Chemistry & Biology, Volume *19*

Supplemental Information

An Unusual Role for a Mobile Flavin in StaC-like Indolocarbazole Biosynthetic Enzymes

Peter J. Goldman, Katherine S. Ryan, Michael J. Hamill, Annaleise R. Howard-Jones, Christopher T. Walsh, Sean J. Elliott, and Catherine L. Drennan

Inventory of Supplemental Information

Figure S1

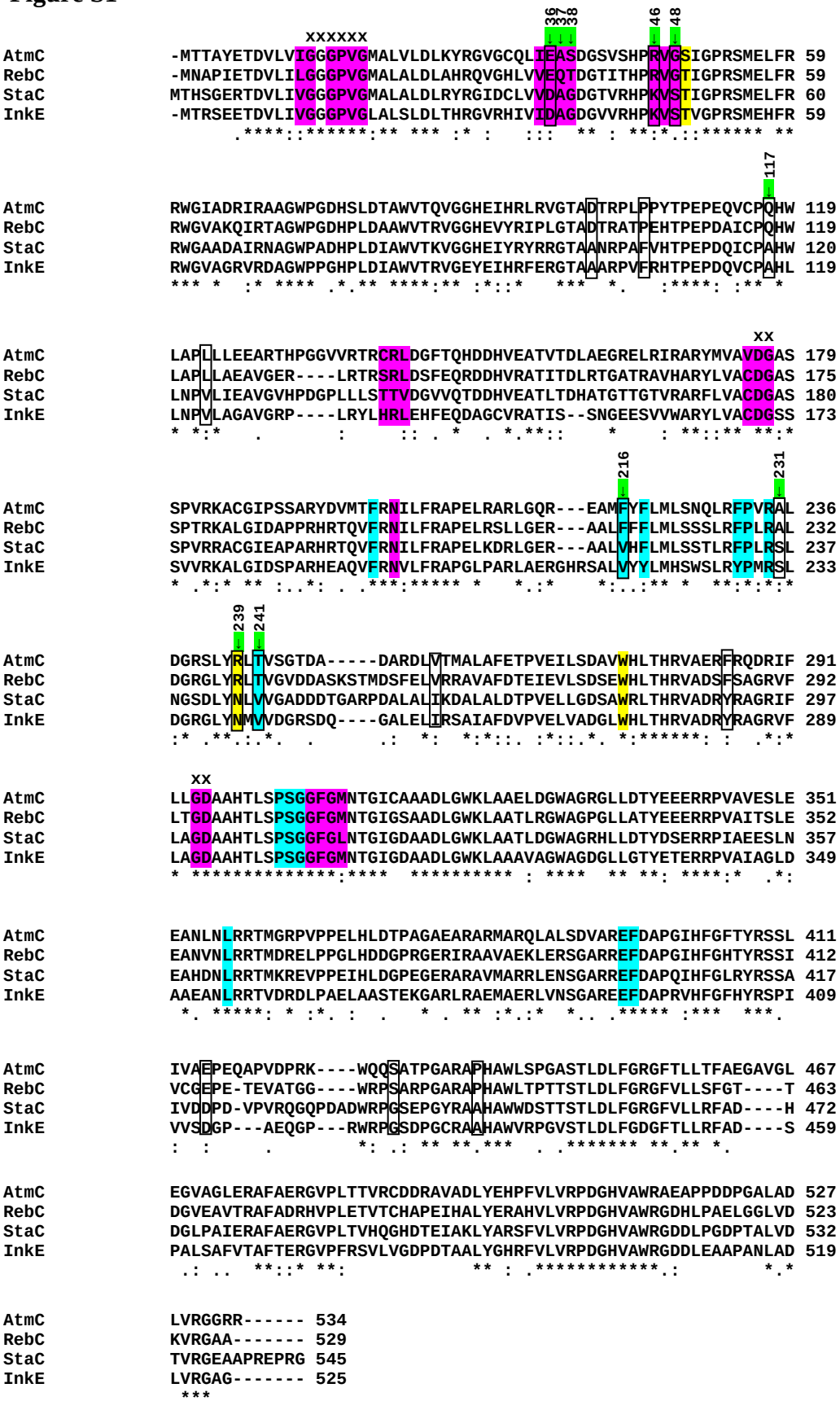
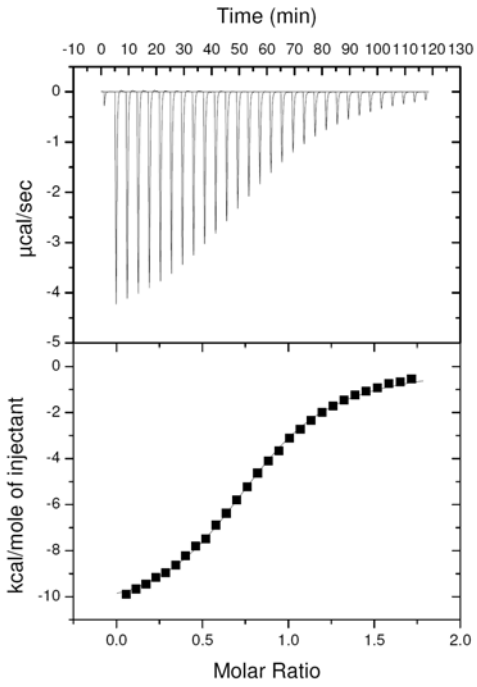

xxxxxx
36
37
38
46
48
117
xx
216
231
239
241
xx
AtmC
RebC
StaC
InkE

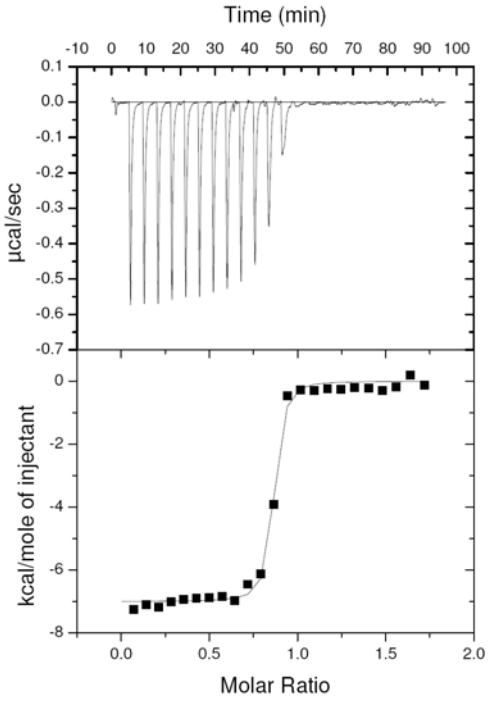
Figure S1, related to Table 1. ClustalW sequence alignment between the highly homologous enzymes AtmC, RebC, StaC, and InkE. High conservation is seen across all four proteins (percent identity between RebC and StaC, RebC and AtmC, StaC and InkE, and AtmC and InkE are 65, 64, 62, and 56, repectively). A star (*) represents residues conserved across all four proteins, and double dots or single dots represent similar (not identical) residues across all four proteins, according to the conventions of ClustalW. A black "x" above the alignment denotes residues involved in conserved FAD binding motifs common to flavin-dependant hydroxylases (Eppink et al., 1997; Howard-Jones and Walsh, 2006; Ryan et al., 2007). To identify amino acids that may be responsible for the functional differences seen between the two pairs of enzymes (AtmC and RebC; StaC and InkE), the crystal structure of RebC in complex with FAD and 7-carboxy-K252c (PDB ID 2R0G) was analyzed. Residues where any atom is within 4 Å of FAD are highlighted in pink, residues where any atom is within 4 Å of putative substrate 7-carboxy-K252c are highlighted in turquoise, and residues where any atom is within 4 Å of both FAD and 7-carboxy-K252c are highlighted in yellow. Boxes indicate amino acids where the aligned residues are identical between StaC and InkE but distinct from the identical residues of RebC and AtmC. Note that some of these residues are not within 4 Å of either FAD or 7-carboxy-K252c. Green arrows point to mutation sites for the generation of RebC-10x and StaC-10x. RebC numbering for these residues are shown above the alignment.

Figure S2.

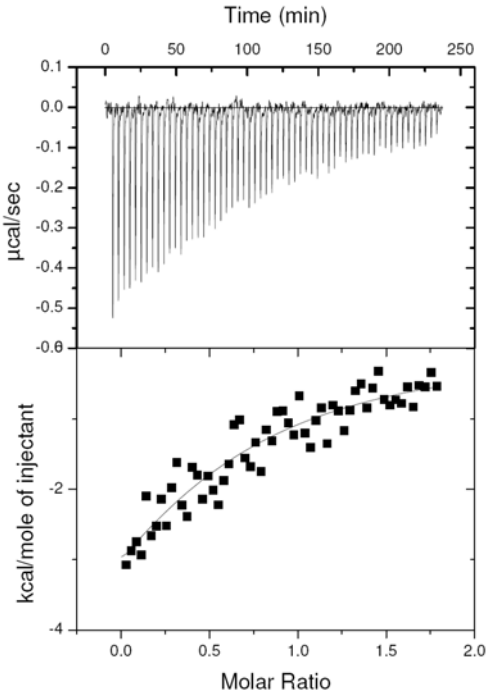
A.



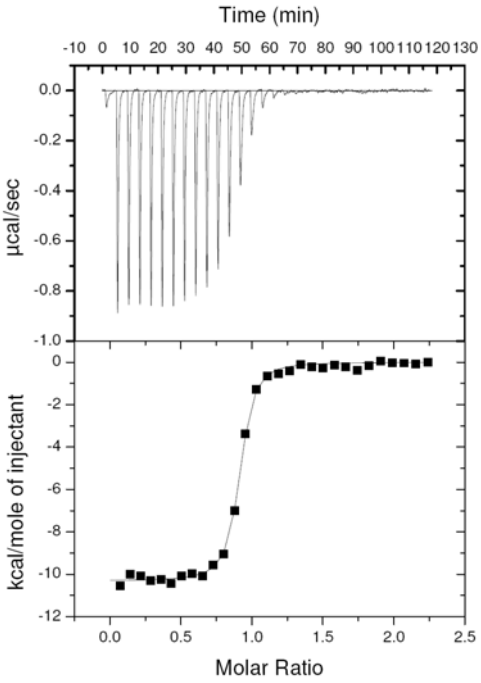
B.



