Synthesis and Characterization of Cytidine Derivatives that Inhibit the Kinase IspE of the Non-Mevalonate Pathway for Isoprenoid Biosynthesis

Crane, Christine M.; Hirsch, Anna; Alphey, Magnus S.; Sgraja, Tanja; Lauw, Susan; Illarionova, Victoria; Rohdich, Felix; Eisenreich, Wolfgang; Hunter, William N.; Bacher, Adelbert

Published in:
ChemMedChem

DOI:
10.1002/cmdc.200700208

Publication date:
2008

CHEMMEDCHEM
Chemistry & Drug Discovery

Supporting Information

Synthesis and Characterization of Cytidine Derivatives that Inhibit the Kinase IspE of the Non-Mevalonate Pathway for Isoprenoid Biosynthesis

by **Christine M. Crane, Anna K. H. Hirsch, Magnus S. Alphey, Tanja Sgraja, Susan Lauw, Victoria Illarionova, Felix Rohdich[*], Wolfgang Eisenreich, William N. Hunter[*], Adelbert Bacher, François Diederich[*]**

Supporting Information

Table of Contents

NMR experiments

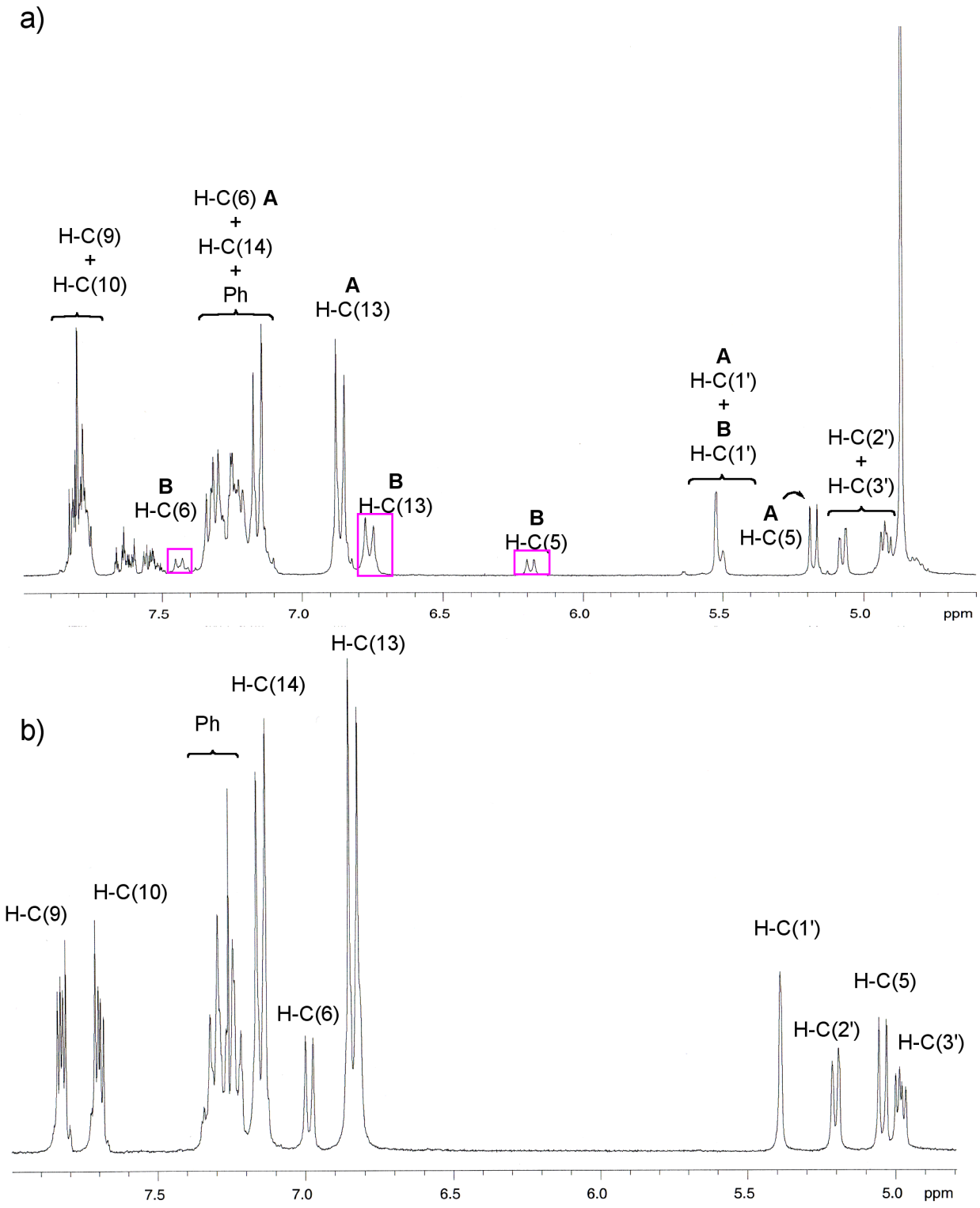


Figure 1SI. a) 300 MHz ^{1}H NMR spectrum (298 K, CD_3OD, region from 5.0 to 8.0 ppm) of (+)-**8** showing the occurrence of two isomers (A/B = 3/1, labeled arbitrarily, without implying a conformational assignment). b) In $CDCl_3$ (300 MHz, 298 K), the rotation is fast, presumably catalyzed by residual traces of HCl, and only one set of signals is observed.

