Supporting Information

Diterpene Cyclases and the Nature of the Isoprene Fold

Rong Cao[1], Yonghui Zhang[2], Francis M. Mann[3], Cancan Huang[2], Dushyant Mukkamala[1], Michael P. Hudock[1], Matthew Mead[3], Sladjana Prisic[3], Ke Wang[2], Fu-Yiang Lin[1], Ting-Kai Chang[2], Reuben Peters[3,*], and Eric Oldfield[1,2,*]

[1]Center for Biophysics & Computational Biology,

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA;

[2]Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue,

Urbana, IL 61801, USA; [3]Department of Biochemistry, Biophysics and Molecular Biology,

Iowa State University, Ames, IA 50011, USA

---

*To whom correspondence may be addressed. E-mail: rjpeters@iastate.edu or eo@chad.scs.uiuc.edu

Table of Contents

Table S1: The number of helices (N_h) and turns (N_t) calculated using the COUDES[1] program versus the number determined experimentally for 10 terpene synthases. The numbers of helices and turns predicted for two bacterial diterpene cyclases (*B. japonicum* CPS, *M. tuberculosis* Rv3377c) and one plant diterpene cyclase (AgABS) are also shown.

		N_h (calc)	N_t (cacl)	N_h (expt)	N_t (expt)
3sqc	SHC	24	31	27	33
1w6k	OSC	26	35	28	34
1kzo	Farnesyl transferase	31	29	31	28
3dra	Geranylgeranyl transferase I	28	20	30	21
3dss	Geranylgeranyl transferase II	27	19	26	22
2a73	complement C3 β subunit	12	7	13	9
2f8z	FPPS	14	9	13	11
2q80	GGPPS	14	10	12	7
1ezf	SQS	13	8	14	11
2zcq	*S. aureus* CrtM	11	5	12	6
	B. japonicum CPS	23	19		
	M. tuberculosis Rv3377c	19	26		
	AgABS	34	32		

Table S2**:** CoMSIA training and test set predictions for AgABS Inhibition using GGPP as the substrate

	Experimental			Test Set pIC_{50} (M)[b, c]							
Compound[a]	IC_{50} (μM)	pIC_{50} (M)[b]	Training	1	2	3	4	5	6	Predicted	Residual
Aza-GGPP	0.01	7.9	7.8	**7.9**	8	7.9	7.9	7.9	7.9	7.9	0
BPH-916	0.15	6.8	6.7	6.8	6.8	**6.7**	6.7	6.8	6.8	6.7	0.1
BPH-882	0.19	6.7	6.7	6.7	6.7	6.8	6.7	**5.6**	6.7	5.6	1.1
BPH-890	0.24	6.6	6.6	6.7	6.7	6.7	6.7	**6**	6.6	6	0.6
BPH-922	0.25	6.6	6.6	6.6	**6.5**	6.7	6.6	6.6	6.5	6.5	0.1
BPH-901	0.29	6.5	6.5	6.5	**6.2**	6.5	6.6	6.7	6.5	6.2	0.3
BPH-899	0.35	6.5	6.5	6.5	6.6	6.5	**6.4**	6.5	6.4	6.4	0.1
BPH-921	0.48	6.3	6.2	6.1	6.1	6.1	6.1	**5.9**	6.2	5.9	0.4
BPH-892	0.53	6.3	6.3	**6.4**	6.2	6.3	6.3	6.3	6.3	6.4	-0.1
BPH-917	0.77	6.1	6.1	6.1	6.1	**6**	6.2	6.1	6.1	6	0.1
BPH-907	0.94	6	6	6	6	**5.9**	6	6	6	5.9	0.1
BPH-909	1.02	6	6.1	**6.3**	6	6.1	6	6	6.1	6.3	-0.3
BPH-908	1.08	6	6	6	5.9	6	6	5.9	**5.9**	5.9	0.1
BPH-924	1.18	5.9	6.2	6.1	6.1	6.2	6.1	**5.9**	6.2	5.9	0
BPH-896	1.45	5.8	5.7	5.8	5.6	5.7	5.8	5.5	**5.4**	5.4	0.4
BPH-863	1.72	5.7	5.7	5.8	5.6	5.7	5.7	5.8	**5.5**	5.5	0.2
BPH-923	1.94	5.7	5.8	**6.2**	5.7	5.7	5.7	5.7	5.7	6.2	-0.5
BPH-905	2.47	5.6	5.5	5.6	5.6	5.5	**5.3**	5.5	5.5	5.3	0.3
BPH-902	3	5.5	5.6	5.5	**5**	5.4	5.5	5.6	5.6	5	0.5
BPH-661	10.68	5	5	5	5	5	**5.2**	5.1	5	5.2	-0.2
BPH-904	12.58	4.9	4.9	4.9	4.9	**4.9**	4.9	4.9	4.9	4.9	0
BPH-913	15	4.8	4.8	4.9	**4.6**	4.8	4.8	4.7	4.8	4.6	0.2
BPH-918	30.34	4.5	4.5	4.5	4.4	4.5	**5.3**	4.5	4.5	5.3	-0.8
BPH-876	41.63	4.4	4.4	4.4	4.6	4.4	4.4	4.5	**5**	5	-0.6
BPH-891	41.72	4.4	4.3	4.4	4.3	**5.6**	4.3	4.4	4.3	5.6	-1.2
BPH-883	67.7	4.2	4.2	4.1	4.2	4.2	**5.6**	4.2	4.2	5.6	-1.4
BPH-673	175.68	3.8	3.7	3.7	**5.2**	3.8	3.7	3.8	3.7	5.2	-1.4
BPH-914	274	3.6	3.6	**4.4**	3.6	3.6	3.5	3.6	3.6	4.4	-0.8
		q^2	0.67	0.56	0.59	0.59	0.64	0.59	0.60		

r^2	0.99	0.99	0.99	0.99	0.99	0.99	0.99
N	6	6	5	6	5	6	6
F	734	474	503	503	602	603	916
n	28	23	23	23	23	24	24
%Steric	0.17	0.18	0.16	0.17	0.17	0.15	0.14
%Electrostatic	0.18	0.19	0.20	0.22	0.20	0.25	0.18
%Donor	0.33	0.31	0.33	0.31	0.37	0.37	0.34
%Acceptor	0.31	0.32	0.29	0.32	0.25	0.29	0.33

N = number of components; n = Number of training set compounds

[a] the structures of all inhibitors shown in Figure S4.
[b] $pIC_{50}=-\log_{10}IC_{50}[M]$
[c] bold values indicated compounds not included in training set.

Table S3**.** Distance matrix for 20 terpene synthase structures obtained by using the Protein Structure Alignment by SSM program[2]. The values shown are the normalized C^{α} rmsd values (1000* (C^{α} rmsd (Å)/N), where N is the number of aligned residues in structures *i* and *j*). The letters shown are the PDB file designations.[a]

	1JFG	1DI1	5EAS	1N1B	2ONG	1PS1	2F7M	2HER	2IHI	1YHK	2FOR	1UBV	2EWG	1RQI	1RTR	2E8T	2J1O	2Q80	2ZCP	1EZF
1JFG[b]	-	17.4	13.1	14.7	14.9	15.6	17.4	19.0	19.9	17.7	21.1	17.4	18.2	19.9	20.2	20.0	21.4	21.1	18	19.9
1DI1	17.4	-	24.9	15.3	13.7	11.3	21.2	20.2	20.9	19.8	22.4	21.1	19.7	22.1	25.0	25.7	27.8	25.5	20.5	22.1
5EAS	13.1	24.9	-	4.3	4.0	12.0	15.6	16.2	16.6	15.1	19.7	15.0	15.5	17.8	19.9	14.5	19.9	14.5	17.2	16.9
1N1B	14.7	15.3	4.3	-	2.5	11.9	15.5	16.7	16.6	16.0	20.9	16.5	16.2	20.8	20.3	15.5	20.0	16.2	18.6	18.2
2ONG	14.9	13.7	4.0	2.5	-	11.2	14.8	16.4	14.9	14.9	19.9	15.6	18.7	17.8	20.2	15.1	20.7	15.6	17	17.3
1PS1	15.6	11.3	12.0	11.9	11.2	-	16.9	17.5	18.3	16.5	22.5	17.1	16.2	21.3	21.8	17.5	23.0	16.7	17.6	18.1
2F7M	17.4	21.2	15.6	15.5	14.8	16.9	-	7.0	5.1	5.9	10.8	2.7	5.8	10.8	10.5	9.5	12.1	10.3	19.1	18
2HER	19.0	20.2	16.2	16.7	16.4	17.5	7.0	-	5.1	5.7	11.8	7.5	5.4	11.1	11.9	10.3	13.9	10.0	20.1	17.2
2IHI	19.9	20.9	16.6	16.6	14.9	18.3	5.1	5.1	-	4.0	10.4	5.7	4.7	9.1	10.7	10.0	12.3	9.6	19.7	18.2
1YHK	17.7	19.8	15.1	16.0	14.9	16.5	5.9	5.7	4.0	-	10.4	5.8	2.7	9.6	10.9	10.2	12.4	10.1	19.6	17.4
2FOR	21.1	22.4	19.7	20.9	19.9	22.5	10.8	11.8	10.4	10.4	-	11.1	10.3	3.8	4.8	18.5	5.8	10.8	19.6	18.7
1UBV	17.4	21.1	15.0	16.5	15.6	17.1	2.7	7.5	5.7	5.8	11.1	-	6.0	11.0	10.5	9.1	12.3	10.5	19.9	18.8
2EWG	18.2	19.7	15.5	16.2	18.7	16.2	5.8	5.4	4.7	2.7	10.3	6.0	-	8.3	11.3	9.3	12.7	8.9	19	16.3
1RQI	19.9	22.1	17.8	20.8	17.8	21.3	10.8	11.1	9.1	9.6	3.8	11.0	8.3	-	6.2	9.7	7.4	8.8	18.6	17.5
1RTR	20.2	25.0	19.9	20.3	20.2	21.8	10.5	11.9	10.7	10.9	4.8	10.5	11.3	6.2	-	11.8	6.3	11.4	18.1	19
2E8T	20.0	25.7	14.5	15.5	15.1	17.5	9.5	10.3	10.0	10.2	18.5	9.1	9.3	9.7	11.8	-	19.2	4.8	20.3	18.6
2J1O	21.4	27.8	19.9	20.0	20.7	23.0	12.1	13.9	12.3	12.4	5.8	12.3	12.7	7.4	6.3	19.2	-	16.4	21.3	23
2Q80	21.1	25.5	14.5	16.2	15.6	16.7	10.3	10.0	9.6	10.1	10.8	10.5	8.9	8.8	11.4	4.8	16.4	-	17.8	15.7
2ZCP	18	20.5	17.2	18.6	17	17.6	19.1	20.1	19.7	19.6	19.6	19.9	19	18.6	18.1	20.3	21.3	17.8	-	10.8
1EZF	19.9	22.1	16.9	18.2	17.3	18.1	18	17.2	18.2	17.4	18.7	18.8	16.3	17.5	19	18.6	23	15.7	10.8	-

[a]These normalized C^{α} rmsd values were used to construct the phylogenic tree shown in Fig 6a.

[b]1JFG: *Fusarium sporotrichioides* trichodiene synthase
1DI1: *Penicillium roqueforti* aristolochene synthase
5EAS: *Nicotiana tabacum* 5-epi-aristolochene synthase
1N1B: *Salvia officinalis* (+)-bornyl diphosphate Synthase
2ONG: *Mentha spicata* 4S-limonene synthase
1PS1: *Streptomyces sp*. pentalenene synthase
2F7M: *Homo sapiens* farnesyl diphosphate synthase (FPPS)
2HER: *Cryptosporidium parvum* FPPS
2IHI: *Plasmodium vivax* FPPS
1YHK: *Trypanosoma cruzi* FPPS
2FOR: *Shigella flexneri* FPPS
1UBV: *Gallus gallus* FPPS
2EWG: *Trypanosoma brucei* FPPS
1RQI:*Escherichia coli* FPPS
1RTR:*Staphylococcus aureus* FPPS
2E8T: *Saccharomyces cerevisiae* geranylgeranyl diphosphate synthase (GGPPS)
2J1O: *Sinapis alba* GGPPS
2Q80: *Homo sapiens* GGPPS
2ZCP: *S. aureus* dehydrosqualene synthase (crtM)
1EZF: *Homo sapiens* squalene synthase