C.



D.



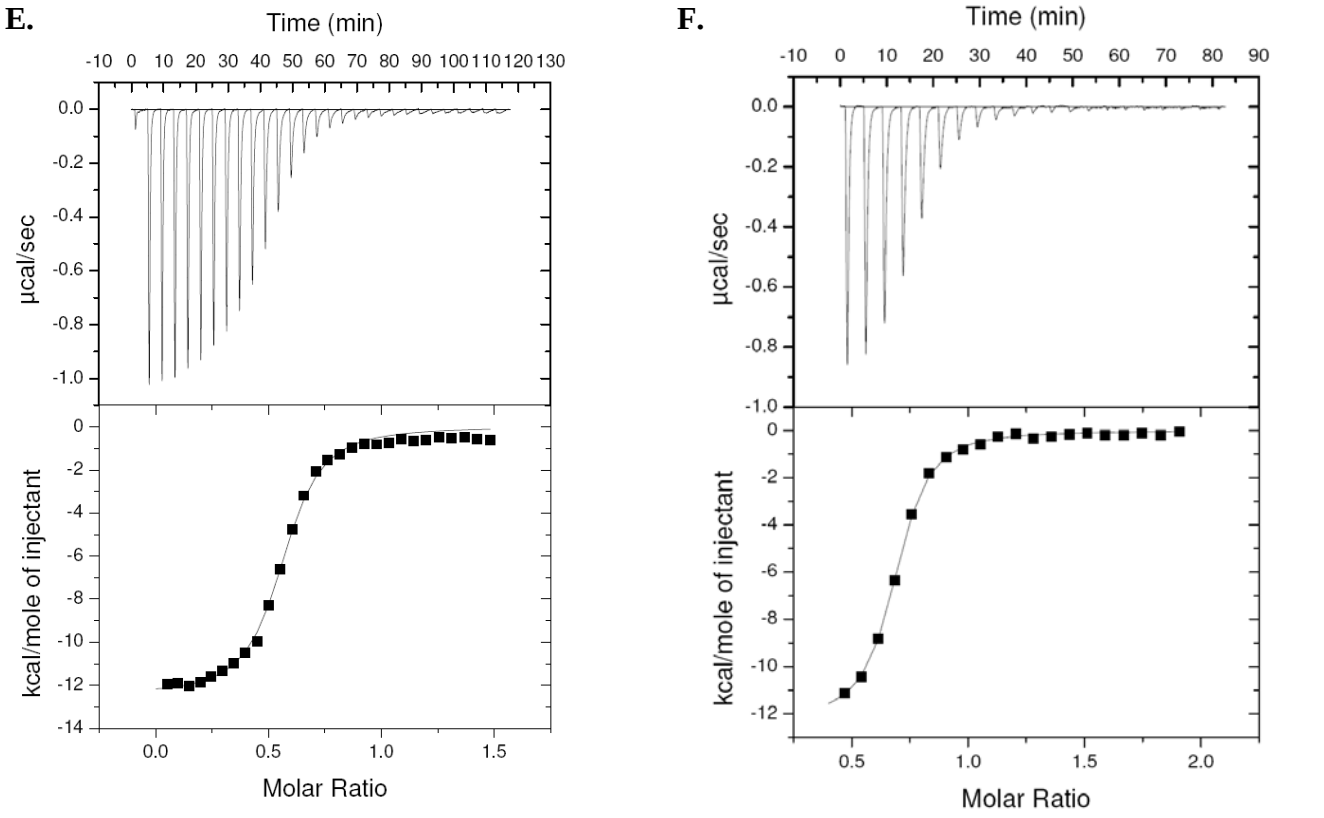


Figure S2, related to Table 2. Isothermal titration calorimetry data for FAD binding to StaC and homologues. ITC trace for (*A*). StaC; (*B*). RebC; (*C*). InkE; (*D*). AtmC; (*E*). RebC-10x; and (*F*). StaC-10x. Protein concentrations are: 29.6 μM (RebC), 30 μM (AtmC), 166 μM (StaC), 132 μM (InkE), 30 μM (StaC-10x), and 44 μM (RebC-10x). FAD was dissolved in the identical gel filtration buffer, and its concentration was calculated using A_{450} using ε_{450} = 11,300 cm^{-1} M^{-1}. FAD concentrations used in ITC experiments are: 300 μM (RebC experiment), 300 μM (AtmC experiment), 405 μM (StaC experiment), 1.05 mM (InkE experiment), 300 μM (StaC-10x experiment) and 300 μM (RebC-10x experiment). ITC experiments were carried out with protein in the cell and FAD in the syringe with the following parameters: 30 injections (one of 0.5 μL, 29 of 10 μL; with a duration time of 1 sec for the first injection and 20 sec for the remaining 29 injections), 240 sec spacing, 2 sec filter, 25°C, reference power of 15 μCal/sec, initial delay of 60 sec, stirring speed of 310, feedback set at high, ITC equilibrium set at fast and auto. For the InkE experiment, all conditions for ITC were identical, except that 60 injections were used (one of 0.5 μL, 59 of 5 μL; with a duration time of 1 sec for the first injection and 10 sec for the remaining 59 injections).

Figure S3

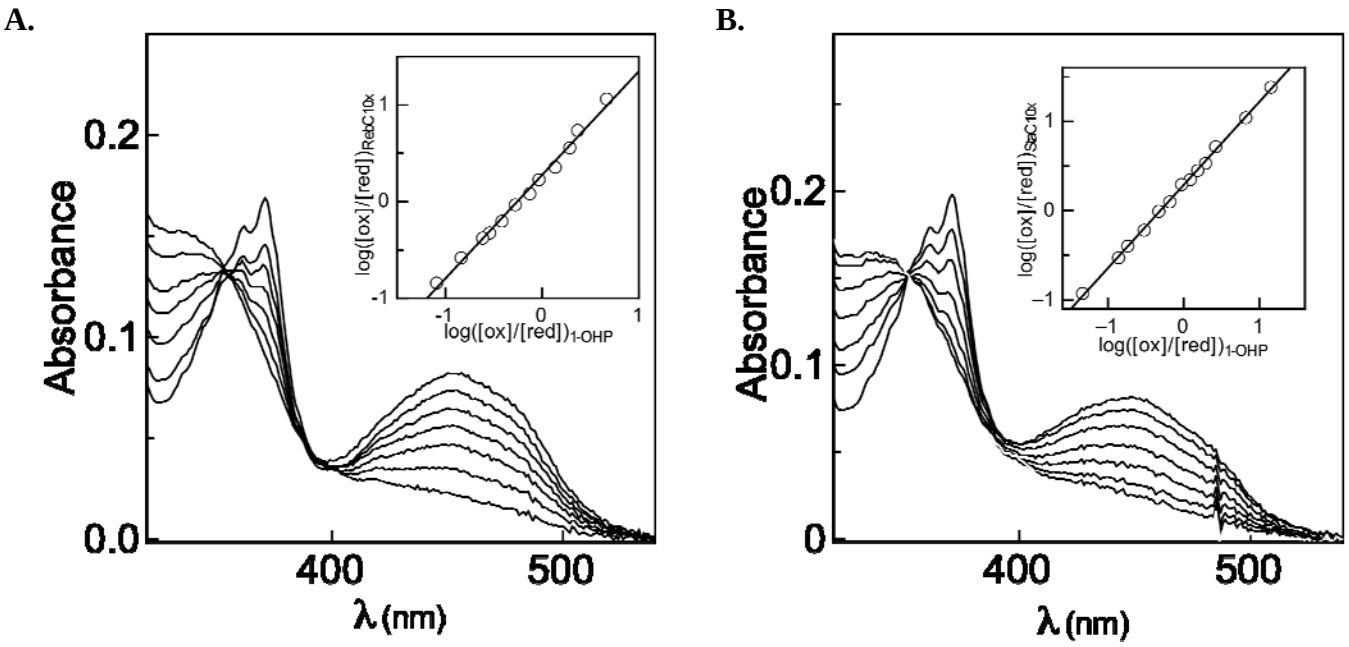


Figure S3. Reduction potential of RebC-10x and StaC-10x. The redox potentials of (*A*). RebC-10x and (*B*). StaC-10x were determined using the xanthine/xanthine oxidase method (Massey, 1991). Inset: Plot of log([ox]/[red]) for 1-OHP vs log([ox]/[red]) for each enzyme, used to calculate the midpoint potential by comparing to 1-OHP (-172 mV).