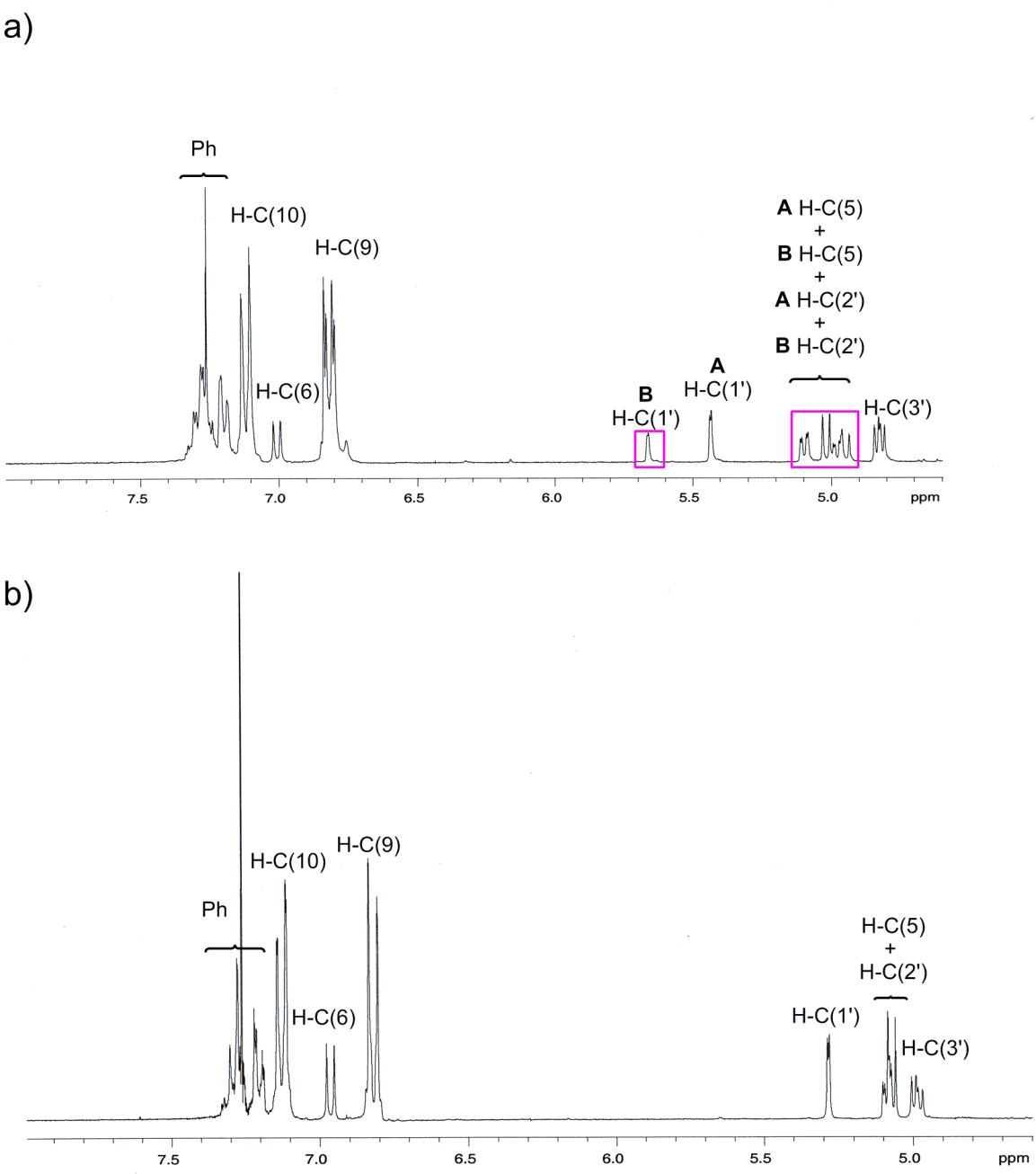


Figure 2SI. a) 300 MHz ^{1}H NMR spectrum (298 K, $CDCl_3$, region from 5.0 to 8.0 ppm) of (+)-**9** showing the occurrence of two isomers (A/B = 3/1, labeled arbitrarily, without implying a conformational assignment). b) 300 MHz ^{1}H NMR spectrum recorded at 298 K in $CDCl_3$ in the presence of 1 M solution of TFA in $CDCl_3$ (5 µL) shows only one set of signals, indicating fast equilibration.

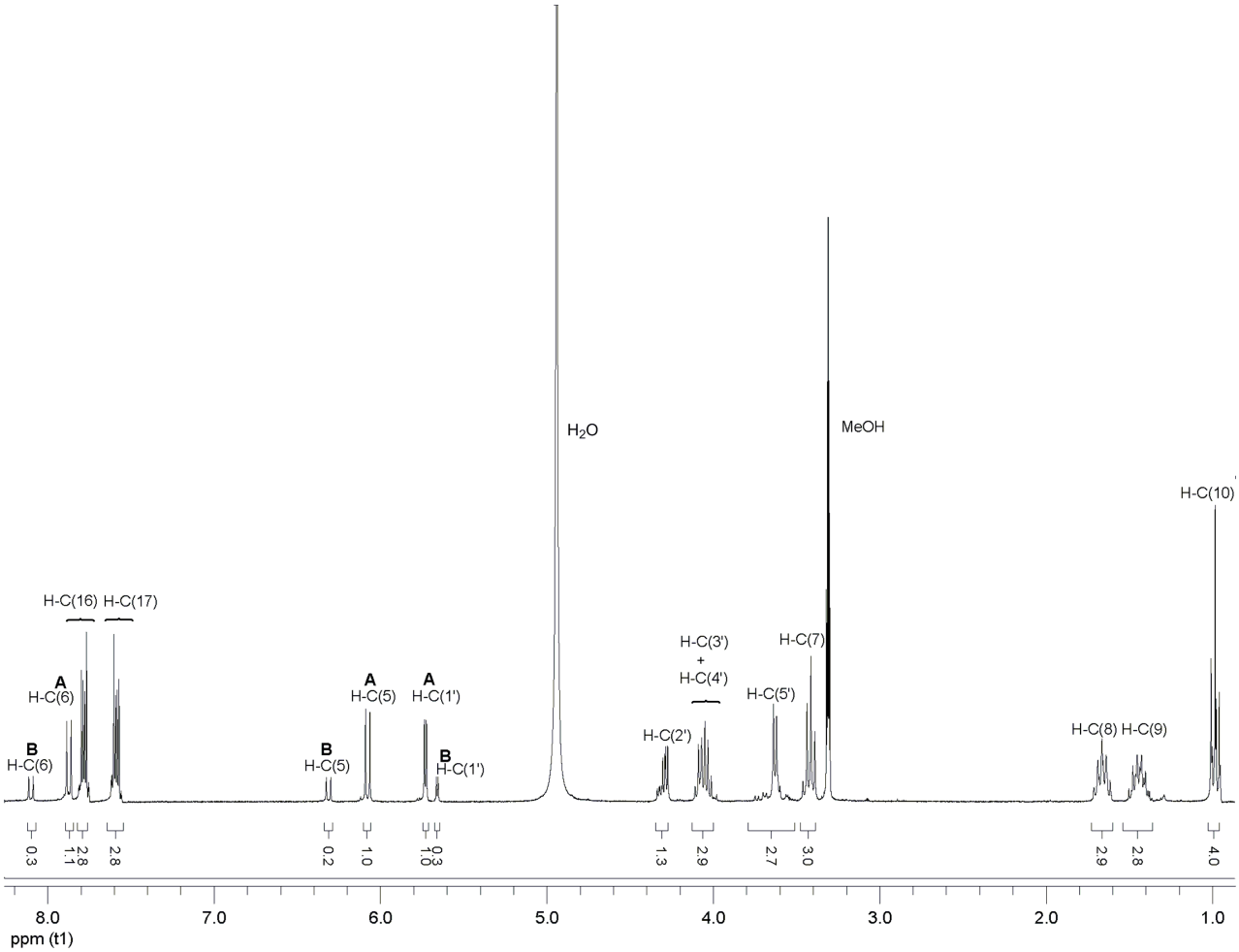


Figure 3SI. 300 MHz ^{1}H NMR spectrum (298 K, CD_3OD) of (-)-**2** showing the occurrence of two isomers (A/B = 3/1, labeled arbitrarily, without implying a conformational assignment). Exchangeable protons are not visible (including the acidic methylene protons (H-C(12) of the benzimidazole linker).