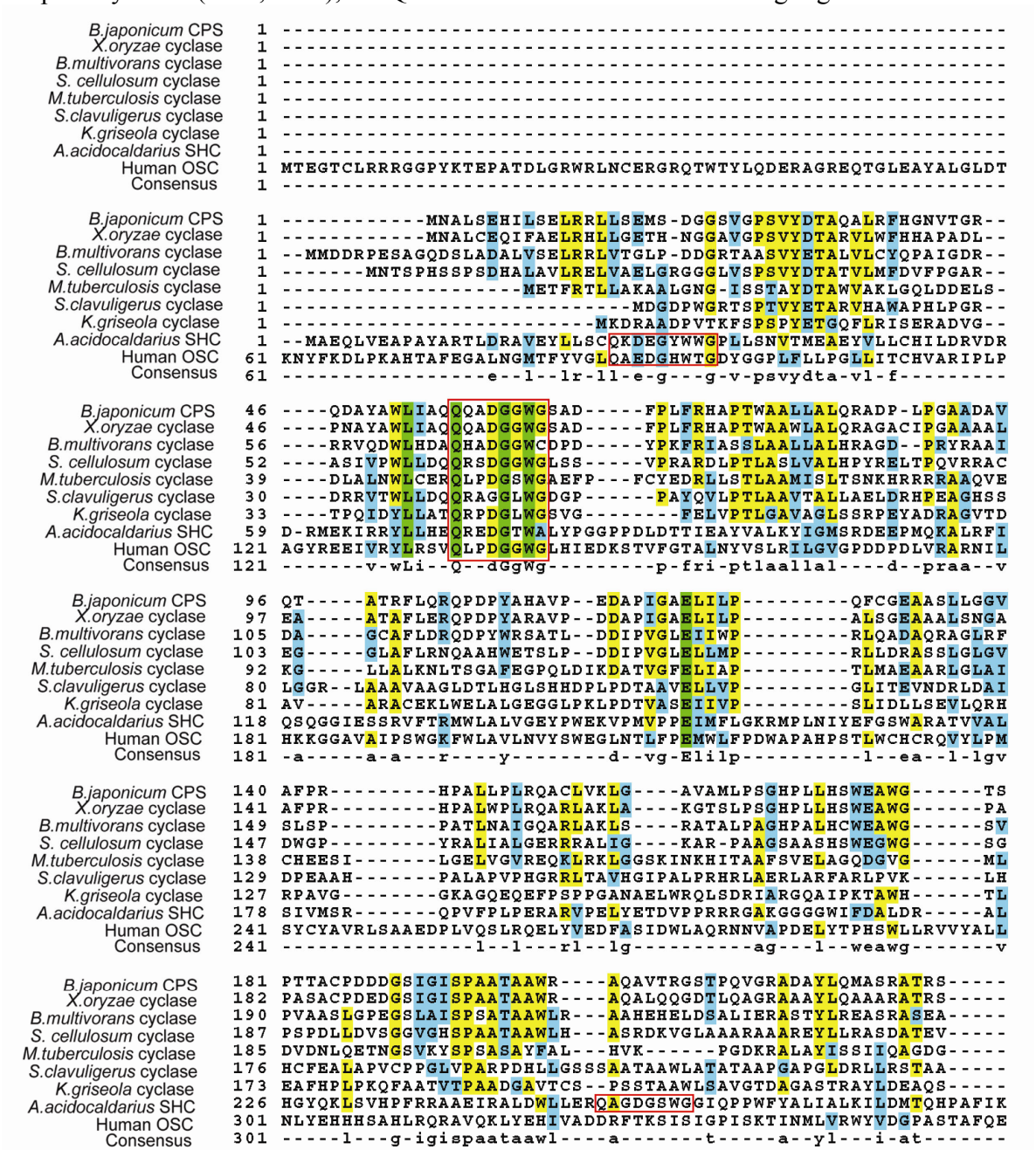
Table S4: Distance matrix for 20 terpene synthase structures obtained by using the Protein Structure Alignment by Incremental Combinatorial Extension (CE) program[3]. The values shown are the normalized C^{α} rmsd values (1000* (C^{α} rmsd (Å)/N), where N is the number of aligned residues in structures *i* and *j*). The letters shown are the PDB file designations.[a]

	1JFG	1DI1	5EAS	1N1B	2ONG	1PS1	2F7M	2HER	2IHI	1YHK	2FOR	1UBV	2EWG	1RQI	1RTR	2E8T	2J1O	2Q80	1EZF	2ZCP
1JFG[b]	-	20.0	13.8	14.8	15.4	13.1	16.9	16.0	33.1	15.9	17.4	16.8	16.5	20.8	16.6	18.7	18.4	23.4	20.4	19.7
1DI1	20.0	-	13.8	15.7	14.3	10.3	18.3	16.6	32.8	20.6	35.0	22.6	22.7	18.8	23.2	20.7	23.5	22.5	20.4	17.9
5EAS	13.8	13.8	-	4.2	4.0	13.4	14.5	18.6	14.2	13.8	16.8	13.9	14.8	16.5	15.7	16.7	18.1	17.3	16.4	15.1
1N1B	14.8	15.7	4.2	-	2.6	15.4	14.5	20.0	17.0	16.7	20.0	16.6	16.7	19.3	20.4	16.0	19.4	17.0	17.3	17.4
2ONG	15.4	14.3	4.0	2.6	-	14.2	16.3	19.5	18.0	16.0	19.5	15.9	16.0	18.6	17.8	16.8	23.1	18.0	17.1	16.9
1PS1	13.1	10.3	13.4	15.4	14.2	-	18.8	17.3	19.0	16.8	22.4	18.2	16.9	20.8	20.1	21.5	22.7	18.2	17.4	17.1
2F7M	16.9	18.3	14.5	14.5	16.3	18.8	-	6.3	4.5	5.0	9.9	5.1	5.3	9.0	9.7	9.3	11.5	9.0	16.9	17.8
2HER	16.0	16.6	18.6	20.0	19.5	17.3	6.3	-	5.9	6.0	9.9	8.9	5.9	9.2	9.8	9.7	11.1	8.8	17.6	19.2
2IHI	33.1	32.8	14.2	17.0	18.0	19.0	4.5	5.9	-	4.9	8.9	5.8	4.7	8.5	8.7	8.8	10.5	9.3	17	17.6
1YHK	15.9	20.6	13.8	16.7	16.0	16.8	5.0	6.0	4.9	-	9.2	6.4	2.5	7.6	9.8	9.1	10.9	8.9	17.1	18.3
2FOR	17.4	35.0	16.8	20.0	19.5	22.4	9.9	9.9	8.9	9.2	-	9.8	9.2	3.4	5.6	10.4	9.3	9.7	21.1	19.6
1UBV	16.8	22.6	13.9	16.6	15.9	18.2	5.1	8.9	5.8	6.4	9.8	-	6.4	9.8	8.7	9.5	10.6	9.8	17.1	17.8
2EWG	16.5	22.7	14.8	16.7	16.0	16.9	5.3	5.9	4.7	2.5	9.2	6.4	-	7.9	9.2	9.1	10.9	8.5	19.3	19.2
1RQI	20.8	18.8	16.5	19.3	18.6	20.8	9.0	9.2	8.5	7.6	3.4	9.8	7.9	-	6.7	10.7	12.9	8.5	21.2	20.6
1RTR	16.6	23.2	15.7	20.4	17.8	20.1	9.7	9.8	8.7	9.8	5.6	8.7	9.2	6.7	-	10.9	9.6	10.4	20.3	20
2E8T	18.7	20.7	16.7	16.0	16.8	21.5	9.3	9.7	8.8	9.1	10.4	9.5	9.1	10.7	10.9	-	12.7	4.5	19.6	21.8
2J1O	18.4	23.5	18.1	19.4	23.1	22.7	11.5	11.1	10.5	10.9	9.3	10.6	10.9	12.9	9.6	12.7	-	16.5	23.8	21
2Q80	23.4	22.5	17.3	17.0	18.0	18.2	9.0	8.8	9.3	8.9	9.7	9.8	8.5	8.5	10.4	4.5	16.5	-	16.5	17.8
1EZF	20.4	20.4	16.4	17.3	17.1	17.4	16.9	17.6	17	17.1	21.1	17.1	19.3	21.2	20.3	19.6	23.8	16.5	-	11.4
2ZCP	19.7	17.9	15.1	17.4	16.9	17.1	17.8	19.2	17.6	18.3	19.6	17.8	19.2	20.6	20	21.8	21	17.8	11.4	-

[a]These normalized C^{α} rmsd values were used to construct the phylogenic tree shown in Fig S5.

[b]1JFG: *Fusarium sporotrichioides* trichodiene synthase
1DI1: *Penicillium roqueforti* aristolochene synthase
5EAS: *Nicotiana tabacum* 5-epi-aristolochene synthase
1N1B: *Salvia officinalis* (+)-bornyl diphosphate Synthase
2ONG: *Mentha spicata* 4S-limonene synthase
1PS1: *Streptomyces sp*. pentalenene synthase
2F7M: *Homo sapiens* farnesyl diphosphate synthase (FPPS)
2HER: *Cryptosporidium parvum* FPPS
2IHI: *Plasmodium vivax* FPPS
1YHK: *Trypanosoma cruzi* FPPS
2FOR: *Shigella flexneri* FPPS
1UBV: *Gallus gallus* FPPS
2EWG: *Trypanosoma brucei* FPPS
1RQI:*Escherichia coli* FPPS
1RTR:*Staphylococcus aureus* FPPS
2E8T: *Saccharomyces cerevisiae* geranylgeranyl diphosphate synthase (GGPPS)
2J1O: *Sinapis alba* GGPPS
2Q80: *Homo sapiens* GGPPS
2ZCP: *S. aureus* dehydrosqualene synthase (crtM)
1EZF: *Homo sapiens* squalene synthase

Figure S1: ClustalW output showing comparison between bacterial diterpene cyclases and two triterpene cyclases (SHC, OSC), the QW motifs and DXDD motif are highlighted in red box.



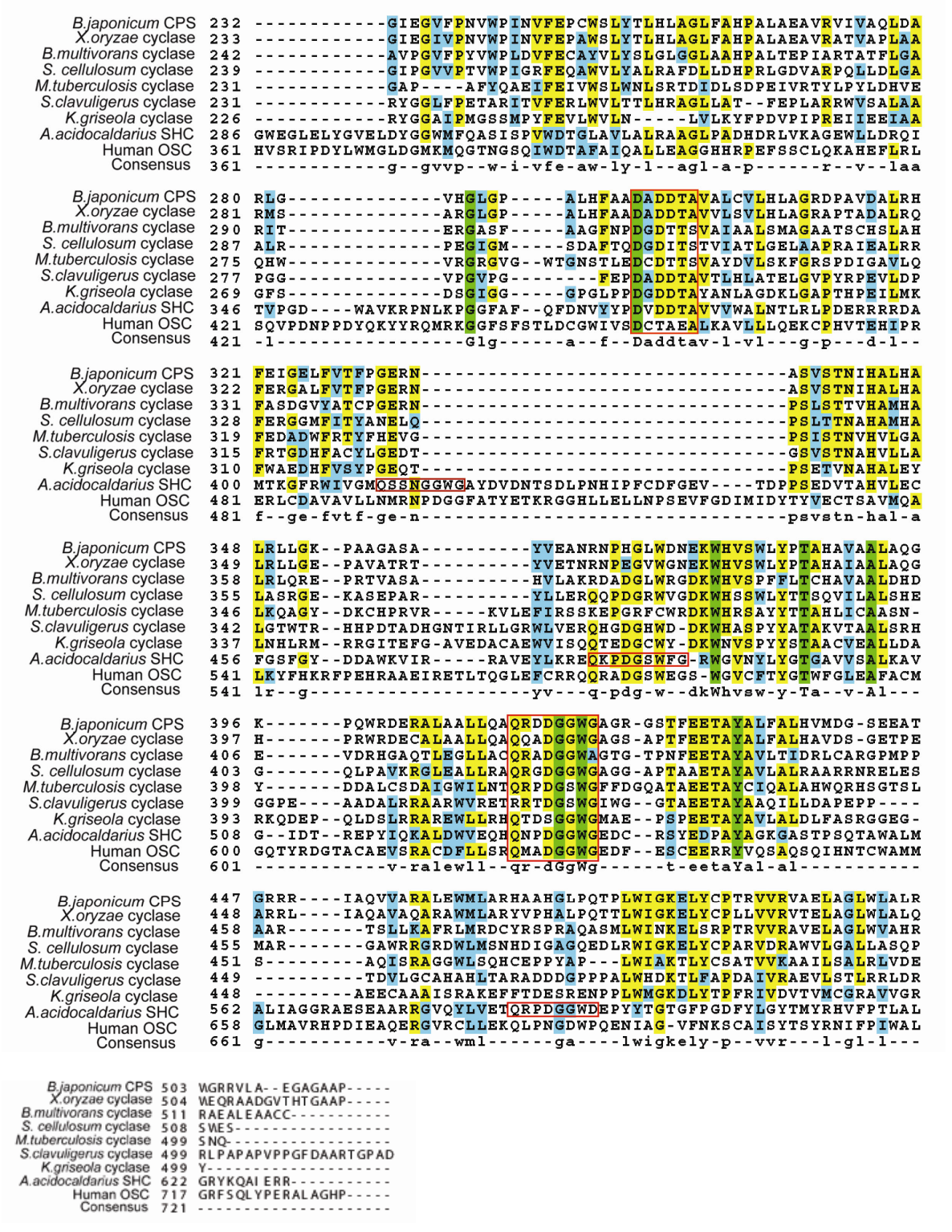

B.japonicum CPS
X.oryzae cyclase
B.multivorans cyclase
S. cellulosum cyclase
M.tuberculosis cyclase
S.clavuligerus cyclase
K.griseola cyclase
A.acidocaldarius SHC
Human OSC
Consensus

Figure S2: Graph showing correlation between the number of predicted helices (red) (COUDES program)[1] and turns (blue) versus number seen experimentally in SHC, OSC, FTase I, GGTase I, GGTase II, FPPS, GGPPS, SQS and CrtM (≥ 5 helical residues define a helix; ≥ 4 turn residues define a turn).

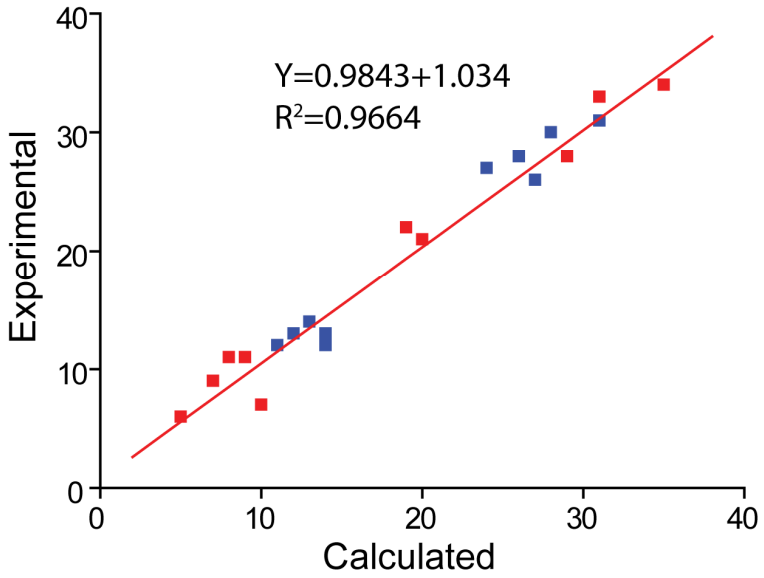


Figure S3: Full alignment of 228 terpene synthases obtained by using the JPRED3[4] program. The target sequence (at the top of the output) is abietadiene synthase from *Abies grandis*, minus the plastid targeting segment[5].