Figure S4

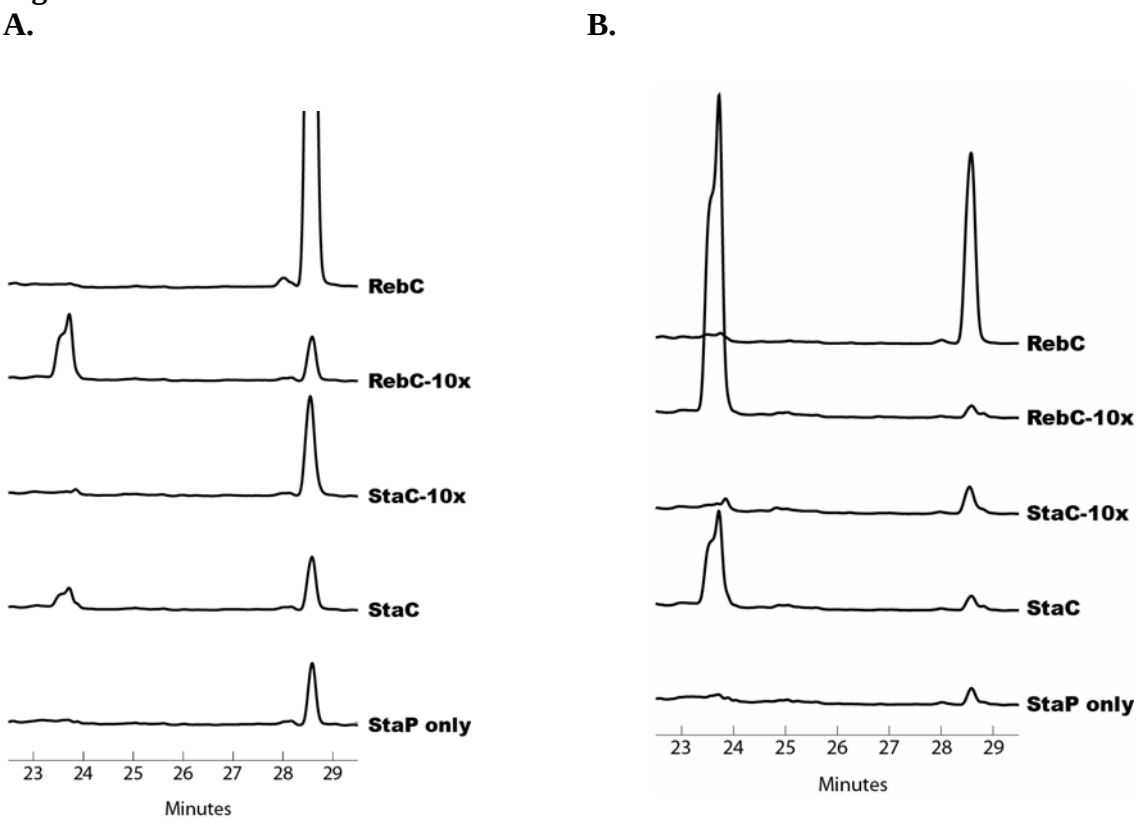


Figure S4, related to Table 3. Activity assay HPLC chromatograms. *(A).* 315 nm chromatogram. *(B).* 289 nm chromatogram. The retention time for arcyriaflavin A is 28.5 minutes. The arcyriaflavin peak in the RebC assay trace has been truncated. The retention time for K252c is 23.5 minutes. Peaks appearing at these times from experimental samples were integrated at the appropriate wavelength. Because reactions are linear at 30 min, integrated peak areas were converted to turnover rates using previously described methods (Howard-Jones and Walsh, 2006). All enzyme assays were also carried out as described previously in the presence of 75 mM HEPES pH 7.5, 1 mg/mL BSA, 5 mM NADPH, 20 μM ferredoxin, and 1 μM flavodoxin $NADP^+$ reductase (Howard-Jones and Walsh, 2006). RebC, StaC, StaC-10x, or RebC-10x were added at 5 μM. All reactions were initiated with the addition of 1 μM StaP and incubated at room temperature for 30 min. Reactions were then quenched with two volumes of methanol and incubated on ice for at least five minutes prior to the removal of precipitated protein by centrifugation. Reactions were assayed using an Agilent 1200 Series reverse-phase HPLC with an Agilent Eclipse C18 analytical column (150 x 4.6 mm). Two buffers (A is 0.2% trifluoroacetic acid in distilled, filtered water, and B is 0.2% trifluoroacetic acid in acetonitrile) were used in the following program of linear gradients at a 1 mL/min flow rate: 10 to 60% B over 30 min, 60 to 100% B over 0.5 min, 100% B for 2 min, 100 to 10% B over 0.5 min, and 10% B for 7 min.

Figure S5

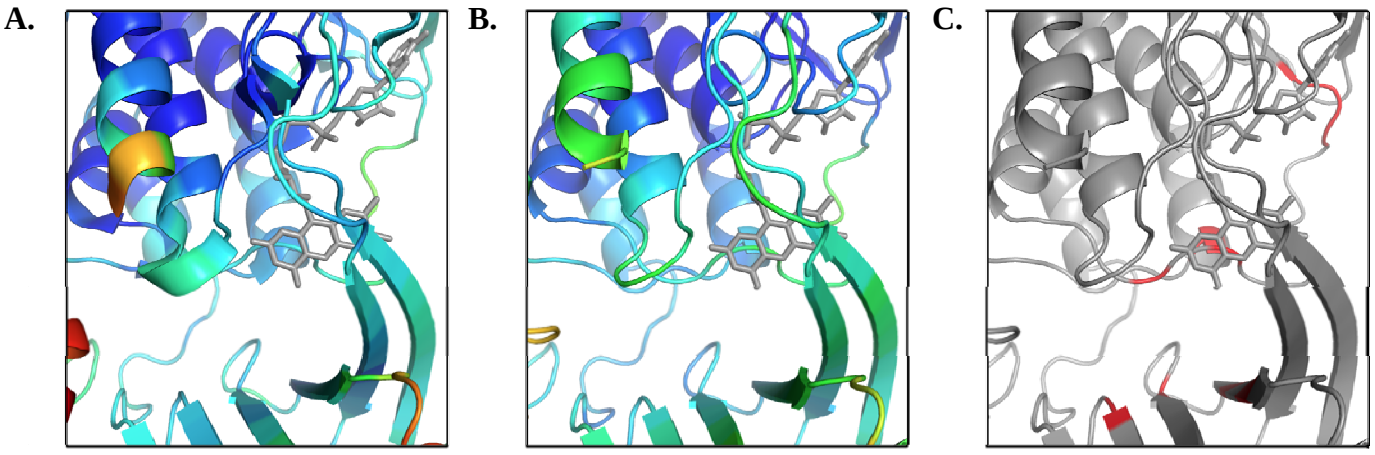
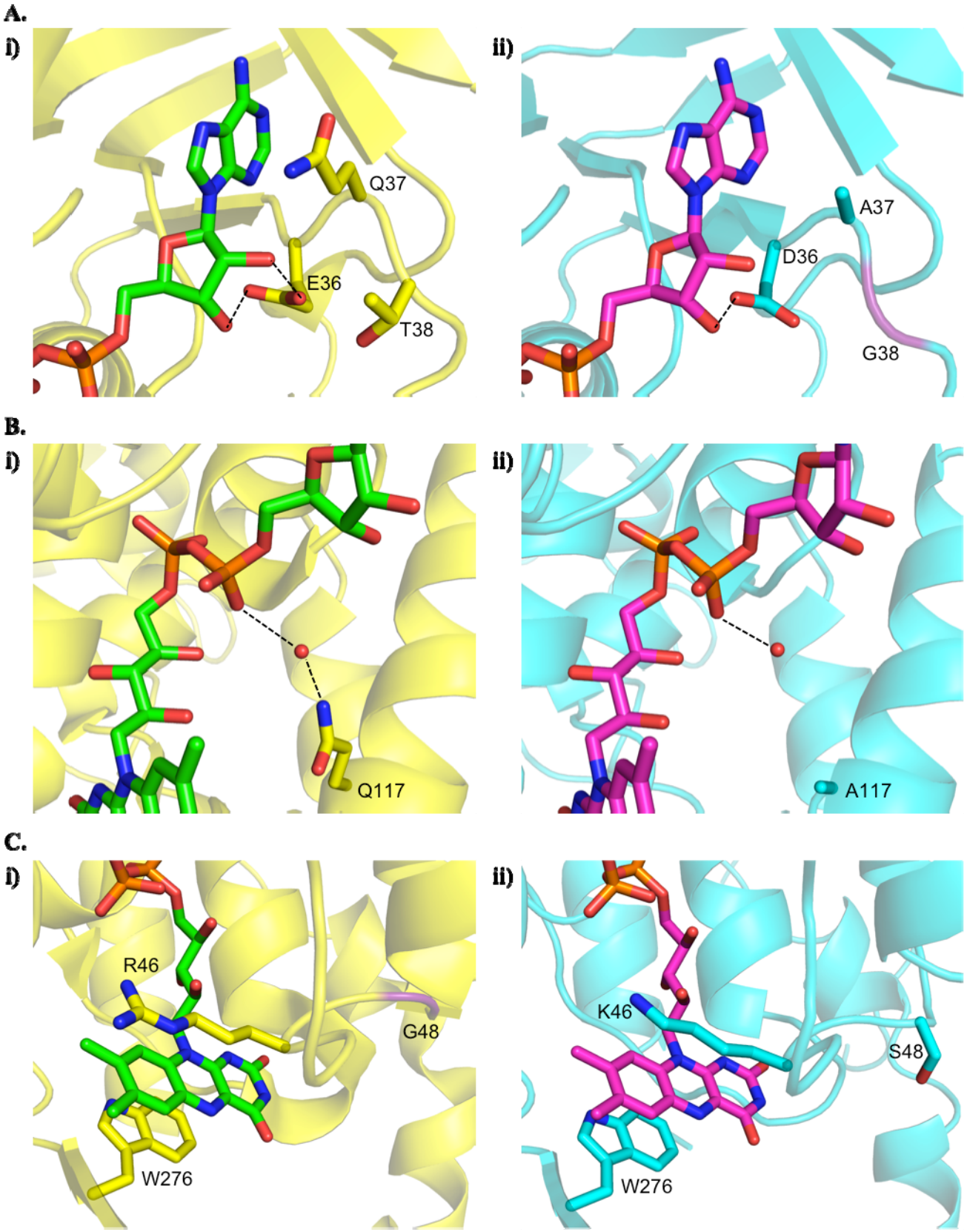


Figure S5, related to Figure 2. B-factor analysis of RebC and RebC-10x in area of the ten mutations. (*A*). RebC (Native structure, PDB ID 2R0C) colored as a heat map from red (high b-factor) to blue (low b-factor). FAD is shown as grey sticks. (*B*). RebC-10x heat map based on "native structure", colored as above. (*C*). The locations of ten mutations in RebC-10x are colored red.