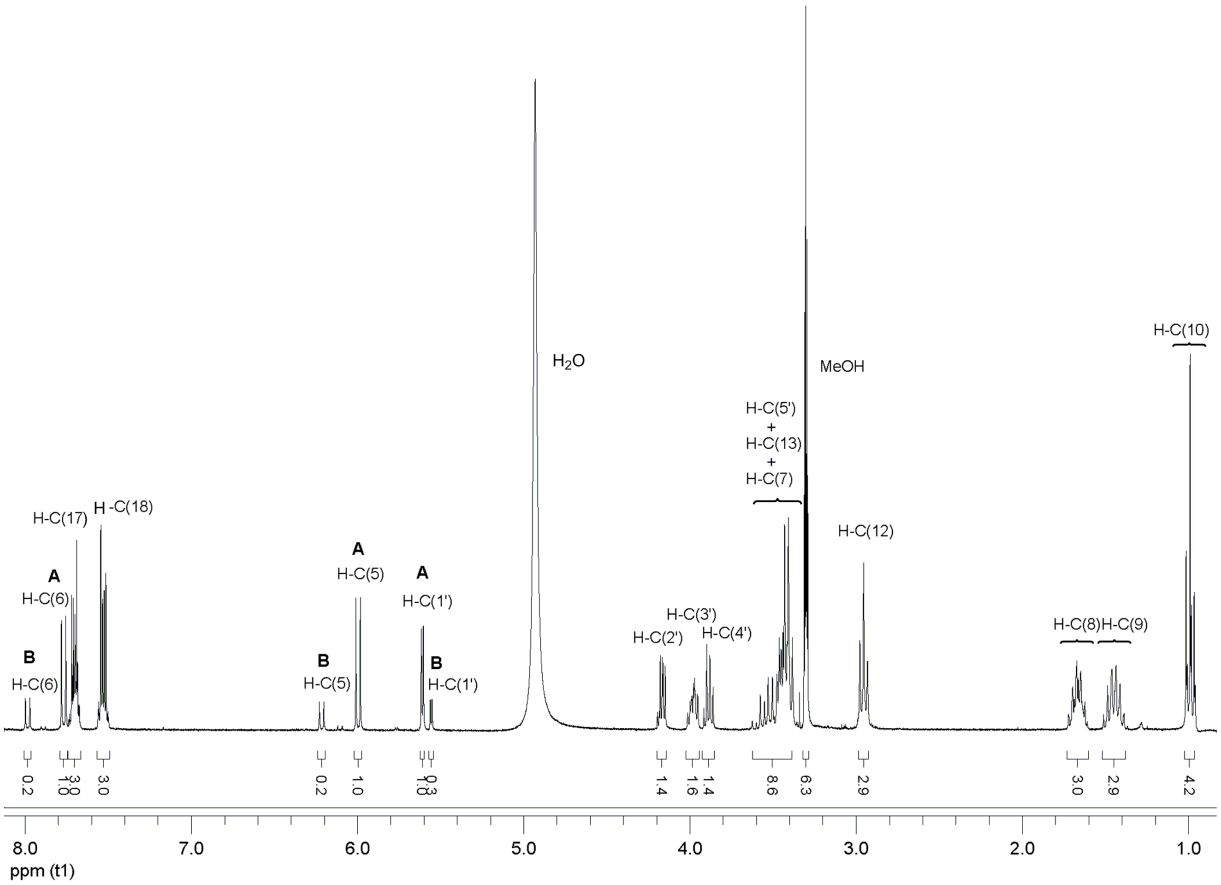


Figure 4SI. 300 MHz ^{1}H NMR spectrum (298 K, CD_3OD) of (+)-**4** showing the occurrence of two isomers (A/B = 3/1, labeled arbitrarily, without implying a conformational assignment).

Biological results

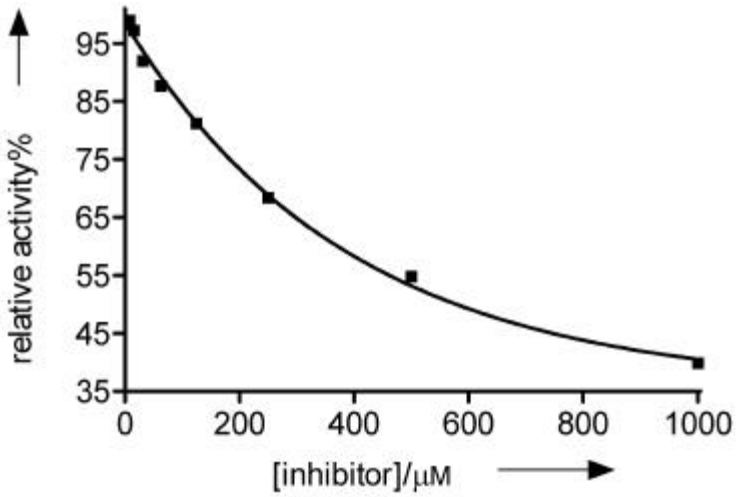


Figure 5SI. Exemplary IC_{50} curve for inhibition of IspE by (-)-**1**. Measured at [CDP-ME] = 1 mM.[1]