Figure S3 (continued).

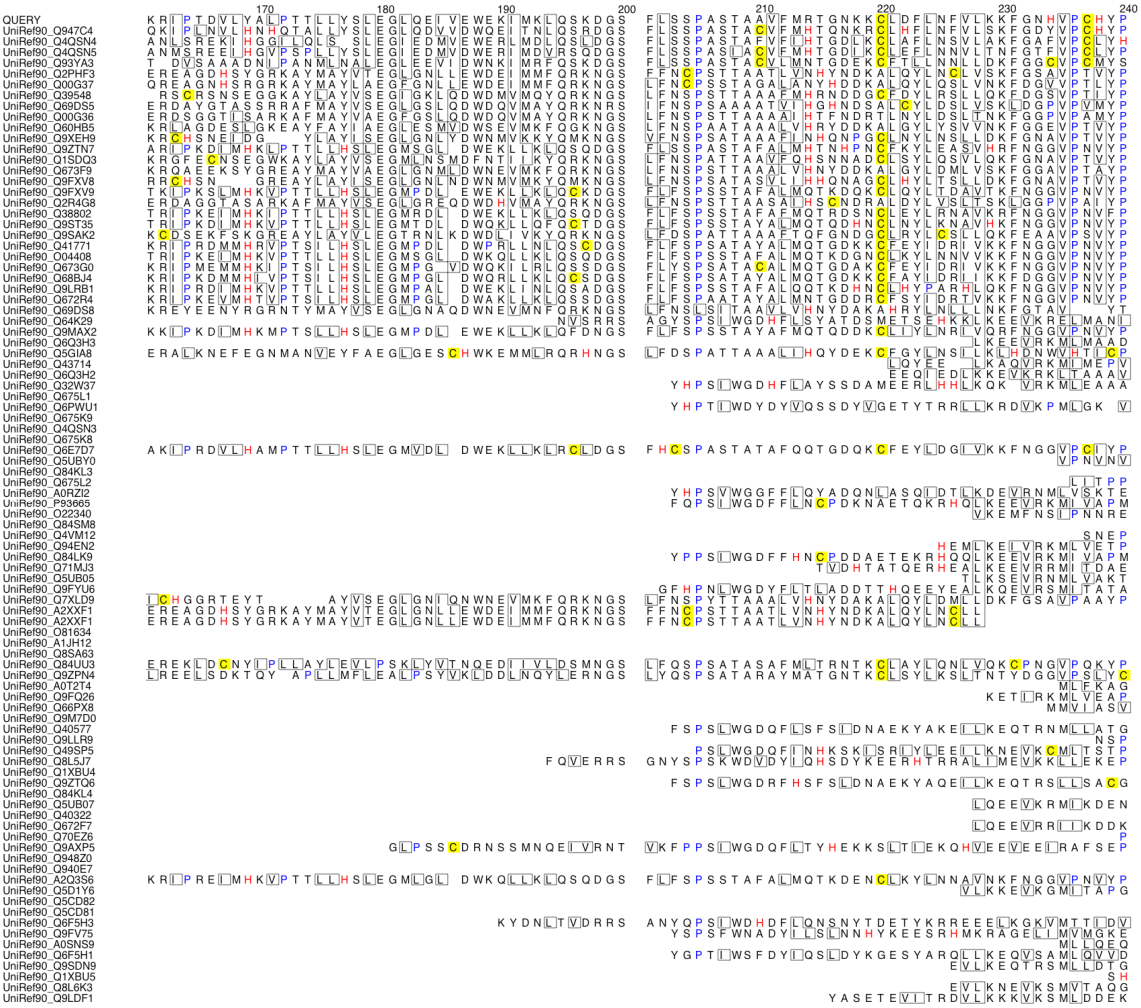


Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

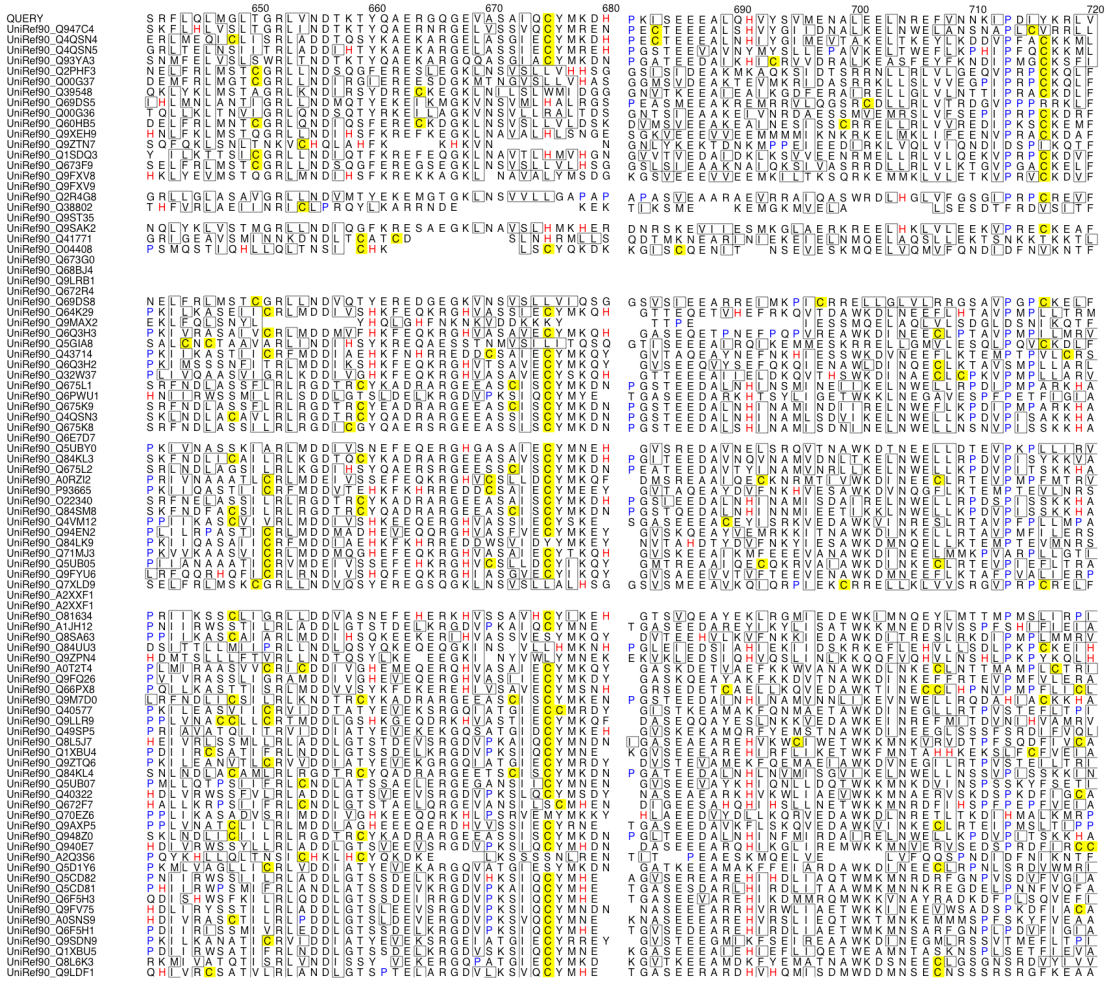


Figure S3 (continued).

Figure S3 (continued).

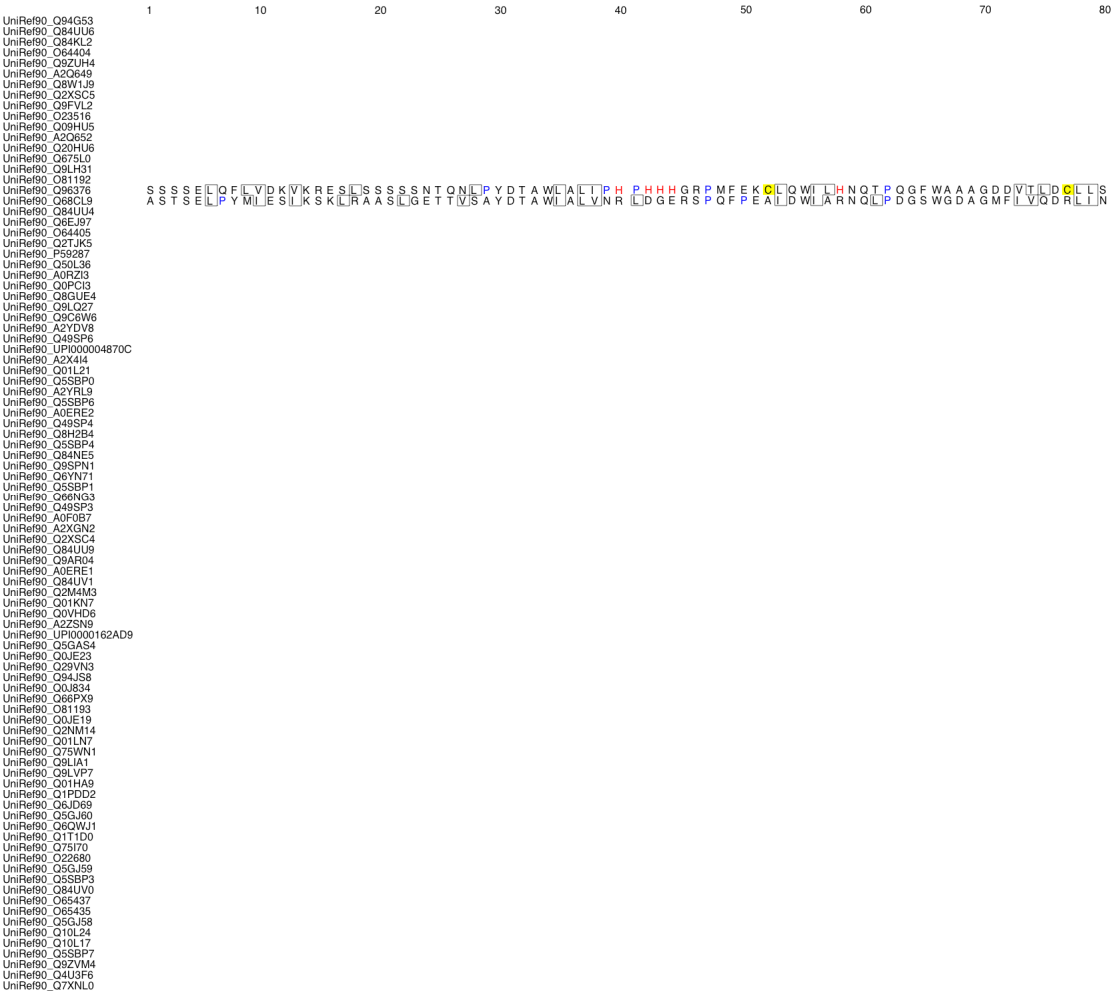


Figure S3 (continued).

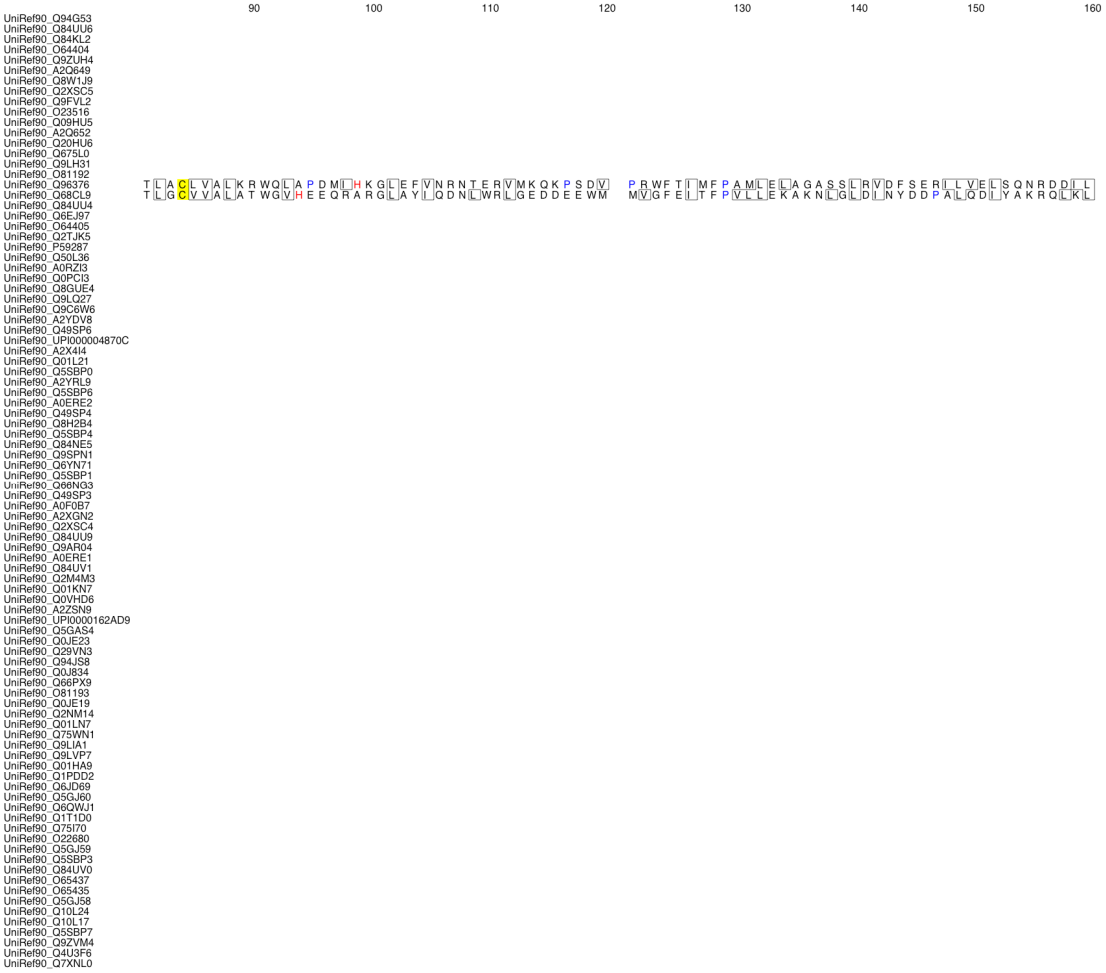


Figure S3 (continued).

