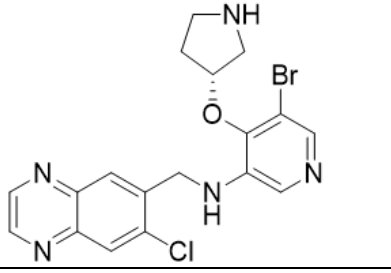
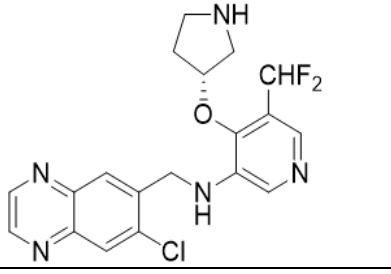
127	
128	
129	
130	
131	
132	
133	
134	

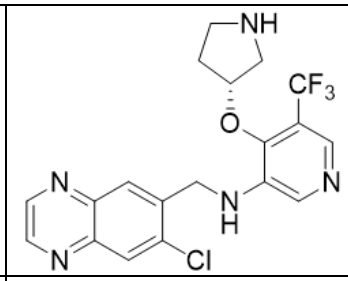
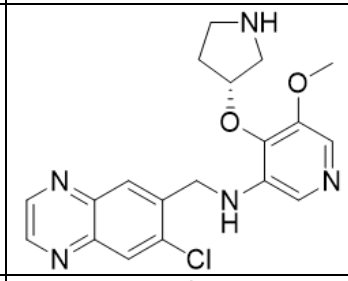
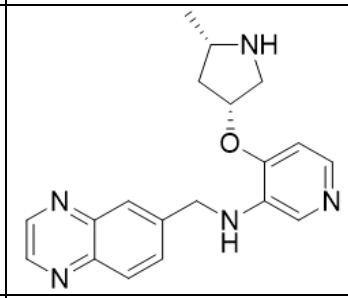
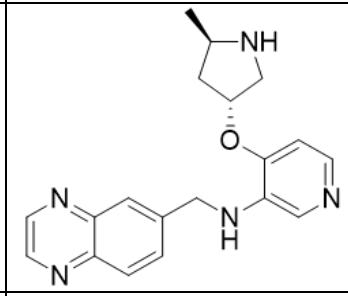
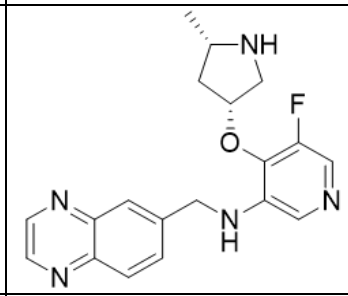
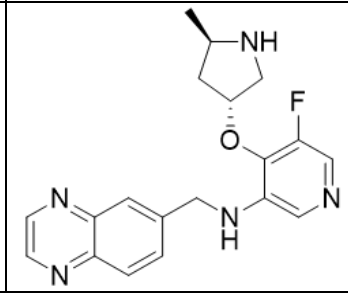
135	
136	
137	
138	
139	
140	
141	