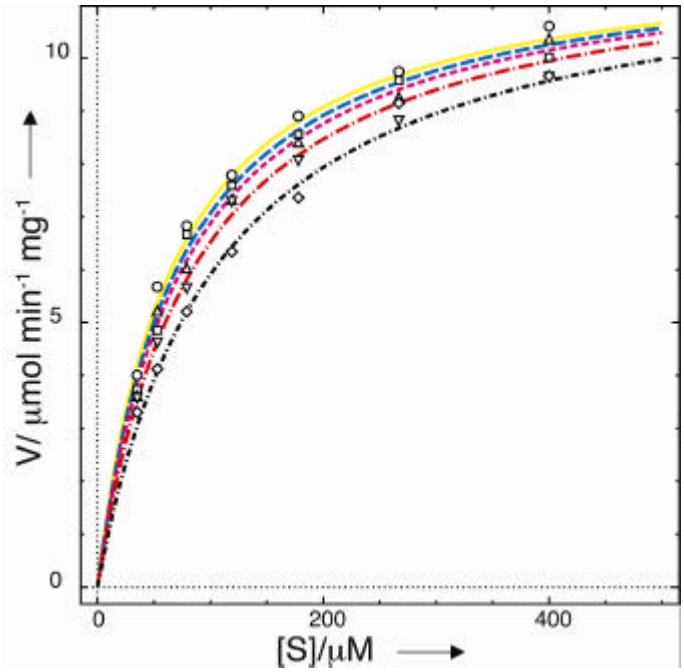
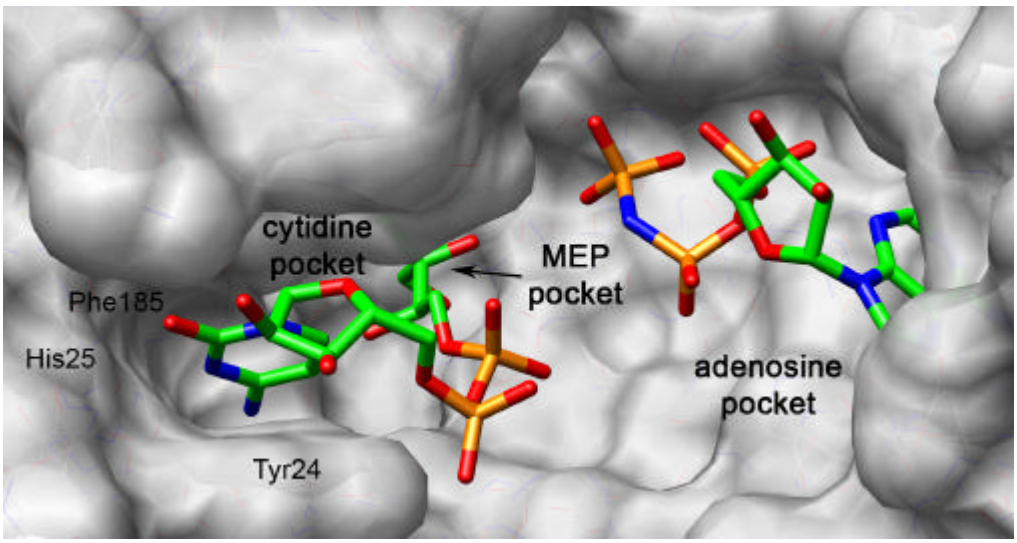


Figure 6SI. The curves for inhibition of IspE by (-)-**1** used to calculate the K_i value. The concentration of (-)-**1** was varied from 3.1 to 50 mM.[1]

a)



b)

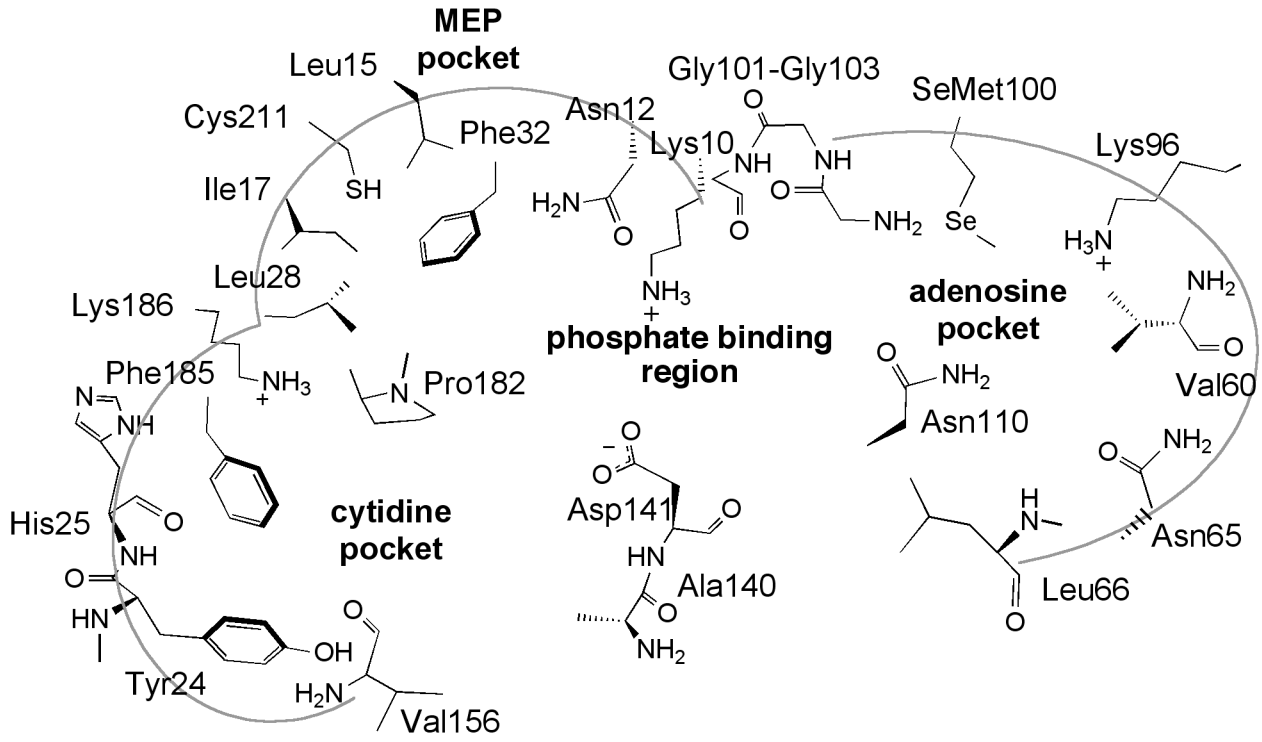


Figure 7SI. Active site of *E. coli* IspE as revealed in the ternary complex with CDP-ME and 5'-adenyl-β,γ-amidotriphosphate (AppNp) (PDB code: 1OJ4).[2] The two nucleotides bind on opposite sides of a solvent-exposed space, normally housing the phosphate moieties of the substrates. a) Stick representation of CDP-ME and AppNp in the active site with the protein depicted as a surface. Color code: substrate skeleton: C: green; O: red; N: blue; P: orange; protein surface: grey. This color code is maintained throughout the article. b) Schematic representation of active site residues. All figures were generated with the molecular graphics program Chimera.[3]

Recognition of the crystallization buffer components

The complexes with (-)-**1** and (+)-**3** feature two co-crystallized sulfate ions in active site A and an additional sulfate ion in active site B. In addition, a glycerol molecule co-crystallized in the adenosine binding region of both structures (Figures 8SI-9SI).