Figure S3 (continued).

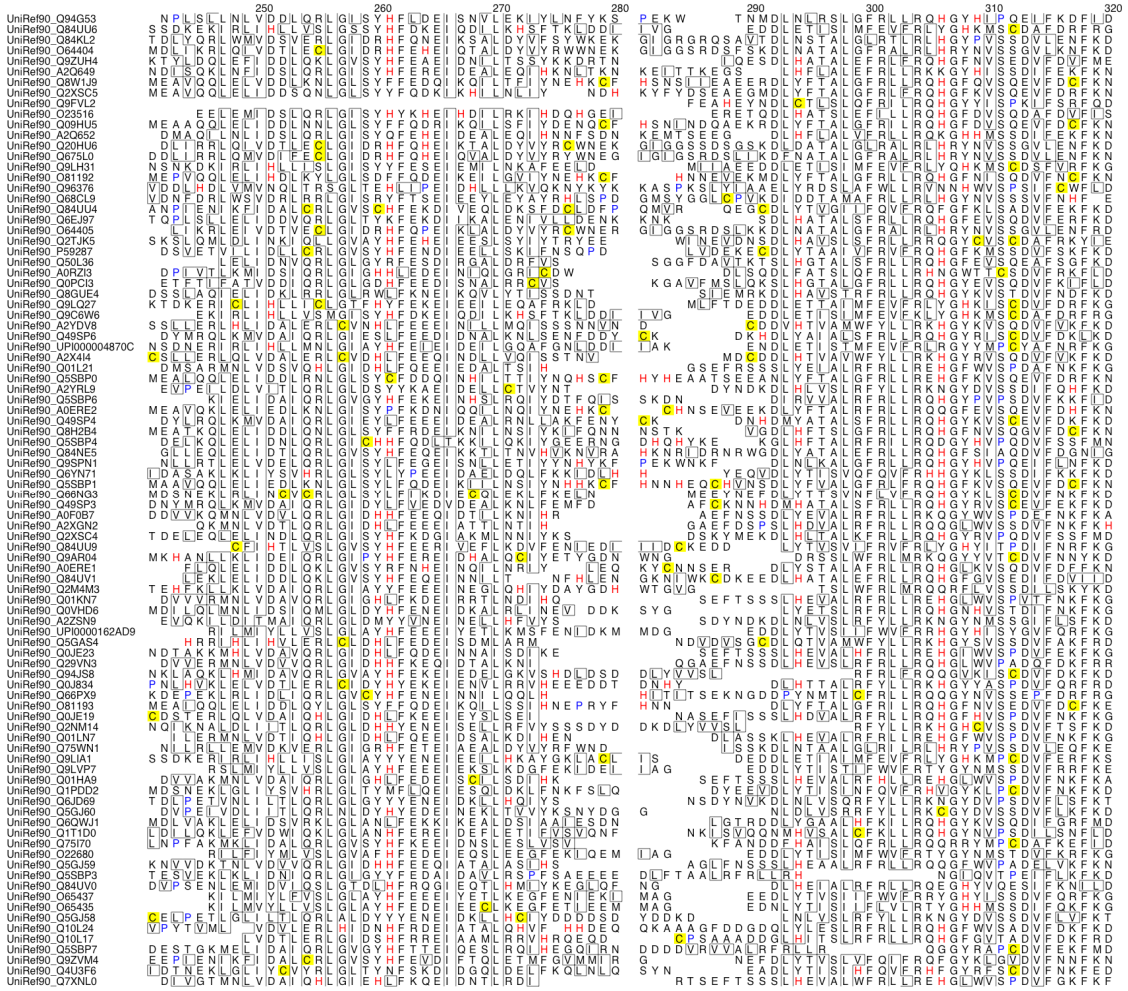


Figure S3 (continued).

Figure S3 (continued).

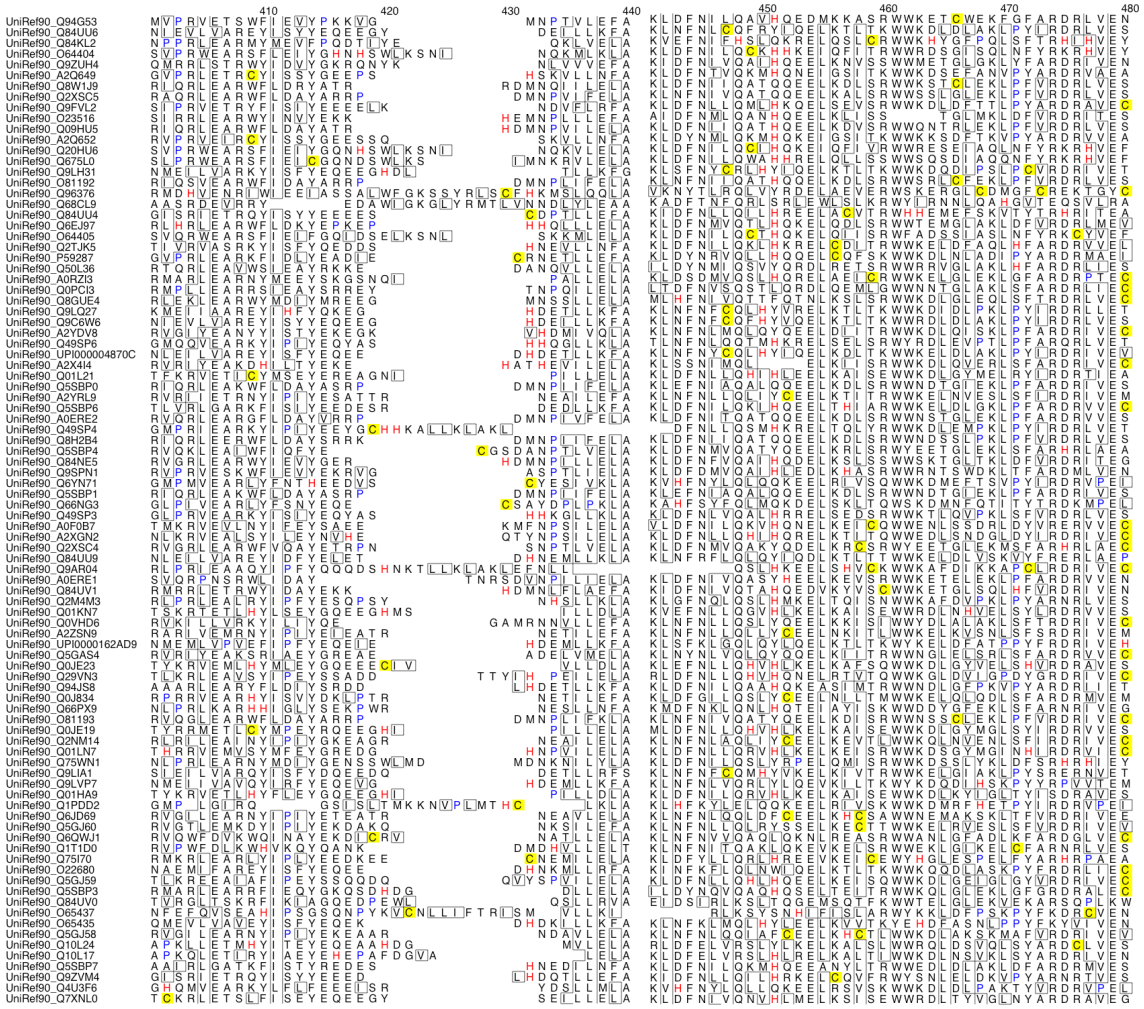


Figure S3 (continued).

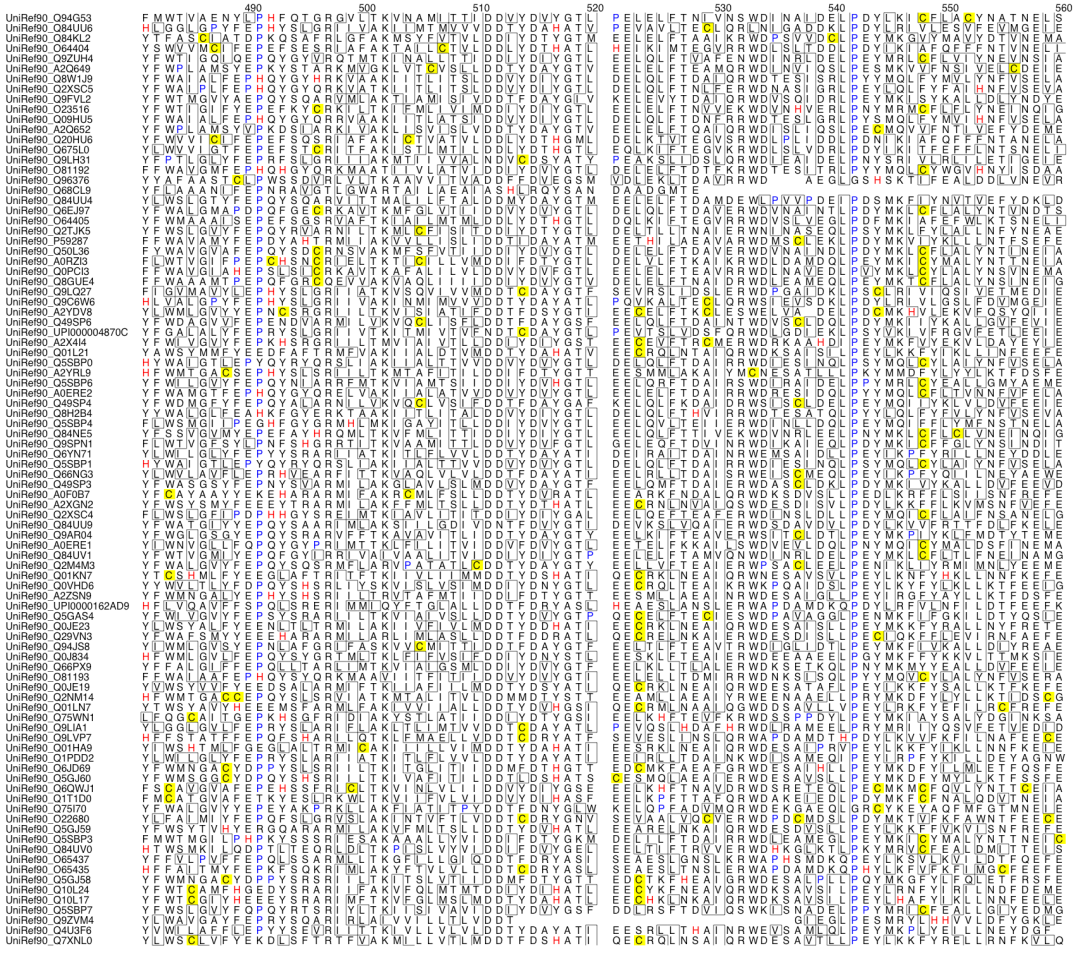


Figure S3 (continued).

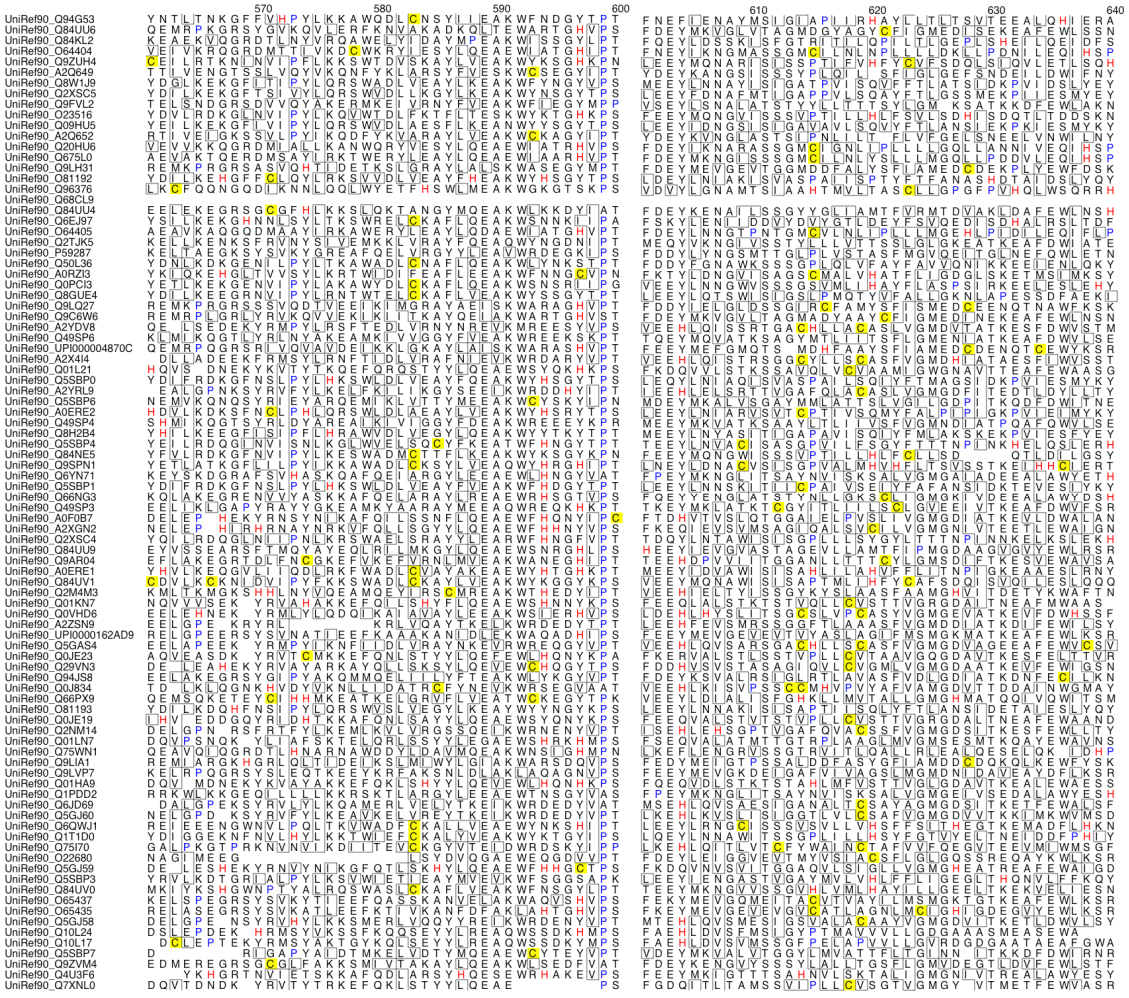


Figure S3 (continued).