Figure S6


A.
i)
Q37
E36
T38
ii)
A37
D36
G38
B.
i)
Q117
ii)
A117
C.
i)
R46
G48
W276
ii)
K46
S48
W276

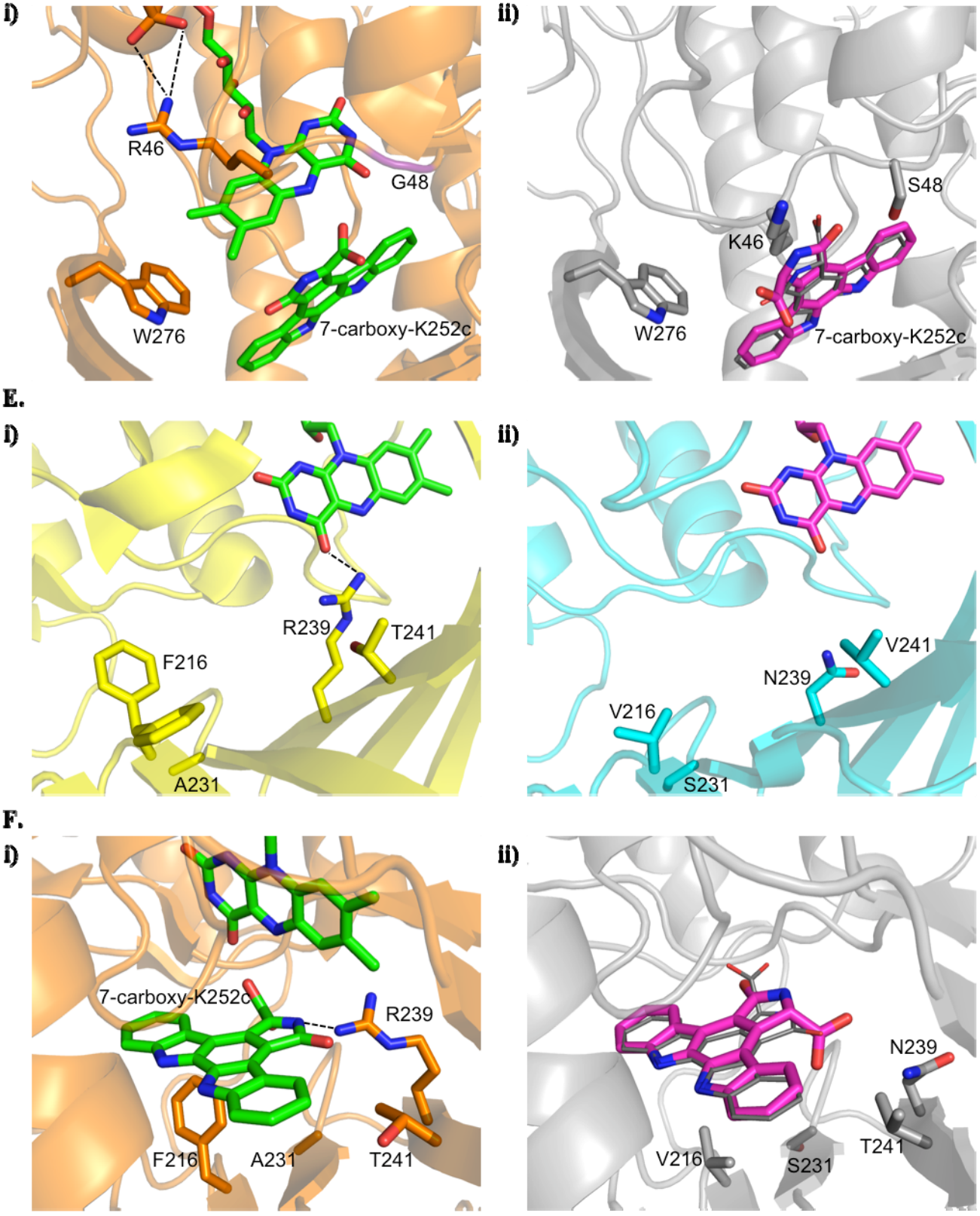

D.
i)
R46
G48
W276
7-carboxy-K252c
ii)
S48
K46
W276
7-carboxy-K252c
E.
i)
R239
T241
F216
A231
ii)
V241
N239
V216
S231
F.
i)
7-carboxy-K252c
R239
F216
A231
T241
ii)
N239
T241
V216
S231

Figure S6, related to Table 1 and Figure 2. Residues chosen from the RebC structures in the design of (i) StaC-10x and RebC-10x and (ii) subsequent RebC-10x residue positioning. Carbon coloring is as follows: native RebC backbone (yellow) and FAD (green), native RebC-10x backbone (cyan) and FAD (magenta), substrate-bound RebC (orange), and substrate-bound RebC-10x (gray). Glycines are shown in purple. *(A)*. A stretch of three residues from RebC – Glu^{36}, Gln^{37}, and Thr^{38} – were installed in StaC-10x, replacing the StaC residues Asp^{37}, Ala^{38}, and Gly^{39}. These residues may modulate binding of the adenosine portion of FAD. These residues are shown in the substrate-free structure of (i) RebC and (ii) RebC-10x, with their orientation unchanged in the structure with bound 7-carboxy-K252c. *(B)*. (i) Gln^{117} in RebC interacts with a phosphate on the FAD via a water molecule. This residue was installed in StaC-10x, replacing Ala^{118}, which was installed in (ii) RebC-10x. Shown are Gln^{117} and Ala^{117} in the substrate-free structure of RebC and RebC-10x. The water binding site is preserved in RebC-10x. The orientation of both waters and residues are unchanged in the structures with bound 7-carboxy-K252c. *(C)*. (i) Arg^{46} in RebC binds on the *si* face of the FAD in the substrate-free structure. Two residues away is Gly^{48} (shown in purple). (ii) The corresponding residues in StaC are Lys^{47} and Ser^{49}, and were installed in RebC-10x at positions 46 and 48. Also shown is Trp^{276}, which is unchanged in StaC, but is shown because it stacks on the *re* face of the FAD. *(D).* As in *(C)*, but in the structures of RebC and RebC-10x are now bound with two different bound tautomers of 7-carboxy-K252c; the (i) enol and (ii) keto. (ii) Two different orientations of 7-carboxy-K252c were found in RebC-10x, one in each monomer; they are overlayed, one in magenta (molecule B) and one in thin grey lines (molecule A). (i) In RebC, Arg^{46} now hydrogen bonds with a phosphate on the FAD. (ii) Flavin is not present in the RebC-10x substrate bound model and the loop containing Lys^{46} has shifted downward. *(E).* (i) Residues Phe^{216}, Ala^{231}, Arg^{239}, and Thr^{241} in RebC line the active site pocket, however, only Arg^{239} directly interacts with FAD, via a hydrogen bond with a carbonyl oxygen on the isoalloxazine ring. These residues were installed in StaC-10x, replacing Val^{221}, Ser^{236}, Asn^{244}, and Val^{246}, respectively. (ii) In RebC-10x, these StaC residues result in a more open binding pocket. Shown are the orientations of the (i) RebC and (ii) RebC-10x residues in the substrate-free structures. *(F)*. As in (*E*), but in the structures of (i) RebC and (ii) RebC-10x with their respective bound tautomers of 7-carboxy-K252c. (i) Arg^{239} now hydrogen bonds with the bound molecule in RebC; (ii) Asn^{239} is too far to hydrogen bond to the molecule in RebC-10x.