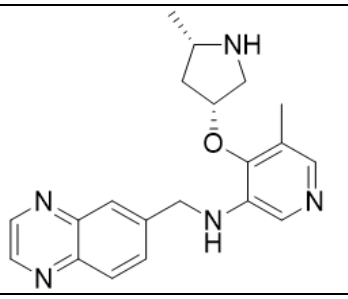
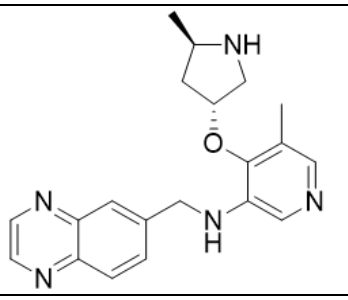
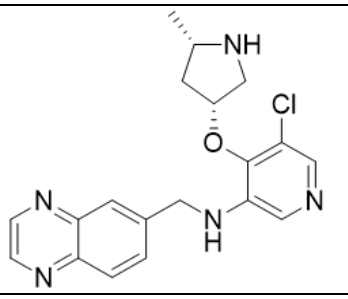
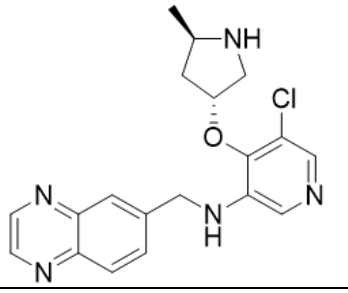
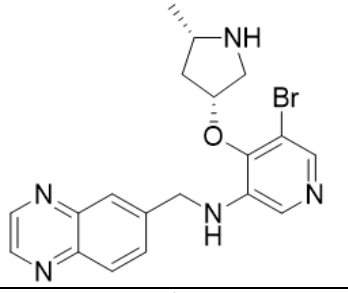
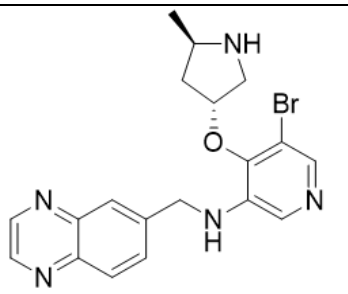
142	
143	
144	
145	
146	
147	
148	

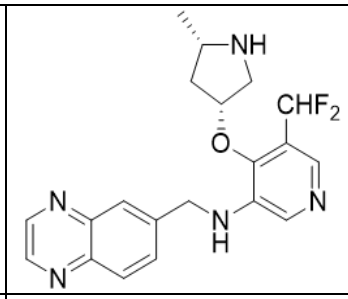
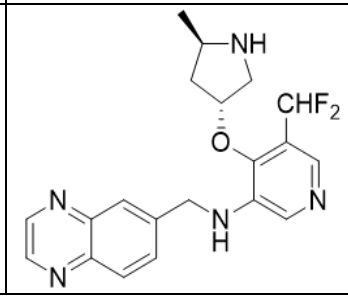
149	
150	
151	
152	
153	
154	
155	
156	

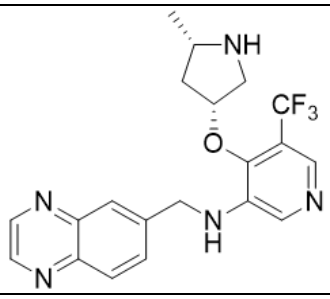
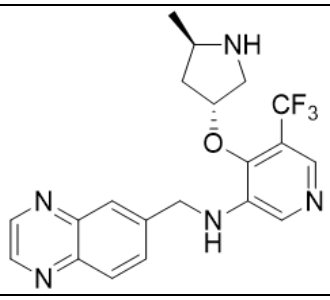
157	
158	
159	
160	
161	
162	
163	

164	
165	
166	
167	
168	
169	

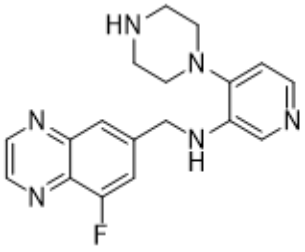
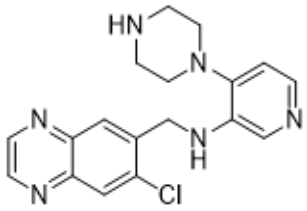
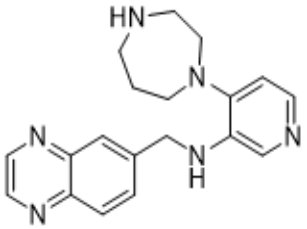
170	
171	
172	
173	
174	
175	

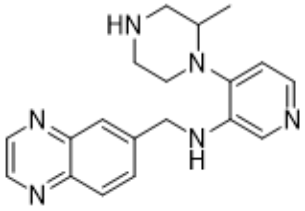
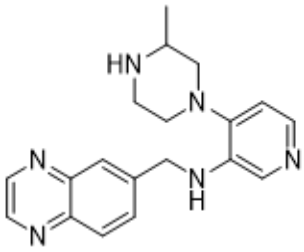
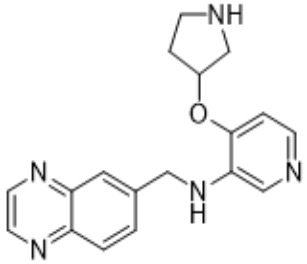
176	
177	
178	
179	
180	
181	