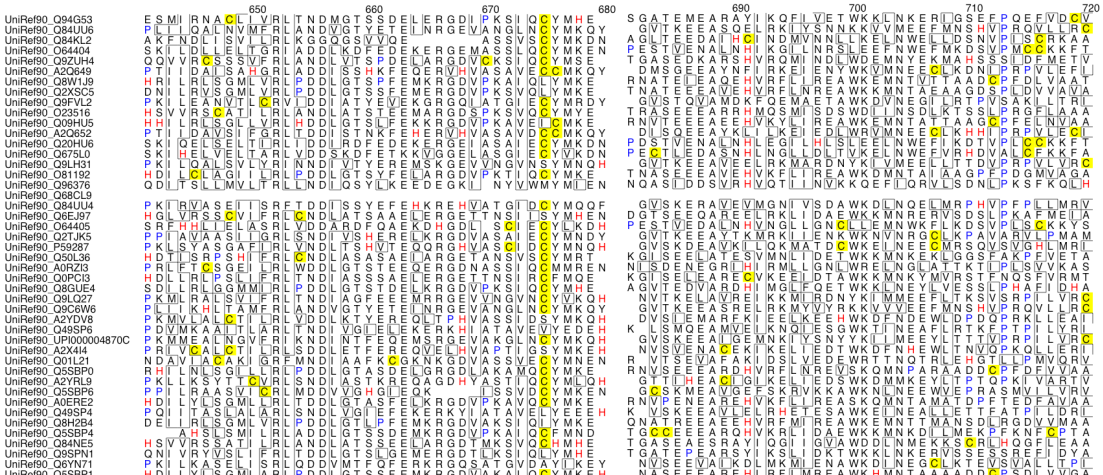


Figure S3 (continued).

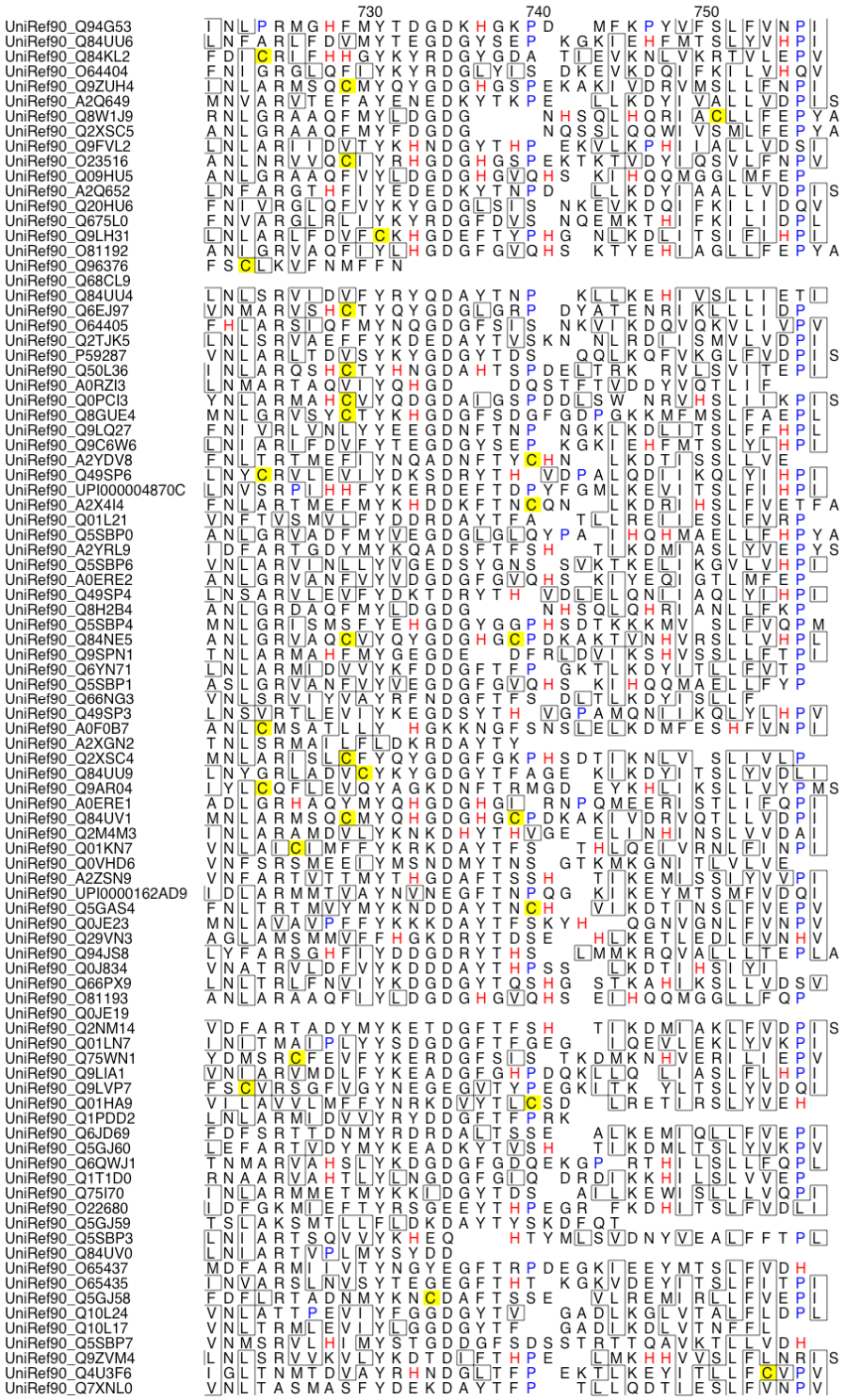


Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

Figure S3 (continued).

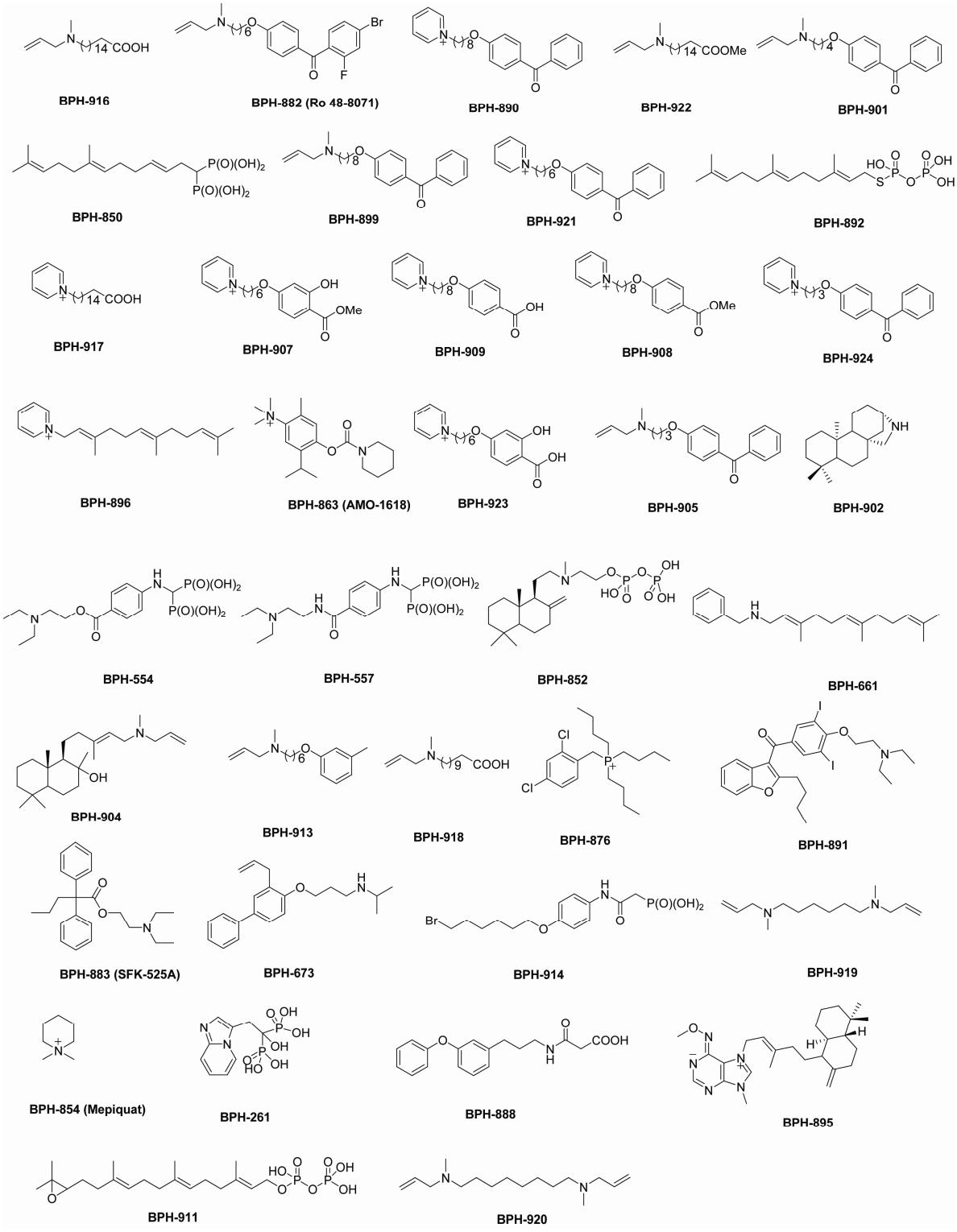


Figure S4: Structures of *Abies grandis* abietadiene synthase inhibitors investigated using GGPP as substrate. Structures are shown in order of decreasing potency (top left is the most potent).

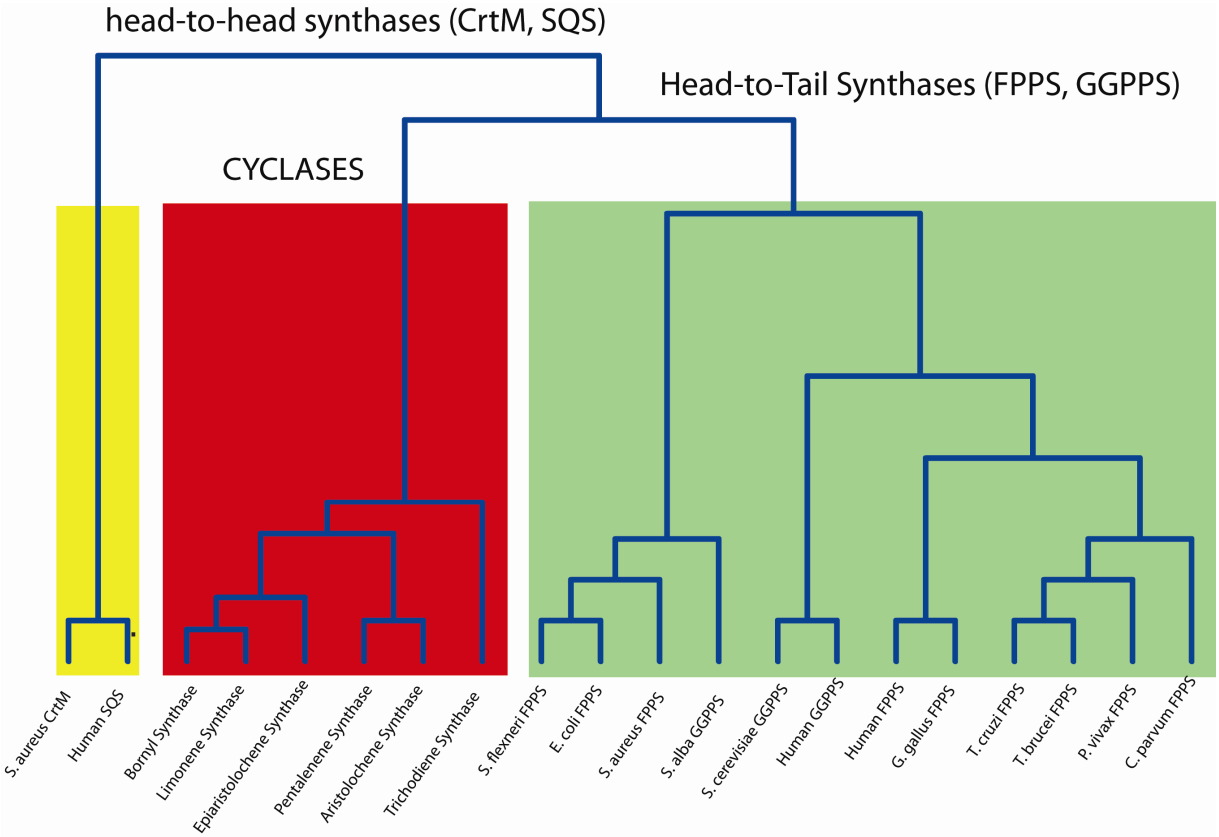
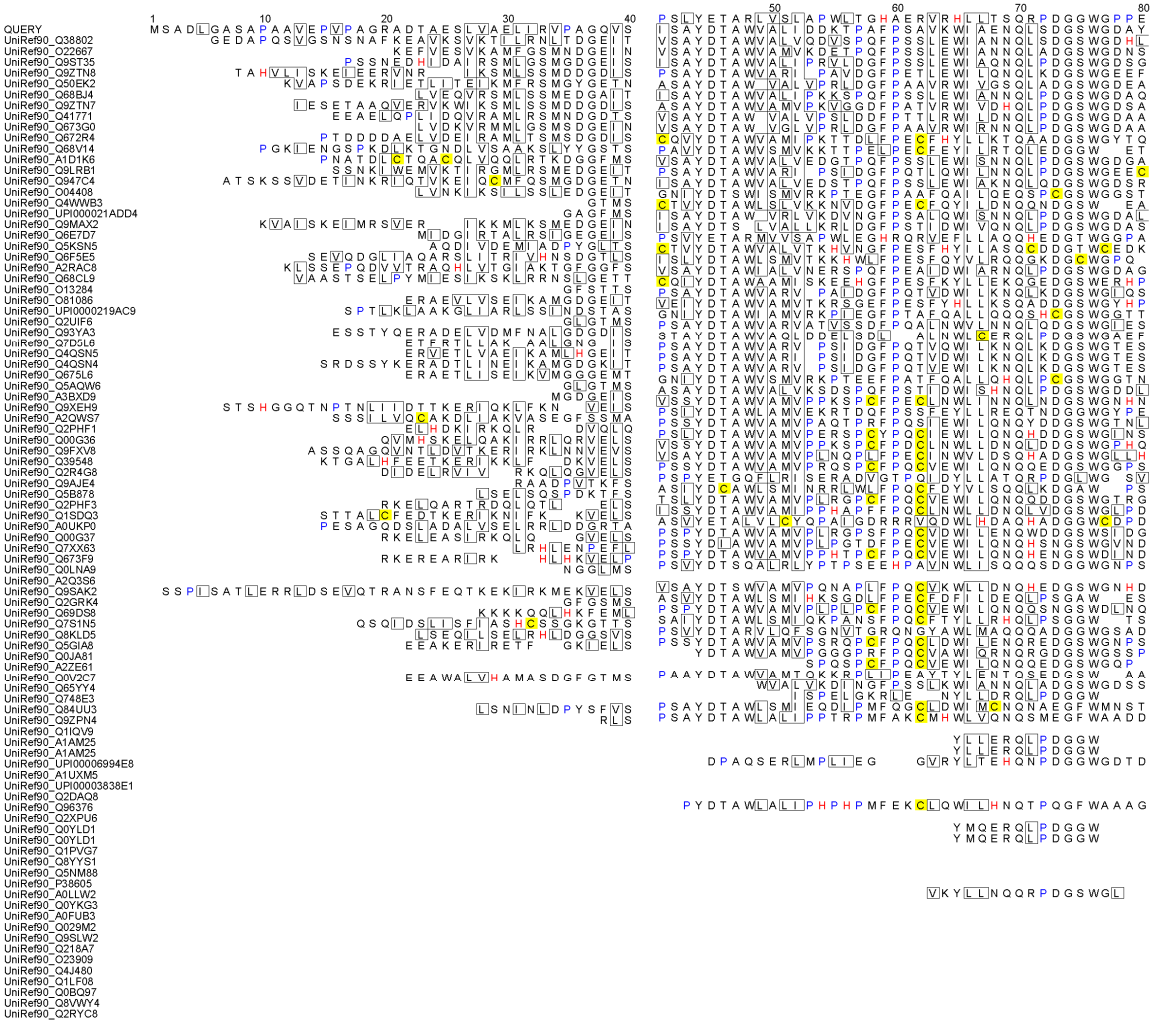