Figure S7

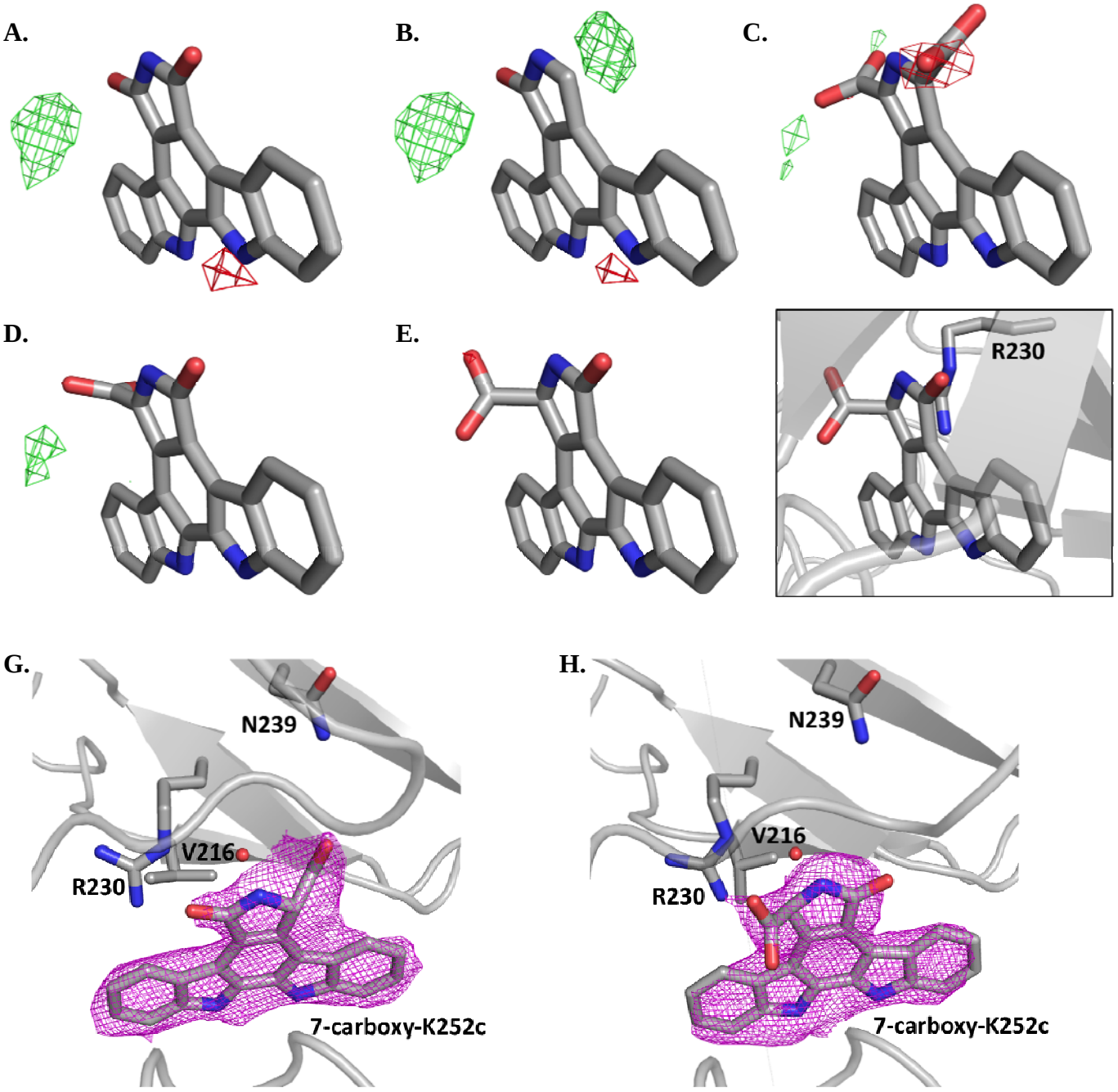


Figure S7, related to Figure 3. Identification and orientation of bound indolocarbazoles in CPA-soaked RebC-10x. Difference electron density after refinement of the following molecules in the RebC-10x active site: (*A*). Arcyriaflavin A; (*B*). K252c; (*C*). Aryl-aryl coupled chromopyrrolic acid; (*D*). Enol tautomer of 7-carboxy-K252c; and (*E*). Keto tautomer of 7-carboxy-K252c. F_o-F_c difference electron density is contoured at +3.0σ (green) and -3.0σ (red), respectively. To show the relationship between this view and others, the box inlay shows the position of Arg230, a residue labeled in (*G)* and *(H)* below as well as in Figures 2 and 3. (*G*). One orientation of the bound keto tautomer of 7-carboxy-K252c with the carboxyl moiety pointed away from Arg230 is found in molecule B. (*H*). Another orientation of the bound keto tautomer of 7-carboxy-K252c with the

carboxyl moiety pointed toward Arg230 is found in molecule A. Shown in pink mesh is a $2F_o$-F_c composite omit map contoured to 1.0σ.

Figure S8

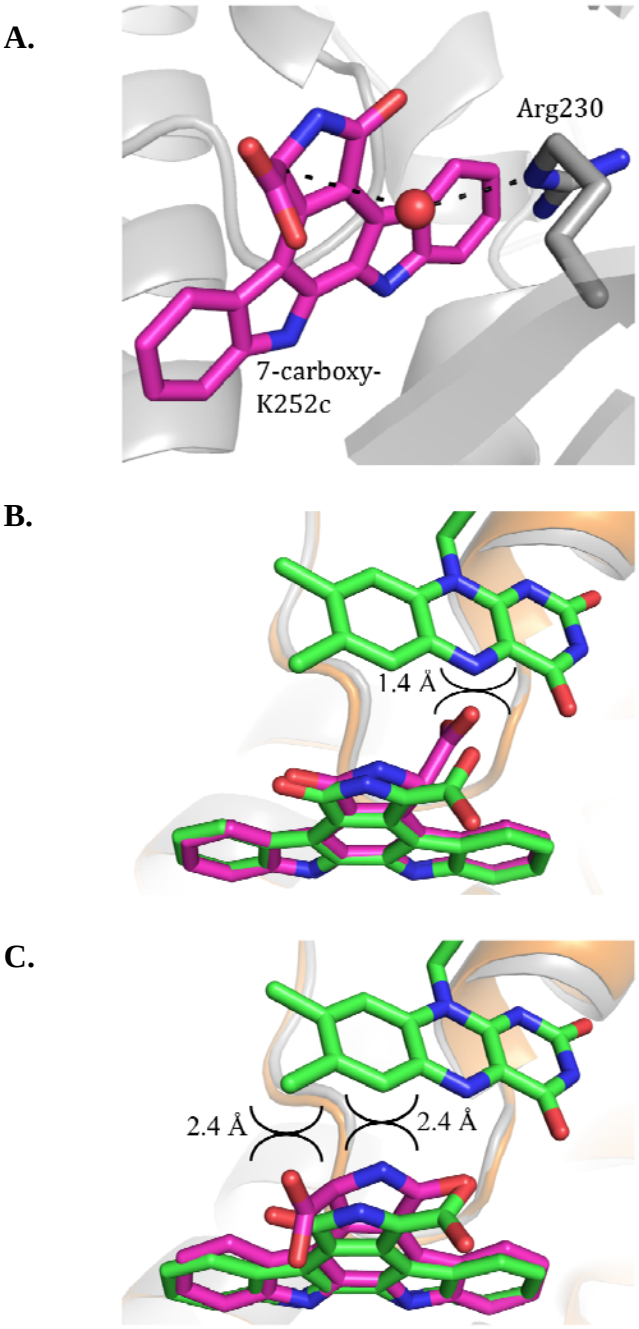
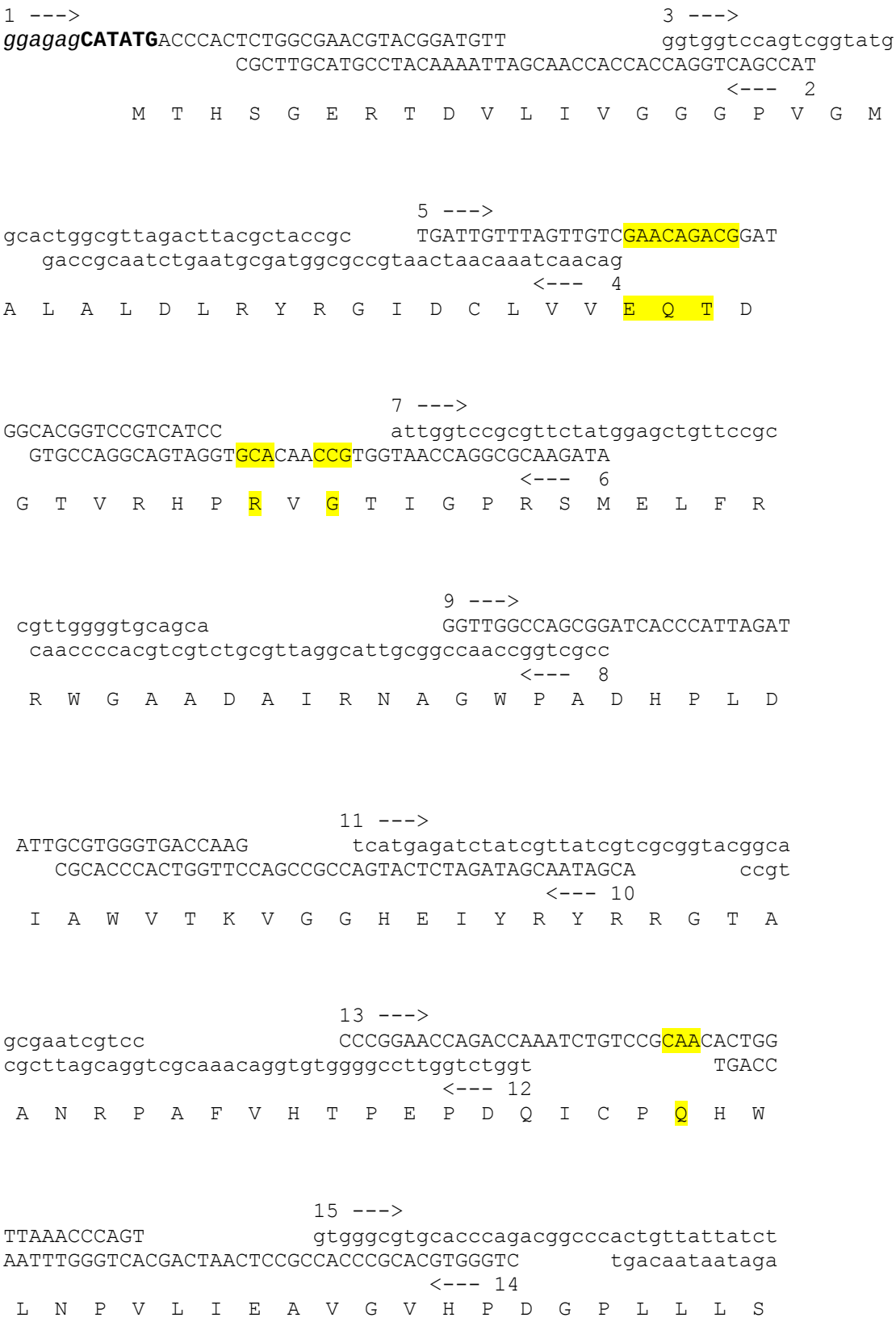


Figure S8, related to Figure 3. Interactions of *S*-keto 7-carboxy-K252c with water and modeled flavin. *(A).* Water binding site for alternative orientation of substrate from that shown in Figure 3E. The water is 2.9 and 3.4 Å from the Arg230 Nε position and the C-7 position of 7-carboxy-K252c, respectively. (*B*). Steric clash between FAD modeled in the "in" position and one orientation of bound *S*-keto 7-carboxy-K252c from the RebC-10x CPA soak structure. Colors: RebC-10x ribbons in grey, *S*-keto 7-carboxy-K252c carbons in magenta, RebC substrate-bound structure (PDB ID 2R0G) ribbons in orange, FAD and enol 7-carboxy-K252c carbons in green. (*C*). Steric clash between FAD

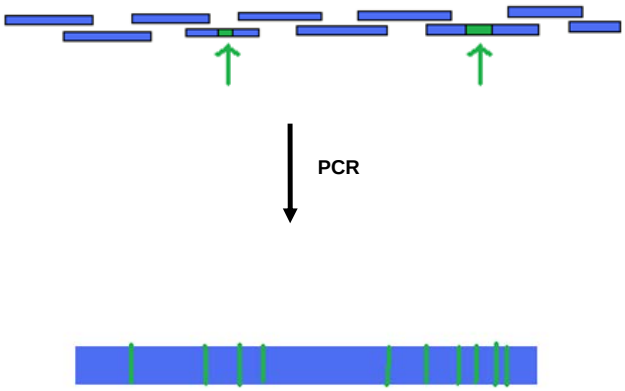
modeled in the “in” position and the other orientation of bound *S*-keto 7-carboxy-K252c from the RebC-10x CPA soak. Colors as in (*B*).