In active site A of the complex with (-)-**1** two co-crystallized sulfate ions are observed. The first resides in the position normally occupied by the β-phosphate of CDP-ME and is H-bonded to Ala129 ($d(N_{Ala}\cdots O)$ = 2.7 Å). The second is bound in the region where the γ-phosphate of AppNp binds and is H-bonded to Gly90 ($d(N_{Gly}\cdots O)$ = 3.1 Å), Gly92 ($d(N_{Gly}\cdots O)$ = 2.8 Å), and several water molecules ($d(O\cdots O)$ = 2.2, 2.7, 2.9 Å) (Figure 2a).

a) b)

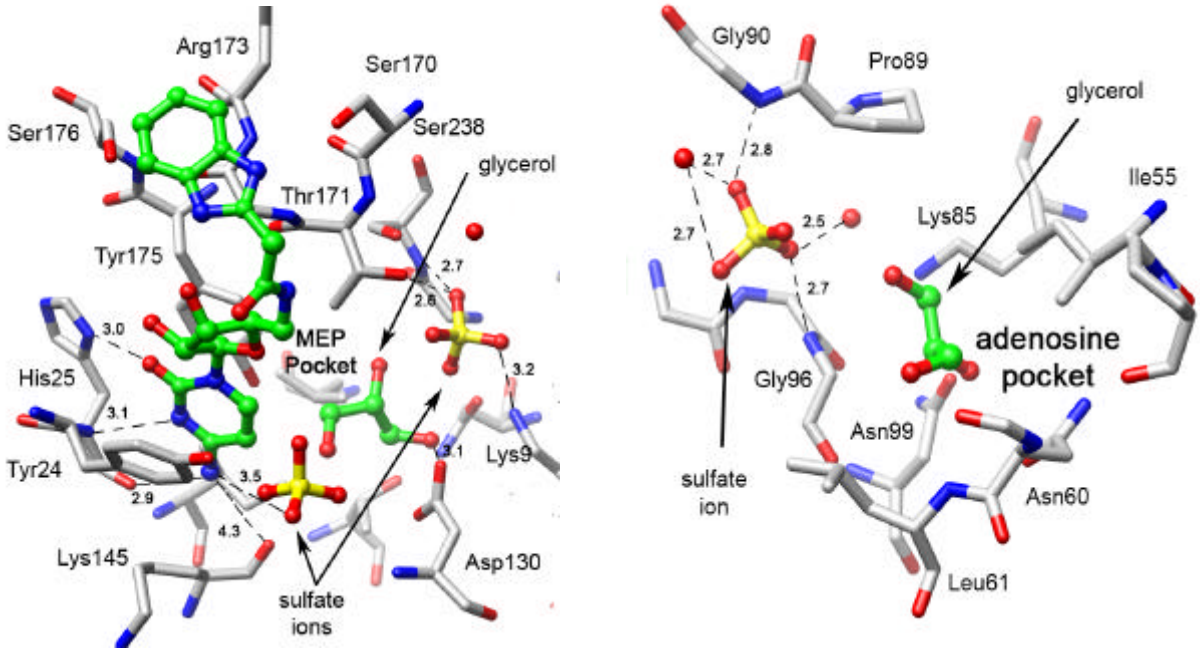


Figure 8SI. Binding mode of inhibitor (-)-**1** in active site B of the complex with *A. aeolicus* IspE solved to 2.4 Å resolution. a) Ball-and-stick representation of (-)-**1** in the cytidine binding pocket. Two sulfate ions are bound in this region of the active site. The first is located in the catalytic domain near residues Lys9 ($d(N_{Lys}\cdots O)$ = 3.2 Å), Thr171 ($d(O_{Thr}\cdots O)$ = 2.6 Å), and Ser238 ($d(N_{Ser}\cdots O)$ = 2.7 Å, $d(O_{Ser}\cdots O)$ = 2.9 Å) and the second is near residue Tyr24 ($d(O_{Tyr}\cdots O)$ = 3.5 Å). A glycerol molecule is also bound in the MEP pocket near the catalytic Asp130 ($d(O\cdots O)$ = 3.1 Å). b) Ball-and-stick representation of the third sulfate ion bound in the adenosine pocket. It is located near residues Gly90 ($d(N_{Gly}\cdots O)$ = 2.8 Å) and Gly96 ($d(N_{Gly}\cdots O)$ = 2.7 Å) and two water molecules ($d(O\cdots O)$ = 2.5 Å, $d(O\cdots O)$ = 2.7 Å). H-bonds are represented as dashed lines. Distances between heavy atoms are given in Å.

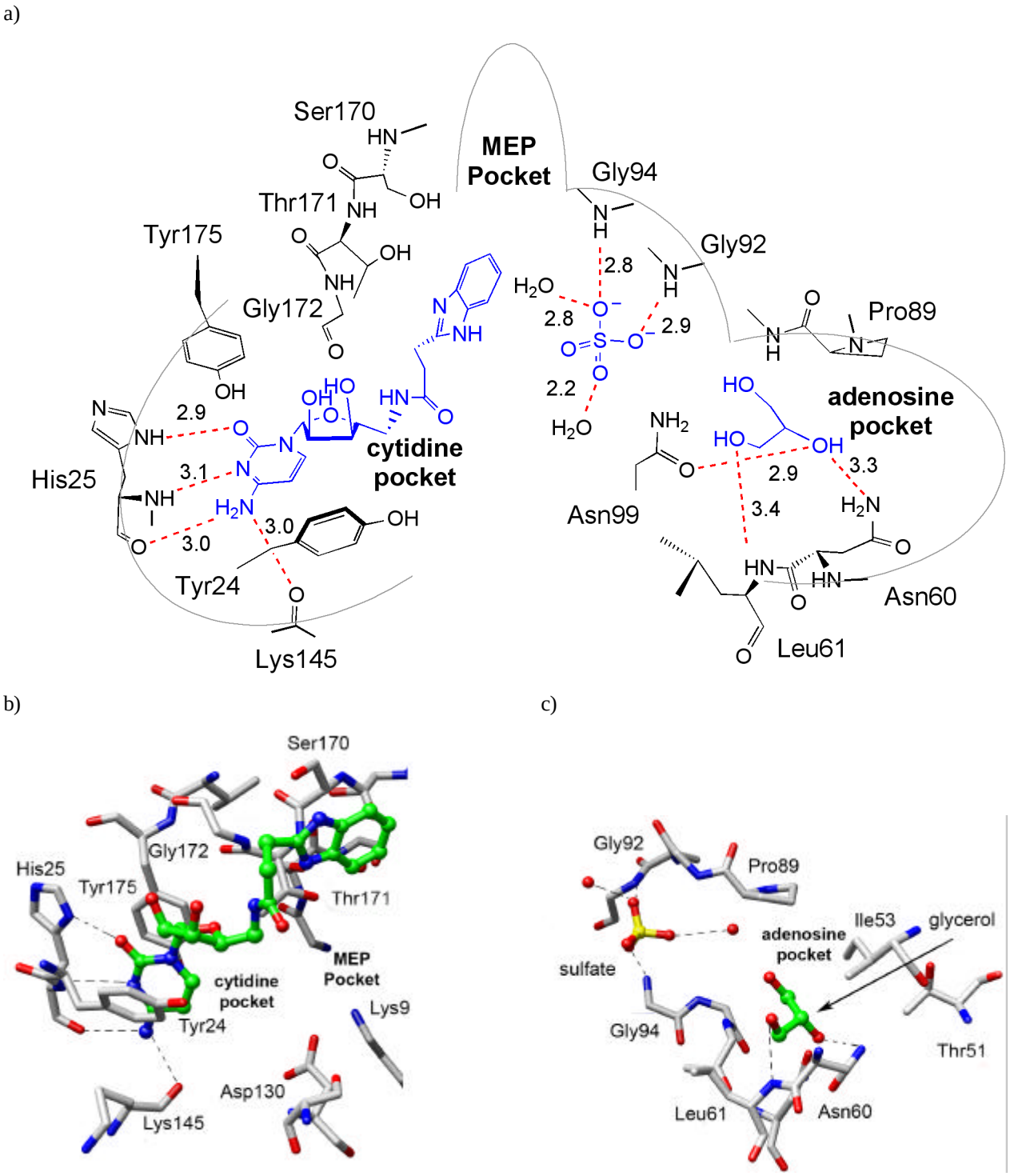
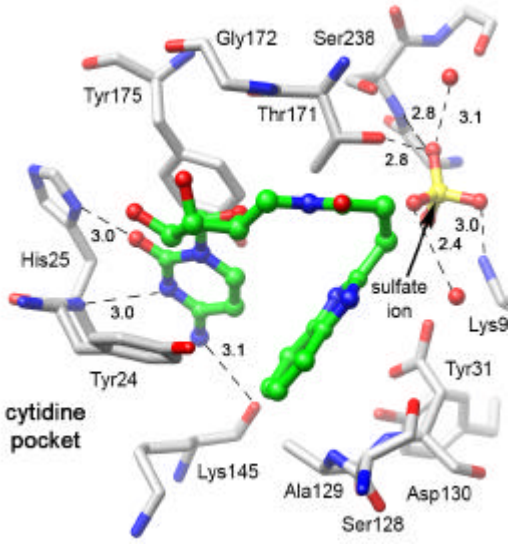