Figure S5: Structure dendrogram for α and αβ proteins showing strong clustering of GGPPS/FPPS, separate from that seen with the monoterpene and sesquiterpene cyclases. This dendrogram was generated from the structural alignments obtained by using the Incremental Combinatorial Extension (CE) Program[3].

Figure S6: JPRED3[4] output using *Bradyrhizobium japonicum* cylcase blr2149 as the query.



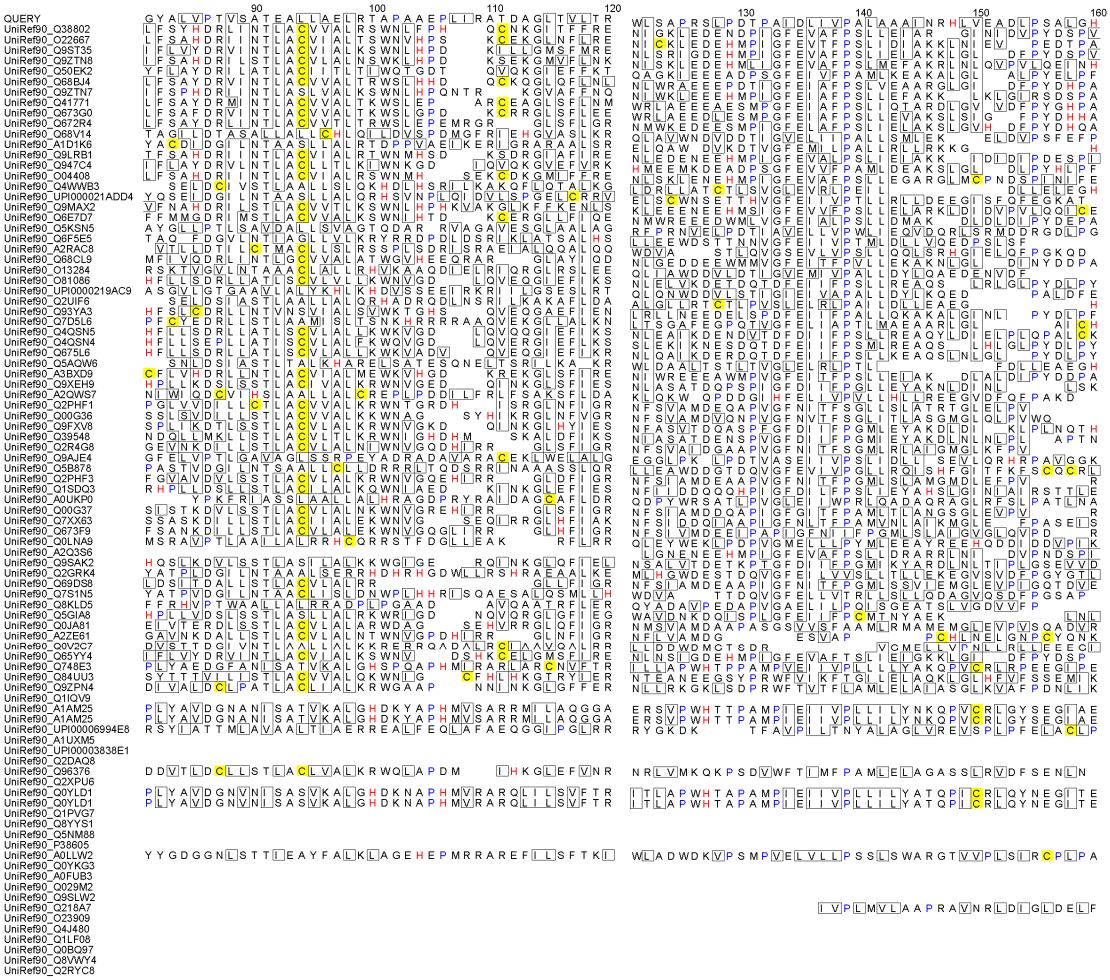


Figure S6 (continued).

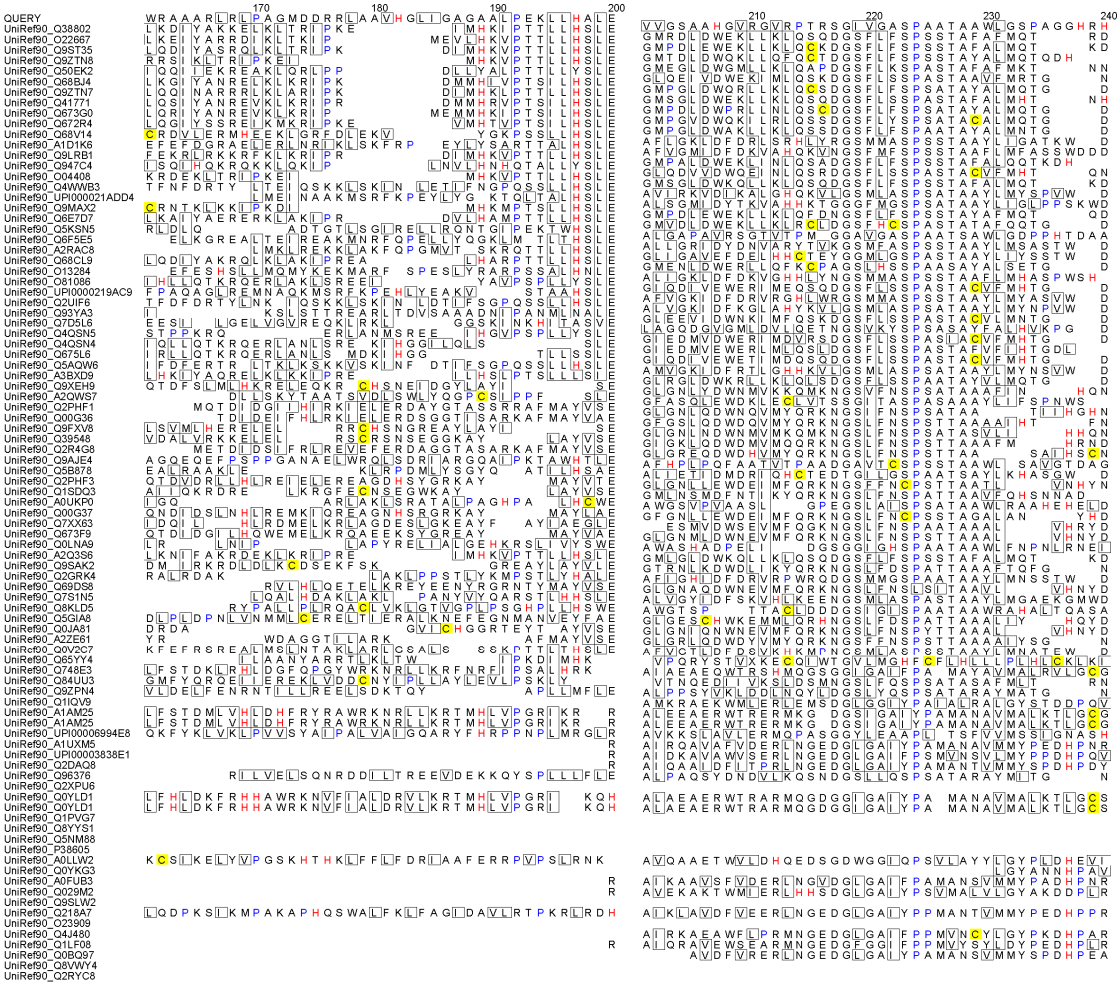


Figure S6 (continued).

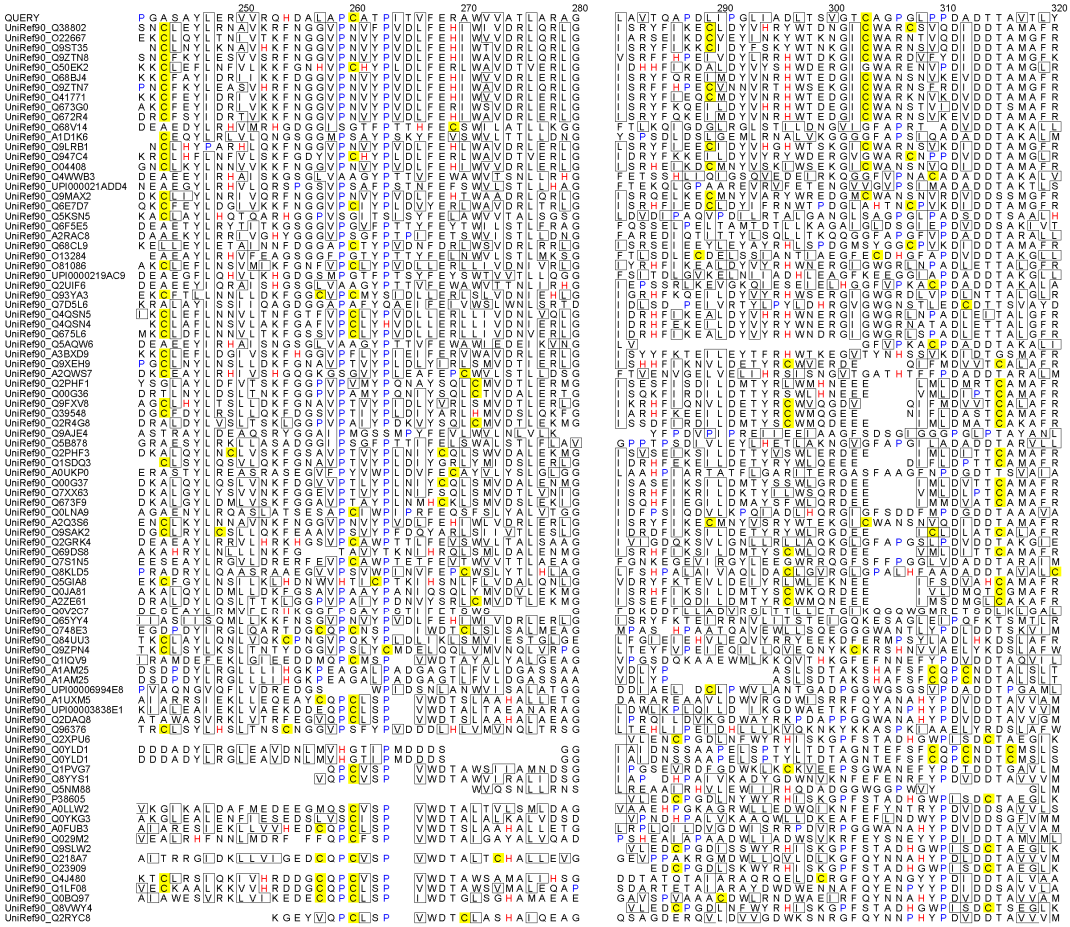


Figure S6 (continued).