Figure S9
A.


1 --->
ggagagCATATGACCCACTCTGGCGAACGTACGGATGTT
3 --->
ggtggtccagtcggtatg
CGCTTGCATGCCTACAAAATTAGCAACCACCACCAGGTCAGCCAT
<--- 2
M T H S G E R T D V L I V G G G P V G M
5 --->
gcactggcgttagacttacgctaccgc
TGATTGTTTAGTTGTCGAACAGACGGAT
gaccgcaatctgaatgcgatggcgccgtaactaacaaatcaacag
<--- 4
A L A L D L R Y R G I D C L V V E Q T D
7 --->
GGCACGGTCCGTCATCC
attggtccgcgttctatggagctgttccgc
GTGCCAGGCAGTAGGTGCACAACCGTGGTAACCAGGCGCAAGATA
<--- 6
G T V R H P R V G T I G P R S M E L F R
9 --->
cgttggggtgcagca
GGTTGGCCAGCGGATCACCCATTAGAT
caaccccacgtcgtctgcgttaggcattgcggccaaccggtcgcc
<--- 8
R W G A A D A I R N A G W P A D H P L D
11 --->
ATTGCGTGGGTGACCAAG
tcatgagatctatcgttatcgtcgcggtacggca
CGCACCCACTGGTTCCAGCCGCCAGTACTCTAGATAGCAATAGCA
ccgt
<--- 10
I A W V T K V G G H E I Y R Y R R G T A
13 --->
gcgaatcgtcc
CCCGGAACCAGACCAAATCTGTCCGCAACACTGG
cgcttagcaggtcgcaaacaggtgtggggccttggtctggt
TGACC
<--- 12
A N R P A F V H T P E P D Q I C P Q H W
15 --->
TTAAACCCAGT
gtgggcgtgcacccagacggcccactgttattatct
AATTTGGGTCACGACTAACTCCGCCACCCGCACGTGGGTC
tgacaataataga
<--- 14
L N P V L I E A V G V H P D G P L L L S

```
                17 --->
acgaccgtt       CGTCCAAACCGACGACCATGTCGAGGCGACGCTGACCGATCAC
tgctggcaactgccgcagcaggtttggctgct              CGACTGGCTAGTG
                        <--- 16
 T  T  V  D  G  V  V  Q  T  D  D  H  V  E  A  T  L  T  D  H


                  19 --->
GC                cgtgcgtgcccgctttctggttgcctgcgacggcgcaagc
CGGTGGCCATGGTGGCCGTGGCACGCACGGGC                  ccgcgttcg
                        <--- 18
 A  T  G  T  T  G  T  V  R  A  R  F  L  V  A  C  D  G  A  S


                21 --->
tctcc           GTGGTATTGAAGCCCCAGCACGCCATCGCACGCAGGTCTTC
agaggccaagcagcacggacaccataacttcggggt           TGCGTCCAGAAG
                           <--- 20
 S  P  V  R  R  A  C  G  I  E  A  P  A  R  H  R  T  Q  V  F


              23 --->
CGTA          gtgccccggagttaaaagatcgcctgggtgaacgtgcggcgctg
GCATTATAGAATAAGGCACGGGGCCTCAATTTT               cacgccgcgac
                        <--- 22
 R  N  I  L  F  R  A  P  E  L  K  D  R  L  G  E  R  A  A  L


                25 --->
                GCAGCACGTTACGCTTTCCATTACGCGCATTAAATGGTAGC
aaagtaaagaattacaattcgtcgtgcaatgcga              AATTTACCATCG
                         <--- 24
 F  H  F  L  M  L  S  S  T  L  R  F  P  L  R  A  L  N  G  S


                27 --->
GATT            gttggtgcagatgatgatacgggtgcccgtccggatgcatta
CTAAATATGGCAGACTGCCAACCACGTCTACTA          ggcaggcctacgtaat
                         <--- 26                          <---
 D  L  Y  R  L  T  V  G  A  D  D  D  T  G  A  R  P  D  A  L


gca
cgtgACTAGTtcctgc
       < --- 28
 A
```

B.



C.

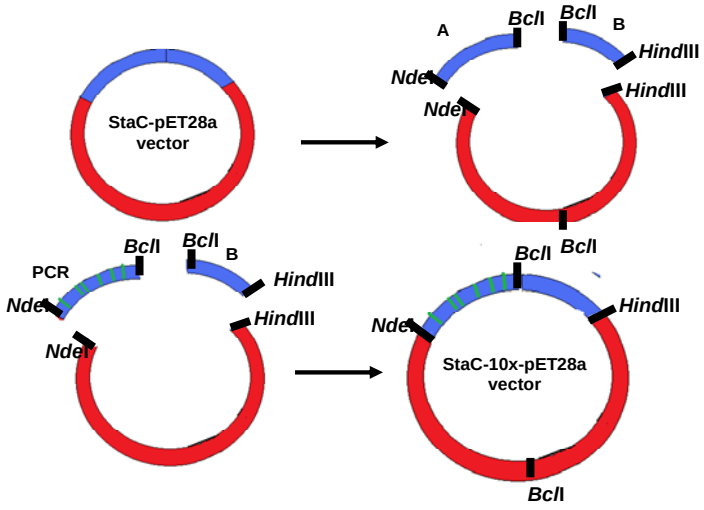


Figure S9. Construction of the *staC-10x* gene. (*A*). Primer-based construction of the N-terminus of StaC-10x. Encoded amino acids are listed below the DNA sequence, with altered amino acids and their corresponding codons highlighted in yellow. DNA sequences in bold represent restriction digestion sites, and DNA sequences in italics represent overhangs. (*B*). PCR gives rise to a fragment encoding ten modified amino acids. Top: partial scheme of PCR-based assembly, with green arrows indicating altered codons (also in green). Bottom: Result of total PCR-based assembly is a fragment with 10 altered codons (shown as green lines). (*C*). A three-part ligation incorporates the fragment into the pET28a-StaC vector, replacing the region encoding the N-terminus of the StaC protein. Restriction digest sites are indicated.

Supplemental Experiemental Procedures

Generation of expression vectors for wild-type proteins

RebC was cloned into pET28a (Novagen) as previously described (Howard-Jones and Walsh, 2006). *StaC* was amplified via PCR from *Streptomyces longisporoflavus* genomic DNA using primers 5'-*ggagag***CATATG**ACGCATTCCGGTGAGCGGAC-3' and 5'-*gtc***AAGCTT**TCAGCCCCGCGGCTCACGGGGCG-3' (Integrated DNA Technologies), where italicized text indicates an overhang and bold text indicates a restriction digest site. The PCR reaction mixture contained 1.25 U of *Pfu* Turbo DNA polymerase (Stratagene), 1x cloned Pfu buffer (Stratagene, 20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/mL bovine serum albumin), 5% DMSO, 250 μM dNTPs (an equimolar mixture dATP, dCTP, dGTP, and dTTP), 1 μL of purified genomic DNA, and 0.5 mM of each primer. The 100 μL reaction mixtures were subjected to the following PCR cycle: 94 °C (3 min); 30 cycles of 94 °C (1 min), 52, 56, 60 or 63 °C (1 min), and 72 °C (2.5 min); 72 °C (10 min). Annealing temperatures of 56, 60 and 63°C all gave successful PCR products, so these were all combined. The PCR fragment was gel purified using the QIAquick Gel Extraction Kit (Qiagen) and digested at 37°C with *NdeI* and *HindIII* (New England Biolabs) in NEBuffer 2 (New England Biolabs, 10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9). Purified pET28a (Novagen) was identically digested with *NdeI* and *HindIII* (New England Biolabs). Both digested PCR product and pET28a were again gel purified. The *staC* insert was ligated into the cut pET28a vector using T4 DNA Ligase (New England Biolabs, 2000 U) in 1x T4 DNA Ligase buffer (New England Biolabs, 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP), with incubation at 16 °C for 12 h**.** Ligation reactions were used to transform TOP10 chemically competent *E. coli* cells, which were then plated onto LB-agar-kanamycin plates. Resulting colonies were amplified in LB-kanamycin, and the plasmid purified (QIAprep Spin Miniprep Kit) then analyzed by restriction enzyme digestion (*NdeI*/*BglI*) for

presence of the insert. The sequence integrity was confirmed by DNA sequencing (Molecular Biology Core Facility, Dana-Farber Cancer Institute, Boston).