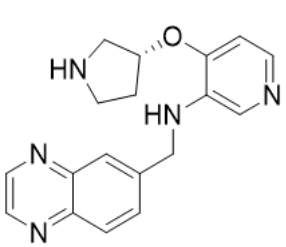
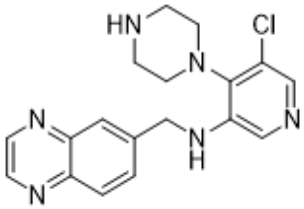
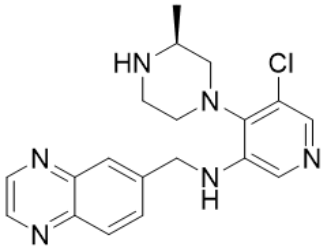
182	
183	

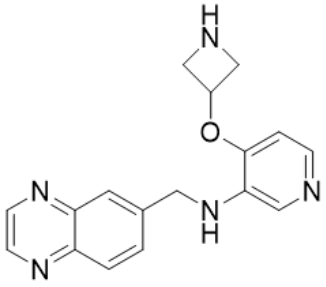
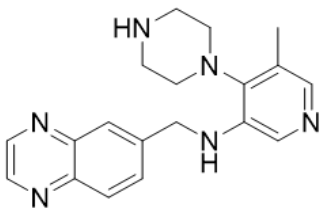
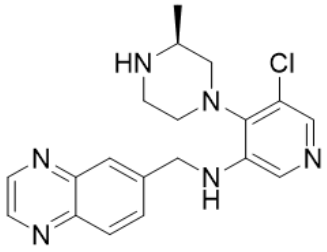
184	
185	