Figure 9SI. Binding mode of (+)-**3** in active site A of *A. aeolicus* IspE (PDB code: 2V2V) as determined by X-ray crystallography to 2.3 Å resolution. In addition to the ligand, a sulfate ion and a glycerol molecule are also bound. a) Schematic representation of the binding mode of (+)-**3** and the co-crystallized sulfate and glycerol molecules (blue) with H-bonds depicted as red dashed lines. Distances between heavy atoms are given in Å. b) Ball-and-stick representation of the binding mode of (+)-**3** in the cytidine binding pocket. The cytosine base, sandwiched between Tyr24 and Tyr175, is H-bonded to the side chain and backbone of His25 ($d(N_{His}\cdots O)$ = 2.9 Å, $d(N_{His}\cdots N)$ = 3.1 Å, $d(O_{His}\cdots N)$ = 3.0 Å) and Lys145 ($d(O_{Lys}\cdots N)$ = 3.0 Å). c) Ball-and-stick representation of the sulfate ion H-bonded to residues Gly92 ($d(N_{Gly}\cdots O)$ = 2.9 Å), Gly94 ($d(N_{Gly}\cdots O)$ = 2.8 Å), and two water molecules ($d(O\cdots O)$ = 2.2 Å, $d(O\cdots O)$ = 2.8 Å) in the adenosine-binding pocket**.** The glycerol molecule is H-bonded to residues Asn60 ($d(N_{Asn}\cdots O)$ = 3.3 Å), Leu61 ($d(N_{Leu}\cdots O)$ = 3.4 Å),

and Asn99 ($d(O_{Asn}\cdots O)$ = 2.9 Å) in the adenine binding region, at the position normally occupied by the γ-phosphate of AppNP. H-bonds are depicted as dashed lines. Color code: protein: Se: cyan; S: yellow.

a)



b)

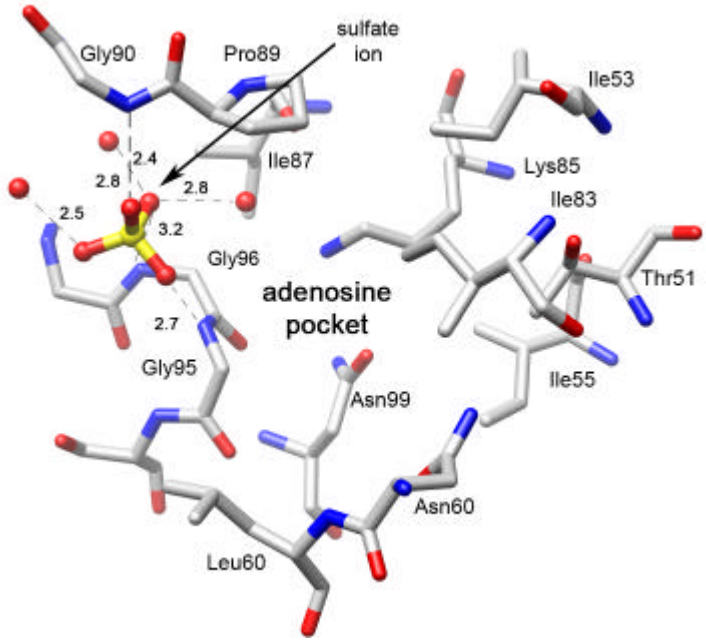
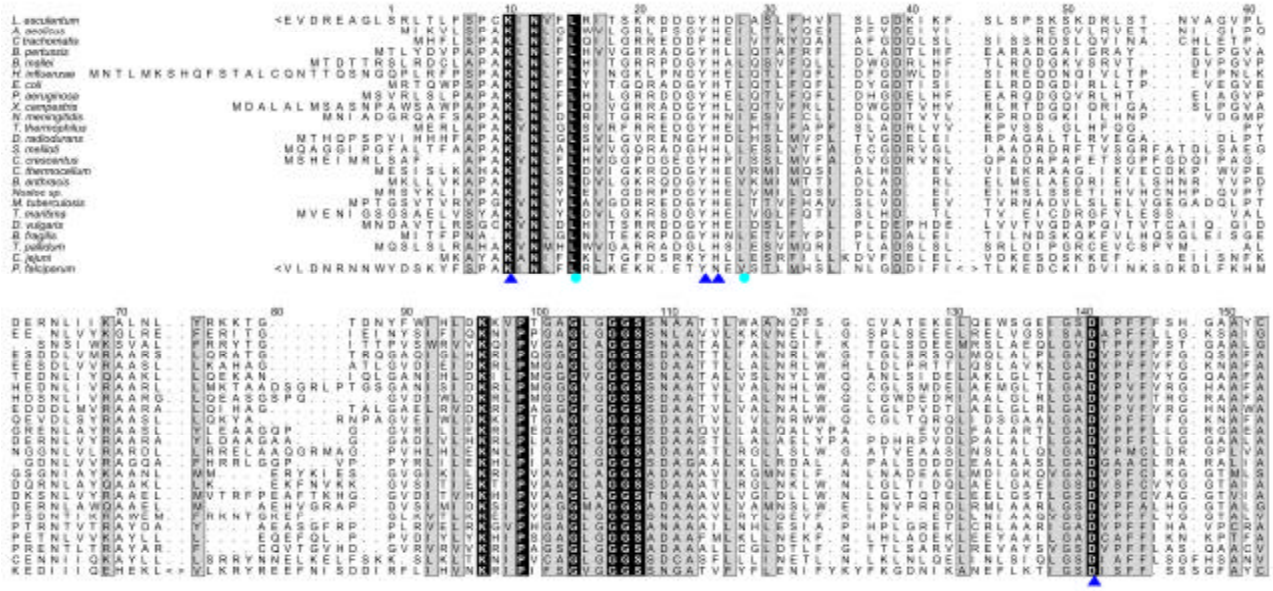