InkE was amplified via PCR from *Nonomuraea longicatena* (Kim et al., 2007) genomic DNA using primers 5'-*ggagag***CATATG**ACTCGCAGCGAAGAGACCGAC-3' and 5'-*ccg***AAGCTT**TCACCCCGCCCCTCGCACGAGATC-3' (Integrated DNA Technologies), where italicized text indicates an overhang and bold text indicates a restriction digest site. The PCR reaction mixture contained 1.25 U of *Pfu* Turbo DNA polymerase (Stratagene), 1x cloned Pfu buffer (Stratagene, 20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/mL bovine serum albumin), 5% DMSO, 250 μM dNTPs (an equimolar mixture dATP, dCTP, dGTP, and dTTP), 0.25 μL of purified genomic DNA, and 0.5 mM of each primer. The 50 μL reaction mixtures were subjected to the following PCR cycle: 98 °C (3 min); 30 cycles of 98 °C (1 min), 52, 59, 66 or 72 °C (1 min), and 72 °C (2.5 min); 72 °C (10 min). All annealing temperatures (52, 59, 66 and 72 °C) gave successful PCR products, so were all combined. The PCR fragment was gel purified using the QIAquick Gel Extraction Kit (Qiagen) and digested at 37°C with *NdeI* and *HindIII* (New England Biolabs) in NEBuffer 2 (New England Biolabs, 10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9). Purified pET28a (Novagen) was identically digested with *NdeI* and *HindIII* (New England Biolabs). Both digested PCR product and pET28a were again gel purified; the *inkE* fragment was then concentrated five-fold by lyophilization and re-suspended in water. The *inkE* insert was ligated into the cut pET28a vector using Quick Ligase (New England Biolabs, 2000 U) in 1x Quick Ligase buffer (New England Biolabs, 66 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM dithiothreitol, 1 mM ATP, 7.5% polyethylene glycol (PEG 6000)), with incubation at 25 °C for 3 h. Ligation reactions were used to transform TOP10 chemically competent *E. coli* cells, which were then plated onto LB-agar-kanamycin plates. Resulting colonies were amplified in LB-kanamycin, and the plasmid purified (QIAprep Spin Miniprep Kit) then analyzed by

restriction enzyme digestion (*NdeI*/*HindIII*) for presence of the insert. The sequence integrity was confirmed by DNA sequencing (Molecular Biology Core Facility, Dana-Farber Cancer Institute, Boston).

AtmC was amplified from *Actinomadura melliaura* (Gao et al., 2006) genomic DNA (a kind gift of Carl Balibar) via PCR. A reaction mixture with 2.5 U of *Pfu* Turbo DNA Polymerase (Stratagene), 1x cloned *Pfu* buffer (Stratagene, 20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/mL bovine serum albumin), 5% DMSO, 300 μM dNTPs (an equimolar mixture dATP, dCTP, dGTP, and dTTP), 1 μL of purified genomic DNA, and 125 ng of each primer (5'-*caagtta***CATATG**ACCACGGCTTACGAGACCGA-3' and 5'-*caagtta***GAATTC**CCATGAAGACCAGCCAGTTCTCCA-3', purchased from Integrated DNA Technologies) in a 50 μL reaction mixture was subjected to the following PCR cycle: 95°C (2 min); 30 cycles of 95°C (30 sec), 57°C (30 sec), and 72°C (3 min); and 72°C (10 min). The PCR fragment was gel purified using the QIAquick Gel Extraction Kit (Qiagen) and digested with *NdeI* and *EcoRI* (New England Biolabs) at 37°C in NEBuffer EcoRI (New England Biolabs, 50 mM NaCl, 100 mM Tris-HCl, 10 mM $MgCl_2$, 0.025% Triton X-100, pH 7.5). Purified pET28a (Novagen) was identically digested with *NdeI* and *EcoRI* (New England Biolabs). Both digested PCR product and pET28a were again gel purified. Various ratios of insert to vector were ligated in a 10 μL reaction volume using 1x T4 DNA ligase buffer (New England Biolabs, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 μg/mL bovine serum albumin) and 400 U of T4 DNA ligase (New England Biolabs), with incubation at room temperature for 50 min. Ligation reactions were transformed into chemically competent DH5α cells and plated onto LB-agar-kanamycin plates. Resulting colonies were amplified in LB with 30 mg/L kanamycin, and the plasmid was purified (QIAprep Spin Miniprep Kit) and assayed via restriction digestion for incorporation of the insert. The

integrity of the sequence was assayed with sequencing reactions spanning the length of the insert. Sequencing was carried out at the MIT Biopolymers Laboratory.

Generation of the StaC-10x expression vector

Twenty-eight overlapping primers encoding the N-terminal portion of the StaC protein with ten amino acid substitutions were designed using DNAWorks 2.4 (http://mcl1.ncifcrf.gov/dnaworks/dnaworks2.html) and synthesized by Integrated DNA Technologies. Primers were designed with each codon for a modified amino acid in a non-overlapping portion of a primer, such that each primer encoding a modified amino acid could be singularly 'swapped' with a primer encoding the wild-type amino acid (Figure S9). This primer design scheme enables facile production of modified StaC proteins with fewer unwanted mutations.

Primer Number	**Sequence**[1]
1	5'-*ggagag***CATATG**ACCCACTCTGGCGAACGTACGGATGTT-3'
2	5'-TACCGACTGGACCACCACCAACGATTAAAACATCCGTACGTTCGC-3'
3	5'-GGTGGTCCAGTCGGTATGGCACTGGCGTTAGACTTACGCTACCGC-3'
4	5'-GACAACTAAACAATCAATGCCGCGGTAGCGTAAGTCTAACGCCAG-3'
5	5'-TGATTGTTTAGTTGTCGAACAGACGGATGGCACGGTCCGTCATCC-3'
6	5'-ATAGAACGCGGACCAATGGTGCCAACACGTGGATGACGGACCGTG-3'
7	5'-ATTGGTCCGCGTTCTATGGAGCTGTTCCGCCGTTGGGGTGCAGCA-3'
8	5'-CCGCTGGCCAACCGGCGTTACGGATTGCGTCTGCTGCACCCCAAC-3'
9	5'-GGTTGGCCAGCGGATCACCCATTAGATATTGCGTGGGTGACCAAG-3'
10	5'-ACGATAACGATAGATCTCATGACCGCCGACCTTGGTCACCCACGC-3'
11	5'-TCATGAGATCTATCGTTATCGTCGCGGTACGGCAGCGAATCGTCC-3'
12	5'-TGGTCTGGTTCCGGGGTGTGGACAAACGCTGGACGATTCGCTGCC-3'
13	5'-CCCGGAACCAGACCAAATCTGTCCGCAACACTGGTTAAACCCAGT-3'
14	5'-CTGGGTGCACGCCCACCGCCTCAATCAGCACTGGGTTTAACCAGT-3'
15	5'-GTGGGCGTGCACCCAGACGGCCCACTGTTATTATCTACGACCGTT-3'
16	5'-TCGTCGGTTTGGACGACGCCGTCAACGGTCGTAGATAATAACAGT-3'
17	5'-CGTCCAAACCGACGACCATGTCGAGGCGACGCTGACCGATCACGC-3'
18	5'-CGGGCACGCACGGTGCCGGTGGTACCGGTGGCGTGATCGGTCAGC-3'
19	5'-CGTGCGTGCCCGCTTTCTGGTTGCCTGCGACGGCGCAAGCTCTCC-3'
20	5'-TGGGGCTTCAATACCACAGGCACGACGAACCGGAGAGCTTGCGCC-3'
21	5'-GTGGTATTGAAGCCCCAGCACGCCATCGCACGCAGGTCTTCCGTA-3'
22	5'-TTTTAACTCCGGGGCACGGAATAAGATATTACGGAAGACCTGCGT-3'
23	5'-GTGCCCCGGAGTTAAAAGATCGCCTGGGTGAACGTGCGGCGCTG -3'
24	5'-AGCGTAACGTGCTGCTTAACATTAAGAAATGAAACAGCGCCGCAC-3'

25	5'-GCAGCACGTTACGCTTTCCATTACGCGCATTAAATGGTAGCGATT-3'
26	5'-ATCATCTGCACCAACCGTCAGACGGTATAAATCGCTACCATTTAA-3'
27	5'-GTTGGTGCAGATGATGATACGGGTGCCCGTCCGGATGCATTAGCA-3'
28	5'-*cgtcct***TGATCA**GTGCTAATGCATCCGGACGG-3'

[1]Bold text within oligonucleotide sequences indicates restriction digest sites, italicized text indicates modified overhangs, and underlined text indicates modified codons. Note that even numbered primers are reverse primers.

To generate the StaC-10x expression vector, a first PCR was carried out using 1x cloned *Pfu* buffer (Stratagene, 20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/mL bovine serum albumin), 2.5 U of *Pfu* Turbo (Stratagene), 250 μM of dNTPs (an equimolar mixture of dATP, dCTP, dGTP, and dTTP), and 12.5 nM of each of the 28 primers, with a temperature cycle of 94°C (2 min); 45 cycles of 94°C (30 s), 58°C (30 s), and 72°C (1 min); and 72°C (2 min). A second PCR was then carried out using 1x cloned *Pfu* buffer, 2.5 U of *Pfu* Turbo, 250 μM dNTPs, 300 nM of each primer 1 and 28, and 1 μL of the first PCR, in a total volume of 50 μL using an identical temperature cycle. The product of the second PCR was gel purified using a QIAquick Gel Extraction Kit (Qiagen). Gel purified DNA from the second PCR was then incubated with 20 U of *Nde*I (New England Biolabs) in NEBuffer 3 (New England Biolabs, 50 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT, pH 7.9) at 37°C for 2 h; then, after addition of 15 U of *Bcl*I (New England Biolabs), the reaction was incubated at 50°C for 2 h.