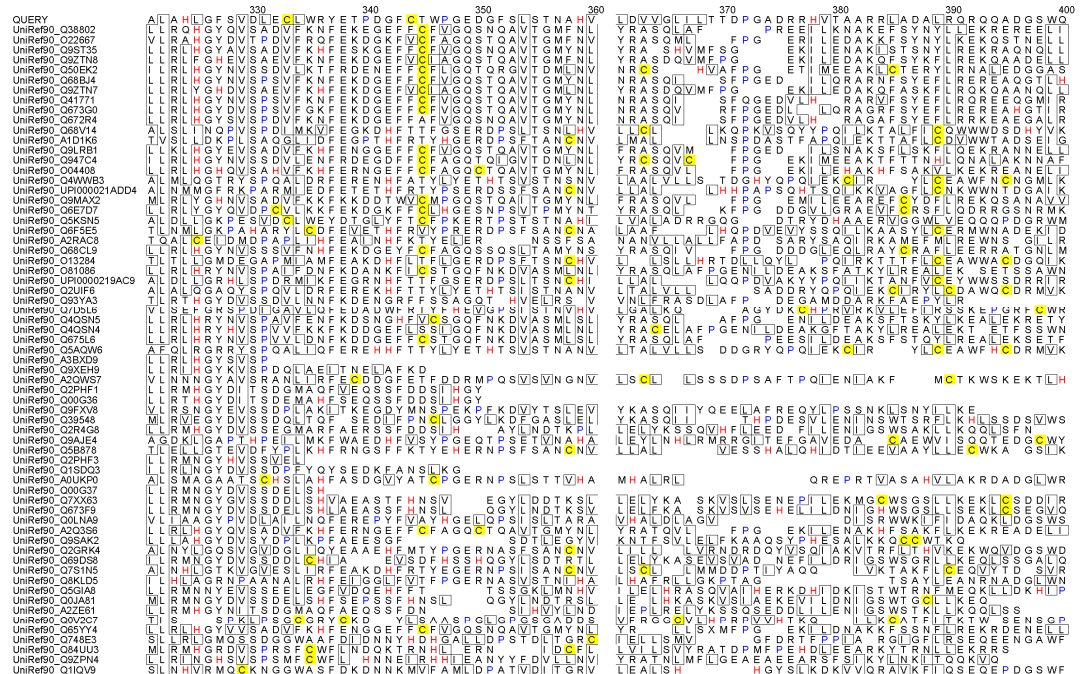


Figure S6 (continued).

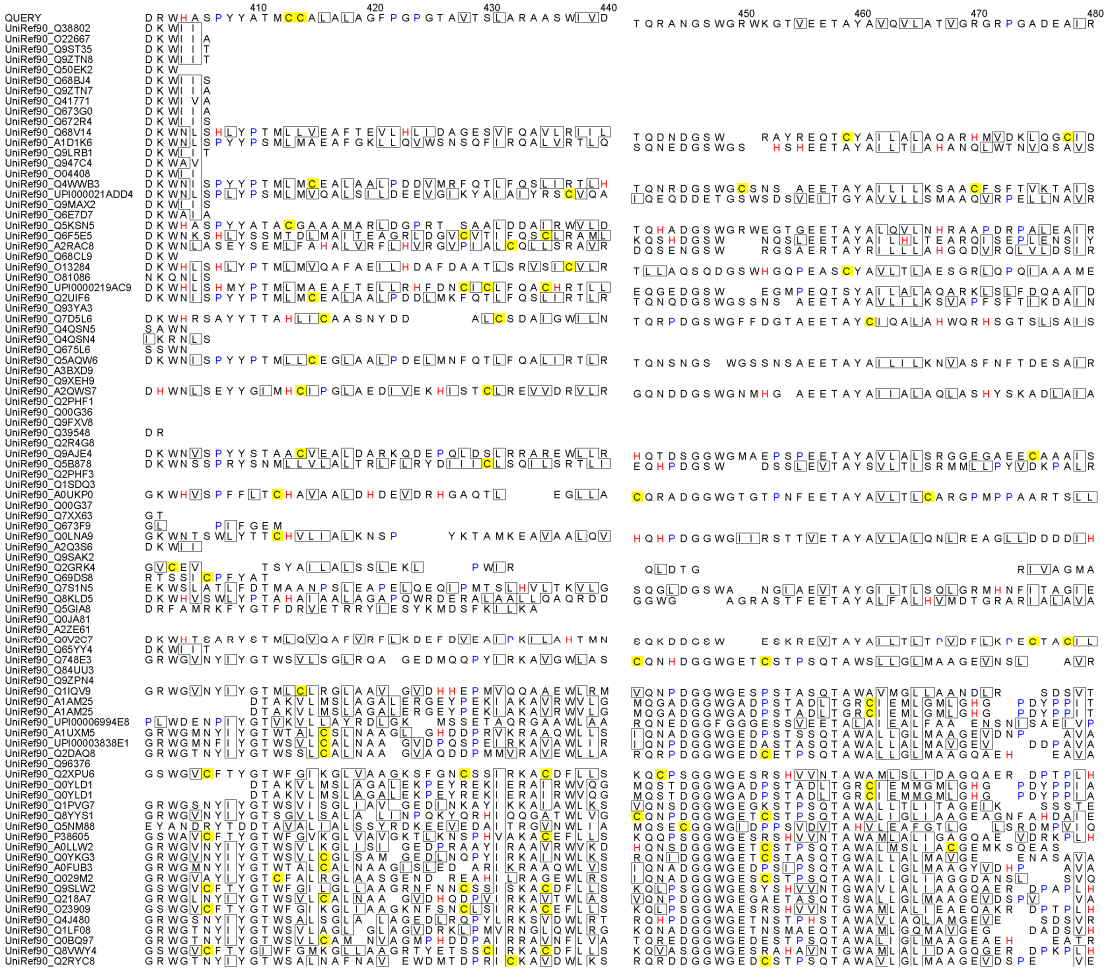


Figure S6 (continued).

Figure S6 (continued).

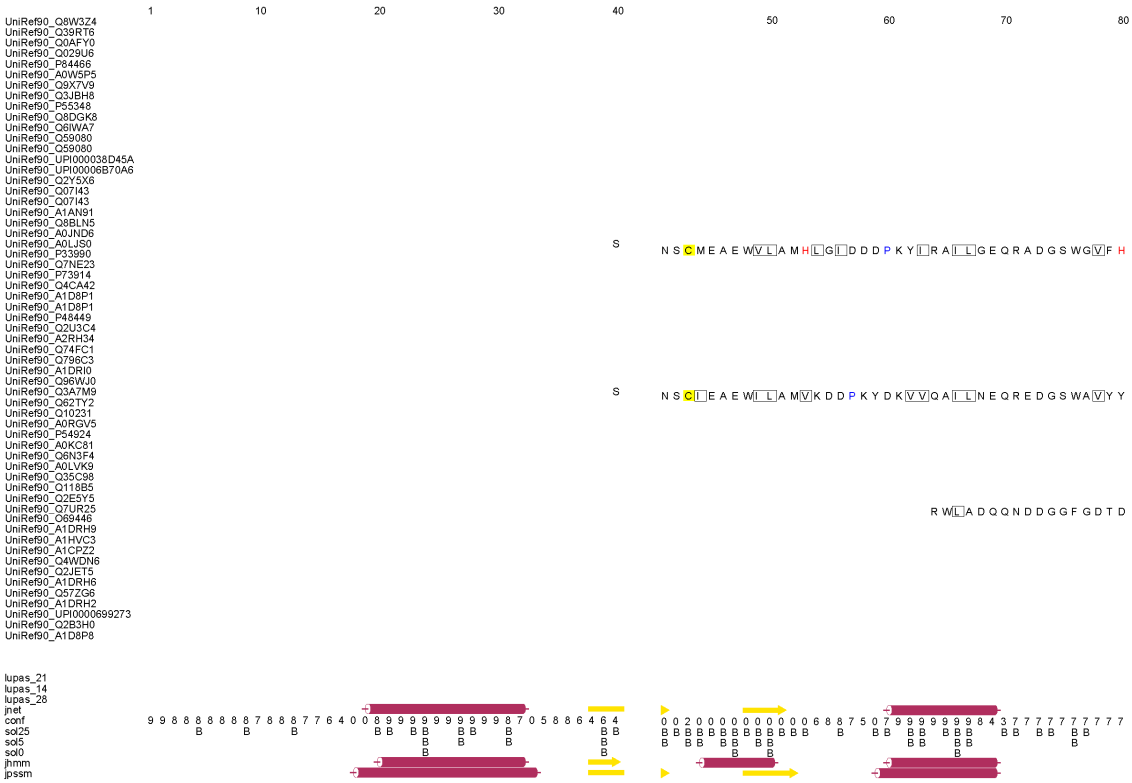
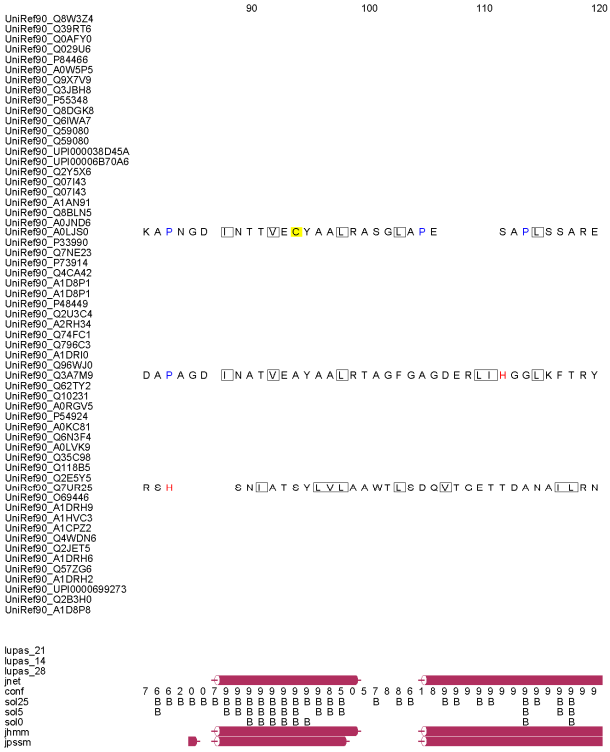


Figure S6 (continued).


90
100
110
120
UniRef90_Q8W3Z4
UniRef90_Q39RT6
UniRef90_Q0AFY0
UniRef90_Q029U6
UniRef90_P84466
UniRef90_A0W5P5
UniRef90_Q9X7V9
UniRef90_Q3JBH8
UniRef90_P55348
UniRef90_Q8DGK8
UniRef90_Q8IWA7
UniRef90_Q59080
UniRef90_Q59080
UniRef90_UPI000038D45A
UniRef90_UPI00006B70A6
UniRef90_Q2Y5X6
UniRef90_Q07I43
UniRef90_Q07I43
UniRef90_A1AN91
UniRef90_Q8BLN5
UniRef90_A0JND6
UniRef90_A0LJS0
UniRef90_P33990
UniRef90_Q7NE23
UniRef90_P73914
UniRef90_C4CAA2
UniRef90_A1D8P1
UniRef90_A1D8P1
UniRef90_P48449
UniRef90_Q2U3C4
UniRef90_A2RH34
UniRef90_Q74FC1
UniRef90_Q796C3
UniRef90_A1DRI0
UniRef90_Q96WJ0
UniRef90_Q3A7M9
UniRef90_Q82TY2
UniRef90_Q10231
UniRef90_A0RGV5
UniRef90_P54924
UniRef90_A0KC81
UniRef90_Q6N3F4
UniRef90_A0LVK9
UniRef90_Q35C98
UniRef90_Q118B5
UniRef90_Q2E5Y5
UniRef90_Q7UR25
UniRef90_O69446
UniRef90_A1DRH9
UniRef90_A1HVC3
UniRef90_A1CPZ2
UniRef90_Q4WDN6
UniRef90_Q2JET5
UniRef90_A1DRH6
UniRef90_Q57ZG6
UniRef90_A1DRH2
UniRef90_UPI0000699273
UniRef90_A1D8P8
lupas_21
lupas_14
lupas_28
jnet
conf
sol25
sol5
sol0
jhmm
jpssm

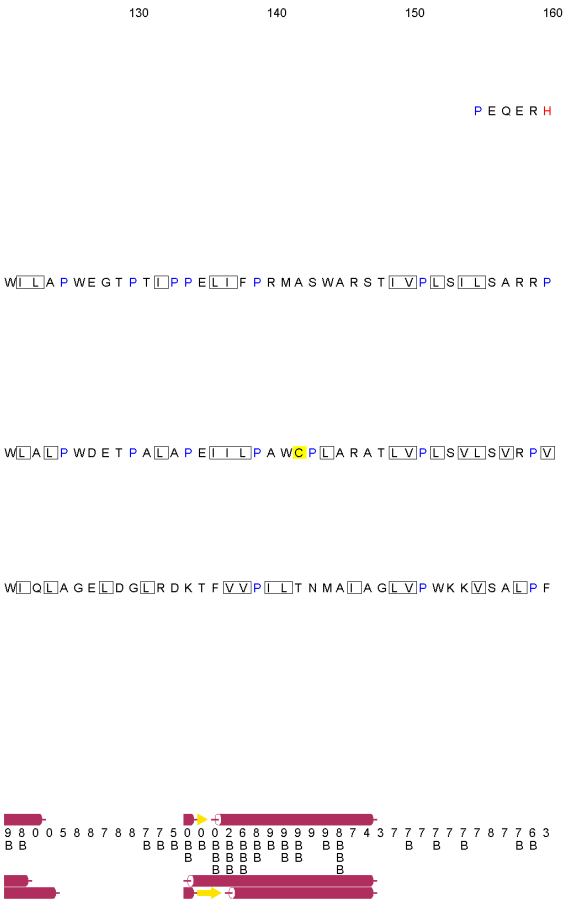


Figure S6 (continued).