Figure 10SI. Binding mode of inhibitor (+)-**3** in active site B of the complex with *A. aeolicus* IspE solved to 2.3 Å resolution. a) Ball-and-stick representation of (+)-**3** in the observed "collapsed" conformation in the cytidine binding pocket. A single sulfate ion is bound in the catalytic domain of the MEP pocket near residues Lys9 ($d(N_{Lys}\cdots O)$= 3.0 Å), Thr171 ($d(O_{Thr}\cdots O)$ = 2.8 Å), and Ser238 ($d(N_{Ser}\cdots O)$ = 2.8 Å, $d(O_{Ser}\cdots O)$ = 2.9 Å) and two water molecules ($d(O\cdots O)$ = 2.4 Å, 3.1 Å). b) Ball-and-stick representation of a second sulfate ion residing near residues Gly95 ($d(N_{Gly}\cdots O)$ = 2.7 Å), Gly96 ($d(N_{Gly}\cdots O)$ = 3.2 Å), and Gly90 ($d(N_{Gly}\cdots O)$ = 2.4 Å) and three water molecules ($d(O\cdots O)$ = 2.5 Å, $d(O\cdots O)$ = 2.4 Å, $d(O\cdots O)$ = 2.8 Å) in the adenosine pocket. This corresponds to the location of the catalytic γ-phosphate of AppNp. A glycerol molecule is not observed in active site B. H-bonds are represented as dashed lines. Distances between heavy atoms are given in Å.

Sequence alignment

a)



b)

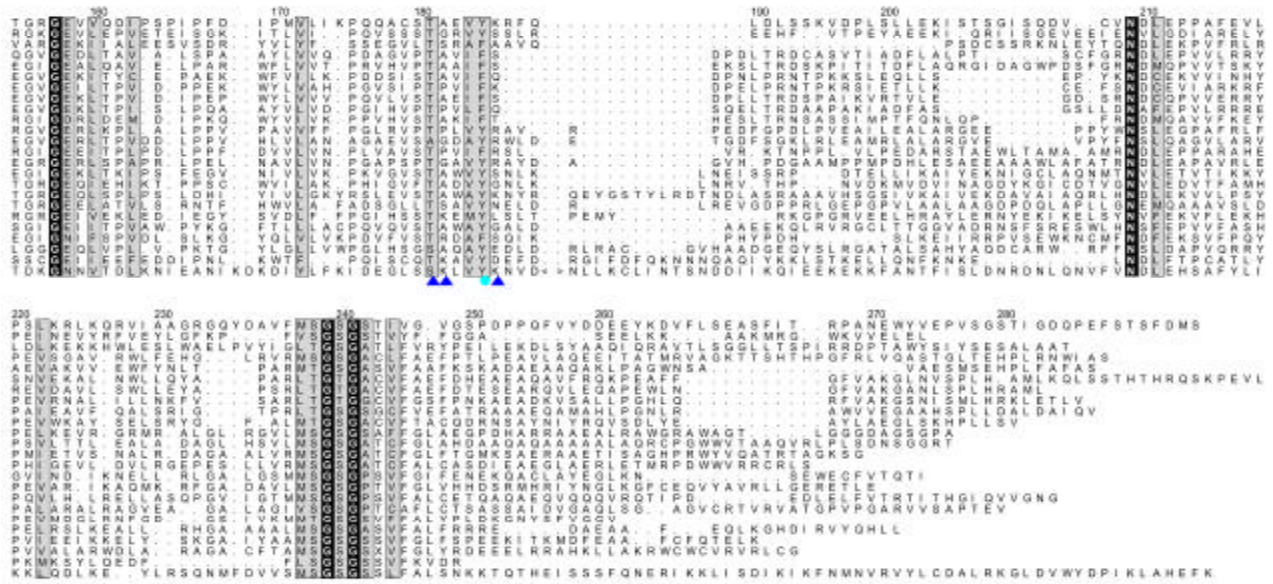


Figure 11SI. Amino acid sequence alignment of IspE proteins from various organisms. The alignment was performed using amino acid sequences of IspE of 22 bacterial groups, one plant (*Lycopersicon esculentum*), and from the apicomplexan (*Plasmodium falciparum*) using the program PileUp (GCG, Madison, Wisconsin). The numbering used is according to the *E. coli* amino acid sequence. Circles (cyan) indicate residues lining the hydrophobic region of the cytidine binding pocket and the arrows (dark blue) indicate important active site residues. a) Amino acid residues 1-153. b) Amino acid residues 154-280.