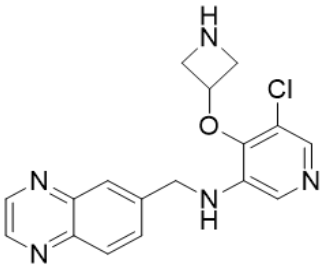
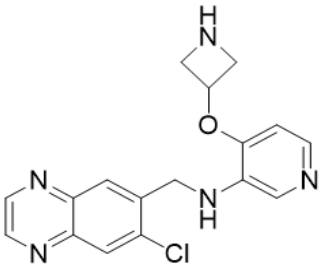
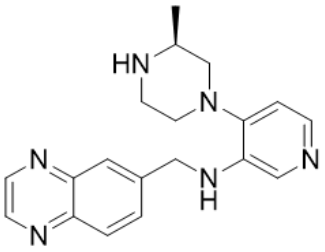
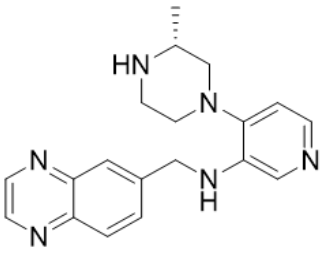
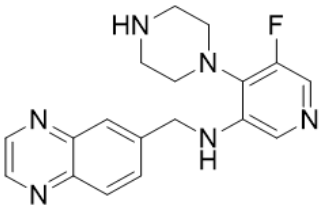
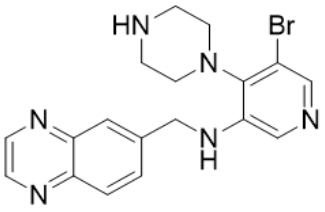
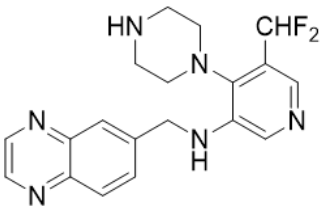
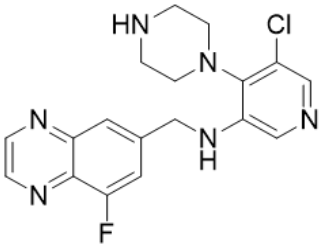
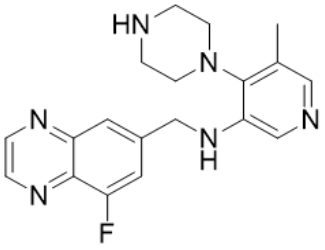
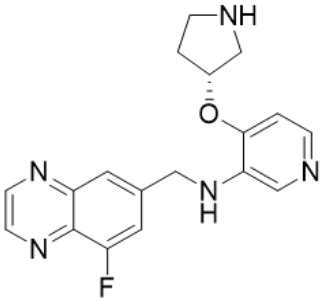
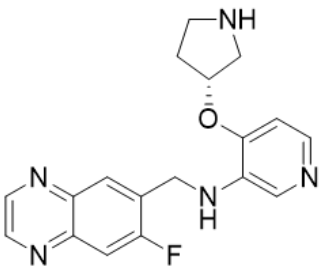
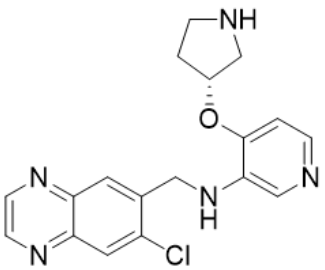
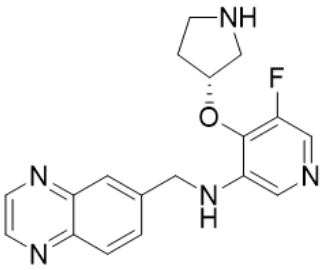
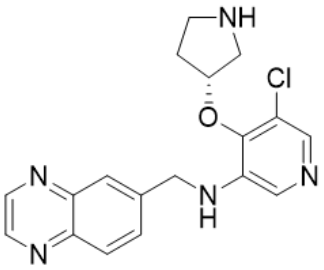
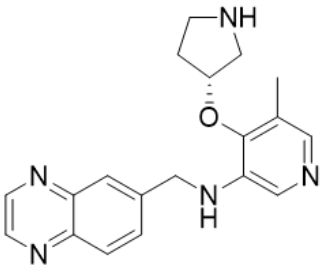
70. The compound of claim 69, having the structure according to one of:

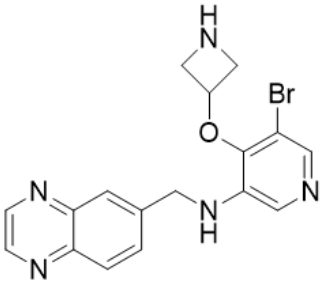
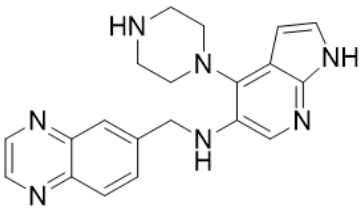
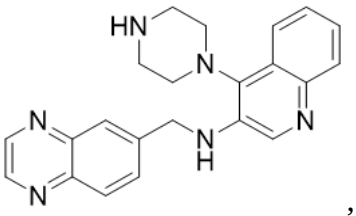
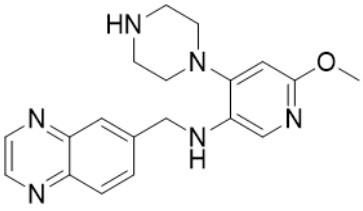
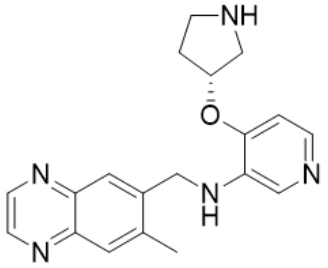
, , ,

, , ,

, , ,

, , ,


,

,

,

,

,

,

,

,

,

,

,

,

,

,

,





, and .