Separately, pET28a-StaC (Howard-Jones and Walsh, 2006) was transformed into *dam*$^-$/*dcm*$^-$ *E. coli* competent cells (New England Biolabs) and purified using a minprep kit (Qiagen). Purified plasmid was incubated with 20 U of *Hind*III (New England Biolabs) in NEBuffer 2 (New England Biolabs, 10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1mM DTT, pH 7.9) at 37°C for 2 h; then, after adjustment of the buffer to 50 mM Tris-HCl and 100 mM NaCl and addition of 15 U of *Bcl*I, incubated at 50°C for 2 h. Additionally, purified plasmid was separately incubated with 20 U of *Nde*I and 20 U of *Hind*III in NEBuffer 2 (New England Biolabs) at 37°C for 2 h, followed by heat inactivation of the restriction enzymes at 65°C for 20 min. 2.5 U of calf alkaline phosphatase (New England Biolabs) was then added to this final reaction mixture, which was incubated at 37°C for 2 h.

Appropriate fragments from the restriction digests were gel purified. Ligation reactions were set up with 20 ng of the purified *Nde*I/*Hind*III digest reaction, 8 ng of the purified *Hind*III/*Bcl*I reaction, 8 ng of the purified *Nde*I/*Bcl*I digest reaction, 1x T4 DNA ligase buffer (New England Biolabs, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 μg/mL bovine serum albumin), and 200 U of T4 DNA ligase (New England Biolabs) and incubated at 16°C for 24 h. Ligation reactions were transformed into DH5α cells and plated on LB-agar-kanamycin plates. Plasmid was purified from overnight cultures of single colonies, analyzed for the presence of the PCR insert via a diagnostic PCR reaction, and sequenced at the MIT Biopolymers Laboratory using three different primers to gather data from the entire length of the *staC-10x* gene.

Generation of the RebC-10x expression vector

Forty-one overlapping primers encoding the N-terminal portion of the RebC protein with ten amino acid substitutions were designed using DNAWorks 2.4 (http://mcl1.ncifcrf.gov/dnaworks/dnaworks2.html) and synthesized by Integrated DNA Technologies.

Primer Number	**Sequence**[1]
1	5'-*ggagag***CATATG**AATGCCCC-3'
2	5'-CCAGGATCAGGACATCGGTTTCAATTGGGGCATTCATATGCTCTC-3'
3	5'-CGATGTCCTGATCCTGGGTGGCGGTCCAGTTGGTATGGCCTTAGC-3'
4	5'-CCCACCTGGCGATGGGCCAGGTCCAGGGCTAAGGCCATACCAACT-3'
5	5'-CATCGCCAGGTGGGTCACTTAGTCGTTGATGCGGGCGACGGCACC-3'
6	5'-TGGACCGATCGTAGAAACTTTTGGGTGCGTGATGGTGCCGTCGCC-3'
7	5'-GTTTCTACGATCGGTCCACGTAGCATGGAATTATTCCGTCGTTGG-3'
8	5'-GCCGTGCGAATCTGTTTTGCGACACCCCAACGACGGAATAATTCC-3'
9	5'-ACAGATTCGCACGGCGGGTTGGCCGGGTGACCACCCGTTAGACGC-3'
10	5'-TCATGGCCACCGACACGCGTCACCCATGCTGCGTCTAACGGGTGG-3'
11	5'-TGTCGGTGGCCATGAAGTTTACCGCATTCCATTAGGTACCGCGGA-3'
12	5'-CCGGGGTGTGTTCCGGCGTTGCACGCGTATCCGCGGTACCTAATG-3'
13	5'-GGAACACACCCCGGAACCAGATGCGATCTGTCCAGCGCATTGGTT-3'
14	5'-TCGCCCACTGCCTCCGCCAGCAGCGGGGCTAACCAATGCGCTGGA-3'
15	5'-AGGCAGTGGGCGAGCGTTTACGTACGCGCTCTCGCTTAGACTCTT-3'
16	5'-CGCGCACGTGGTCGTCACGCTGTTCGAAAGAGTCTAAGCGAGAGC-3'
17	5'-GACCACGTGCGCGCCACGATTACGGATCTGCGTACGGGTGCAACC-3'
18	5'-GGCAACCAGGTAGCGCGCGTGAACTGCACGGGTTGCACCCGTACG-3'
19	5'-CGCTACCTGGTTGCCTGTGACGGCGCGTCTTCTCCGACCCGCAAG-3'
20	5'-GATGACGTGGTGGGGCGTCAATACCTAACGCCTTGCGGGTCGGAG-3'
21	5'-CCCCACCACGTCATCGCACGCAGGTCTTTCGTAATATTCTGTTCC-3'
22	5'-AAGCTGCGTAACTCCGGTGCACGGAACAGAATATTACGAAAGACC-3'

23	5'-CGGAGTTACGCAGCTTACTGGGCGAACGTGCGGCACTGGTGTTCT-3'
24	5'-GAAAACGTAAAGAAGAGCTCAGCATTAAGAAGAACACCAGTGCCG-3'
25	5'-TGAGCTCTTCTTTACGTTTTCCGTTACGCTCTCTGGATGGCCGTG-3'
26	5'-CGTCCACGCCCACAACTAAGTTGTATAAACCACGGCCATCCAGAG-3'
27	5'-GTGGGCGTGGACGATGCCTCTAAGTCTACCATGGACAGCTTTGAA-3'
28	5'-TCGAACGCCACTGCGCGACGAACTAATTCAAAGCTGTCCATGGTA-3'
29	5'-GCAGTGGCGTTCGATACGGAGATTGAAGTGTTATCTGATTCTGAA-3'
30	5'-ACGCGGTGCGTTAAGTGCCATTCAGAATCAGATAACACTTCAATC-3'
31	5'-CTTAACGCACCGCGTCGCGGATTCTTTCTCTGCGGGCCGTGTTTT-3'
32	5'-GCTTAAGGTATGTGCTGCATCGCCCGTCAGGAAAACACGGCCCGC-3'
33	5'-GCAGCACATACCTTAAGCCCAAGCGGTGGTTTCGGCATGAACACG-3'
34	5'-AACCCAGGTCGGCTGCAGAGCCAATGCCCGTGTTCATGCCGAAAC-3'
35	5'-AGCCGACCTGGGTTGGAAGTTAGCAGCGACGCTGCGCGGCTGGGC-3'
36	5'-TTCCTCCTCGTAGGTCGCTAATAAGCCCGGACCGGCCCAGCCGCG-3'
37	5'-CGACCTACGAGGAGGAACGTCGTCCGGTGGCGATTACCAGCCTGG-3'
38	5'-CGTACGGCGCAGGTTAACATTCGCCTCCTCCAGGCTGGTAATCGC-3'
39	5'-CCTGCGCCGTACGATGGATCGTGAGTTACCACCGGGTCTGCACGA-3'
40	5'-CCAAGAGGATCCCTCGCCGCGTGGGCCATCATCGTGCAGACCCGG-3'
41	5'-*ccaaga***GGATCC**CTCGCC-3'

[1]Bold text within oligonucleotide sequences indicates restriction digest sites, italicized text indicates modified overhangs, and underlined text indicates modified codons. Note that even numbered primers are reverse primers.

A first PCR was carried out using 1x cloned *Pfu* buffer (Stratagene, 20 mM Tris-HCl pH 8.8, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 μg/mL bovine serum albumin), 2.5 U of *Pfu* Turbo (Stratagene), 250 μM of dNTPs (an equimolar mixture of dATP, dCTP, dGTP, and dTTP), and 12.5 nM of each of the 41 primers, with a temperature cycle of 94°C (2 min); 45 cycles of 94°C (30 s), 58°C (30 s), and 72°C (2 min); and 72°C (10 min). A second PCR was then carried out using 1x cloned *Pfu* buffer, 2.5 U of *Pfu* Turbo, 250 μM dNTPs, 7.5 μM of each primer 1 and 41, and 1 μL of the first PCR, in a total volume of 50 μL using an identical temperature cycle. The product of the second PCR was gel purified using a QIAquick Gel Extraction Kit (Qiagen). Gel purified DNA from the second PCR was then incubated with 20 U of *Nde*I (New England Biolabs) and 20 U of *Bam*HI (New England Biolabs) in NEBuffer 3 (New England Biolabs, 50 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT, pH 7.9) and 0.1 mg/mL bovine serum albumin (New England Biolabs) at 37°C overnight. Restriction digest reactions were then heat inactivated at 65°C for 20 min. Similarly, pET28a-RebC was incubated with 20 U of *Nde*I (New England Biolabs) and 20 U of *Bam*HI (New England Biolabs) in NEBuffer 3 (New England Biolabs, 50 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT, pH 7.9) and 0.1 mg/mL bovine serum albumin (New England

Biolabs) at 37°C overnight. Restriction digest reactions were then heat inactivated at 65°C for 20 min and then cooled on ice. 5 U of Antarctic Phosphatase (New England Biolabs) and 1x Antarctic Phosphatase Reaction Buffer (New England Biolabs, 50 mM Bis-Tris-Propane-HCl, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, pH 6.0) were then added to this final reaction mixture, which was incubated at 37°C for 1.5 h. Phosphatase reactions were then heat inactivated at 65°C for 20 min.

Appropriate fragments from the restriction digests were gel purified. Ligation reactions with 20 ng of the purified *Nde*I/*Bam*HI pET28a-RebC digest reaction and 27 ng of the *Nde*I/*Bam*HI cut fragment encoding the N-terminus of RebC-10x were set up in 1x T4 DNA ligase buffer (New England Biolabs, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 μg/mL bovine serum albumin), and 200 U of T4 DNA ligase (New England Biolabs) and incubated at 16°C for 16 h. Ligation reactions were transformed into DH5α cells and plated on LB-agar-kanamycin plates. Plasmid was purified from overnight cultures of single colonies, analyzed for the presence of the PCR insert via a diagnostic PCR reaction, and sequenced at the MIT Biopolymers Laboratory using two primers to gather data from the entire length of Reb-10x gene. We observed from this sequence that our design of the fragment encoding the N-terminal fragment of RebC-10x inadvertently introduced a frame shift mutation. Therefore we used QuickChange (Stratagene) to introduce the missing cytosine, generating the final r*ebc-10x* construct. Plasmid was again sequenced at the MIT Biopolymers Laboratory.