Experimental section

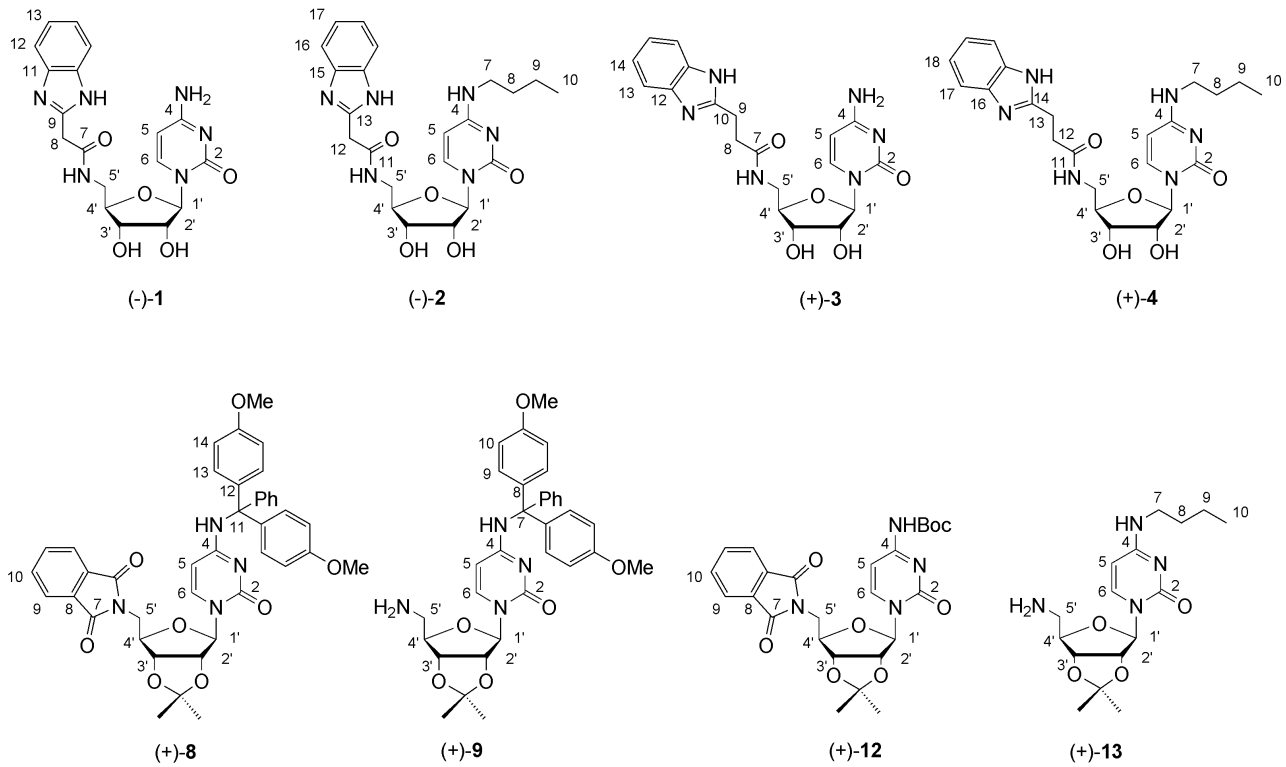


Figure 12SI. Arbitrary labeling of representative compounds.

Table 1SI. Chemical shifts (*d*), coupling pattern, and coupling constants (*J*) for distinguishable resonances in the range of 5.0 – 8.0 ppm of conformers **A** and **B** for compound (+)-**8**, at 298 K. (n.v. = peaks overlap). See Figure 1SI and Experimental Part for assignments.

Resonance		*d* (in CD_3OD)	*d* (in $CDCl_3$)
H-C(1')	**A**	5.53 (s)	5.39 (s)
	B	5.51(s)	n.v.
H-C(2')	**A**	5.08 (dd, *J* = 6.3, 1.0)	5.20 (d, *J* = 6.2)
	B	n.v.	n.v.
H-C(3')	**A**	4.93 (dd, *J* = 6.3, 3.9)	4.98 (dd, *J* = 6.2, 3.6)
	B	n.v.	n.v.
H-C(4')	**A**	4.27-4.33 (m)	4.35-4.41 (m)
	B	n.v.	n.v.
H-C(5')	**A**	4.00 (dd, *J* = 6.9, 2.1)	4.02-4.16 (m)
	B	n.v.	n.v.
H-C(5)	**A**	5.19 (d, *J* = 8.1)	5.04 (d, *J* = 7.2)
	B	6.19 (d, *J* = 7.2)	n.v.
H-C(6)	**A**	n.v.	6.98 (*d*, *J* = 7.5)
	B	7.45 (d, *J* = 7.2)	n.v.
H-C(13)	**A**	6.92 (d, *J* = 9.0)	6.84 (d, *J* = 8.7)
	B	6.77 (d, *J* = 8.7)	n.v.

Table 2SI. Chemical shifts (*d*), coupling pattern, and coupling constants (*J*) for distinguishable resonances of conformers **A** and **B** for compound (+)-**9**, at 298 K. (n.v. = peaks overlap). See Figure 2SI and Experimental Part for assignments.

Resonance		*d* (in $CDCl_3$)	*d* (in $CDCl_3$ + TFA)[a]
H-C(1')	**A**	5.43 (d, *J* = 1.5)	5.28 (d, *J* = 2.1)
	B	5.66 (d, *J* = 1.2)	n.v.
H-C(2')	**A**	5.10 (dd, *J* = 6.6, 1.8)	5.06-5.10 (m)
	B	4.99 (dd, *J* = 6.6, 1.8)	n.v.
H-C(3')	**A**	4.82 (dd, *J* = 6.6, 4.2)	4.98 (dd, *J* = 6.3, 4.2)
	B	n.v.	n.v.
H-C(4')	**A**	4.03 (q, *J* = 5.2)	4.17-4.22 (m)
	B	4.43 (q, *J* = 4.5)	n.v.
H-C(5)	**A**	5.02 (d, *J* = 7.6)	5.06-5.10 (m)
	B	4.95 (d, *J* = 7.8)	n.v.
H-C(6)	**A**	7.01 (d, *J* = 7.6)	6.96 (d, *J* = 7.8)
	B	n.v.	n.v.

[a]in $CDCl_3$ with a 1 M solution of TFA in $CDCl_3$ (5 μL).

Table 3SI. Co-crystal structures: Statistics for data collection and refinement.

Structure	(-)-**1**	(+)-**3**
Space Group	$P2_13$	$P2_13$
Unit cell length (Å)	136.9	137.3
Resolution range (Å)	90.0 - 2.4	96.7 – 2.3
No. of observed / unique reflections	184,802 / 33,908	206,158 / 38,321
Wilson B (Å^2)	51.1	45.1
Completeness (%)	100.0 (100.0)	99.9 (100.0)
Multiplicity/R_{merge} (%)	5.5(5.3)/8.2(58.1)	5.4 (5.3)/8.3 (58.5)
<I/s(I)>	15.8 (2.5)	15.0 (2.7)
R_{work} / R_{free} (%)	21.7(26.8)/27.2(32.8)	21.1(27.4)/23.4(32.4)
r.m.s.d from ideal values, bond lengths (Å)/ bond angles (°)	0.009 / 1.191	0.008 / 1.122
Overall B (Å^2)	40.9	36.4
Main chain	40.6	36.0
Side chain	41.2	36.8
Water molecules	40.4	39.7
Ligands[a]	(SO_4^{2-} 54.5; Br 41.0 glycerol 45.6; (-)-**1** 39.5)	(SO_4^{2-} 53.5; Br 38.3 glycerol 39.3; (+)-**3** 39.0)
Residues in most favorable regions (%)	92.7	93.2
Residues in additionally allowed regions (%)	6.6	6.4
Cruickshanks DPI[b] (Å) based on R_{free}	0.31	0.26

Numbers in parentheses represent the highest resolution bin of width approx. 0.06 Å. [a] The B value is only given for the base and ribose of the inhibitors. [b] Diffraction-component Precision Index (Cruickshank, 1999).

References

[1] V. Illarionova, J. Kaiser, E. Ostrozhenkova, A. Bacher, M. Fischer, W. Eisenreich, F. Rohdich, *J. Org. Chem.* **2006**, *71*, 8824–8834.

[2] L. Miallau, M. S. Alphey, L. E. Kemp, G. A. Leonard, S. M. McSweeney, S. Hecht, A. Bacher, W. Eisenreich, F. Rohdich, W. N. Hunter, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 9173–9178.

[3] a) E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605-1612; b) Chimera home page can be found at http://www.cgl.ucsf.edu/chimera.