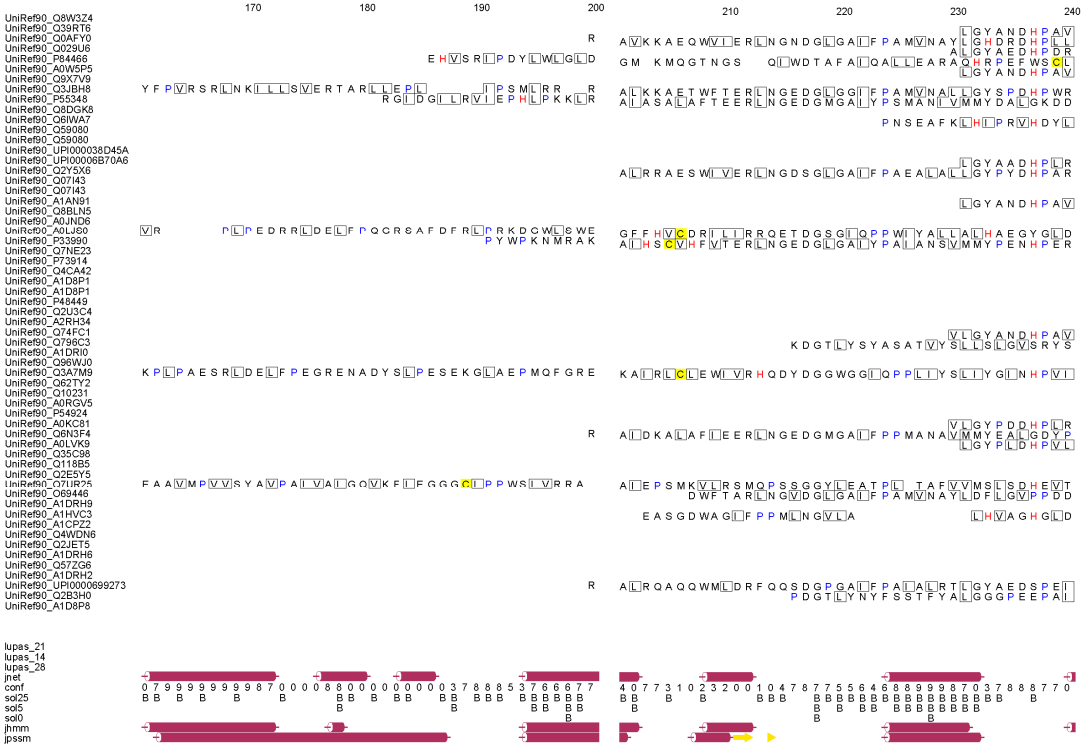


Figure S6 (continued).

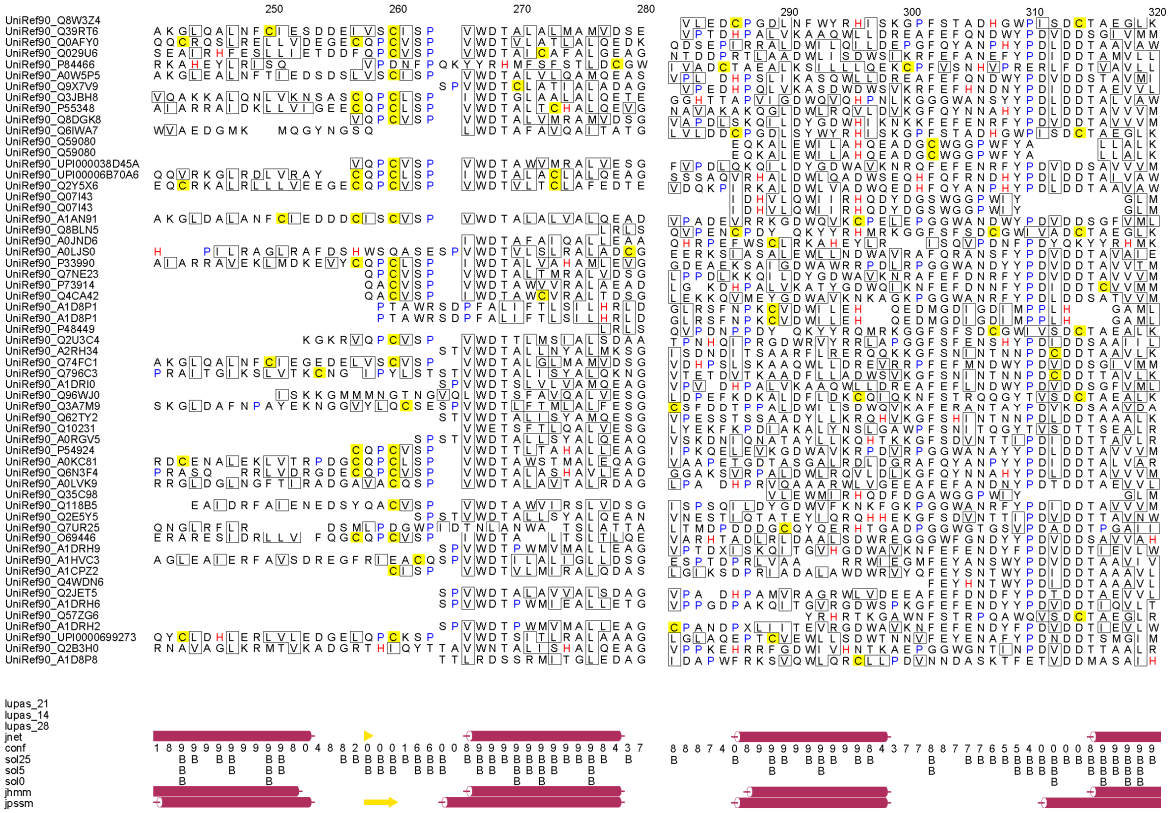


Figure S6 (continued).

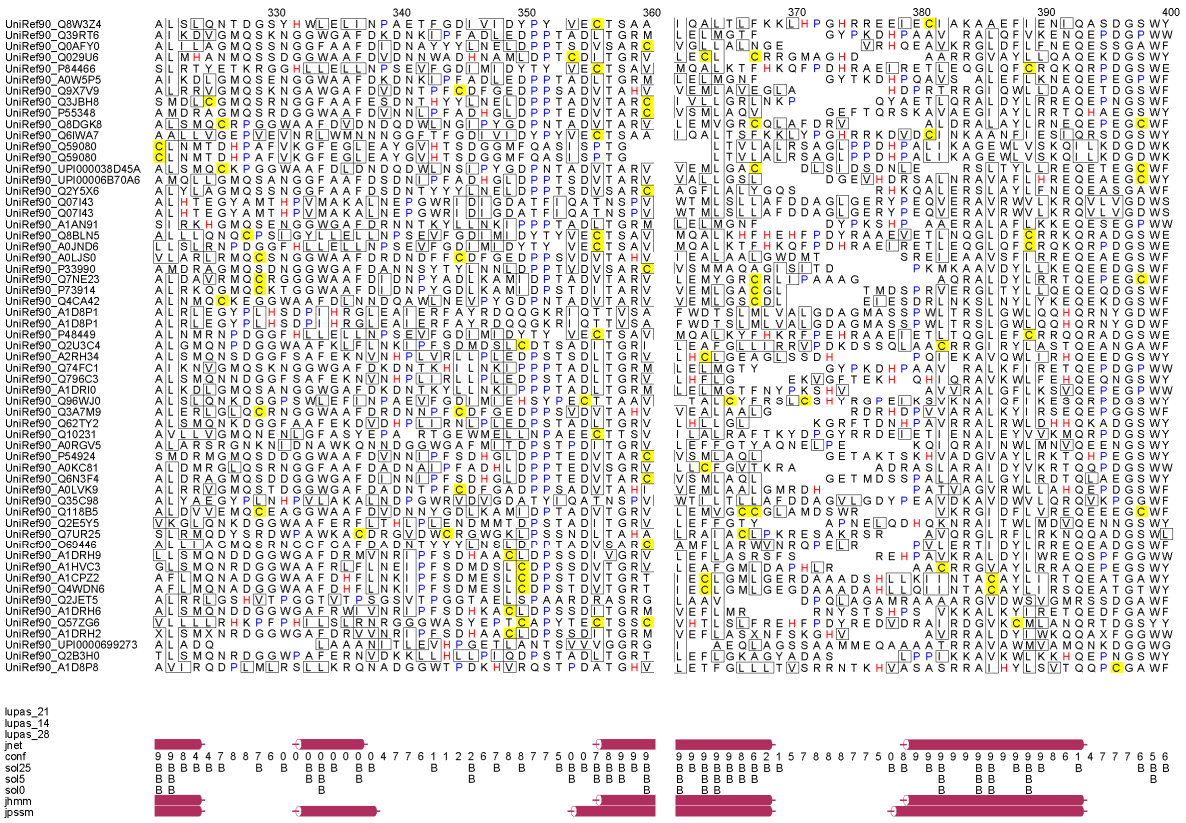


Figure S6 (continued).

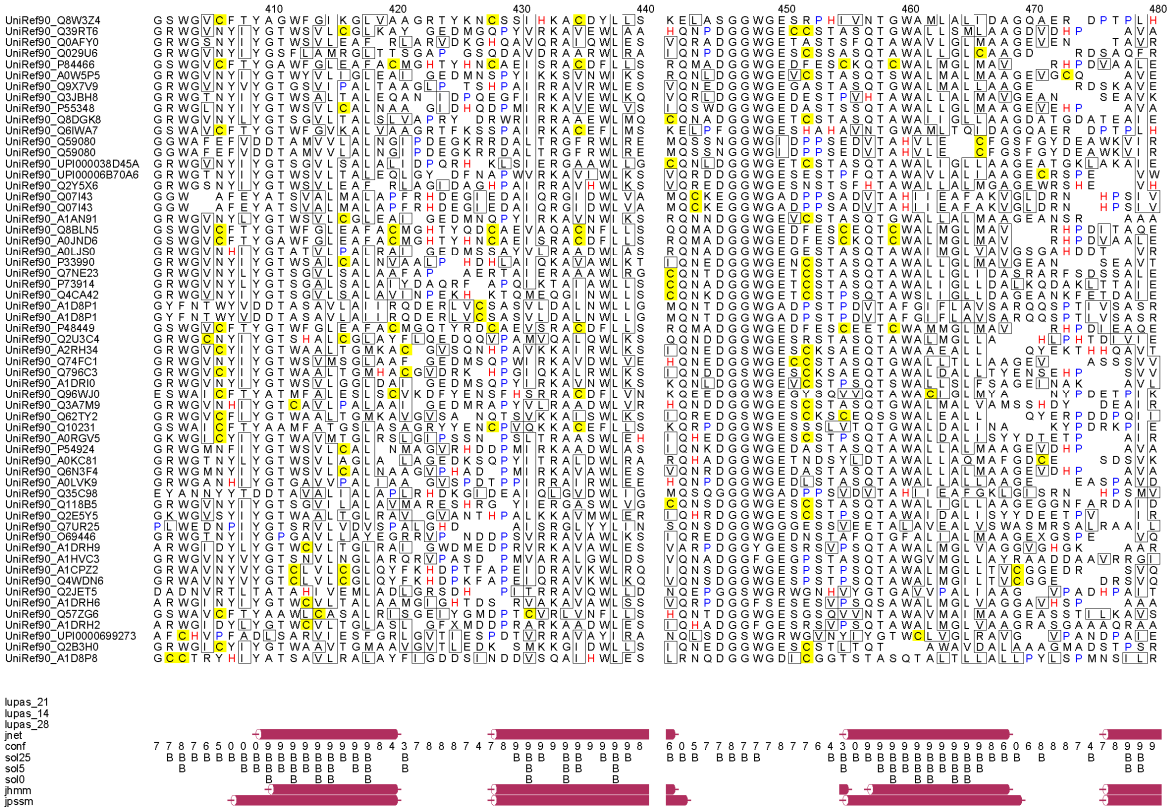


Figure S6 (continued).

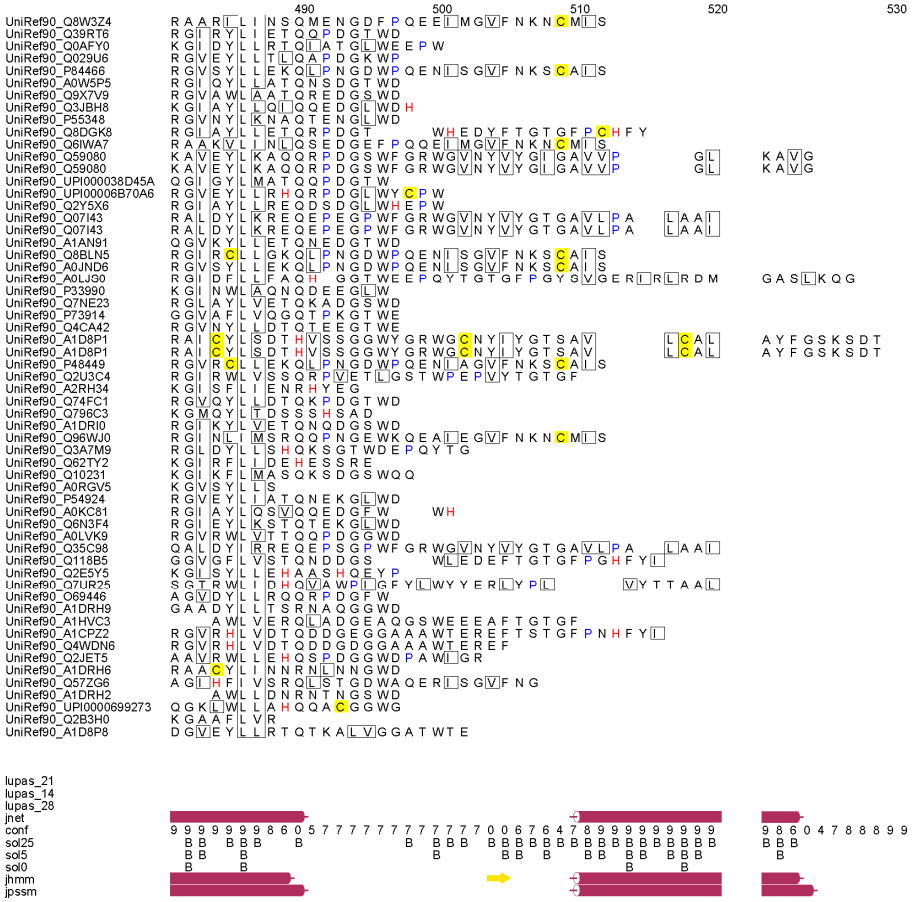


Figure S6 (continued).

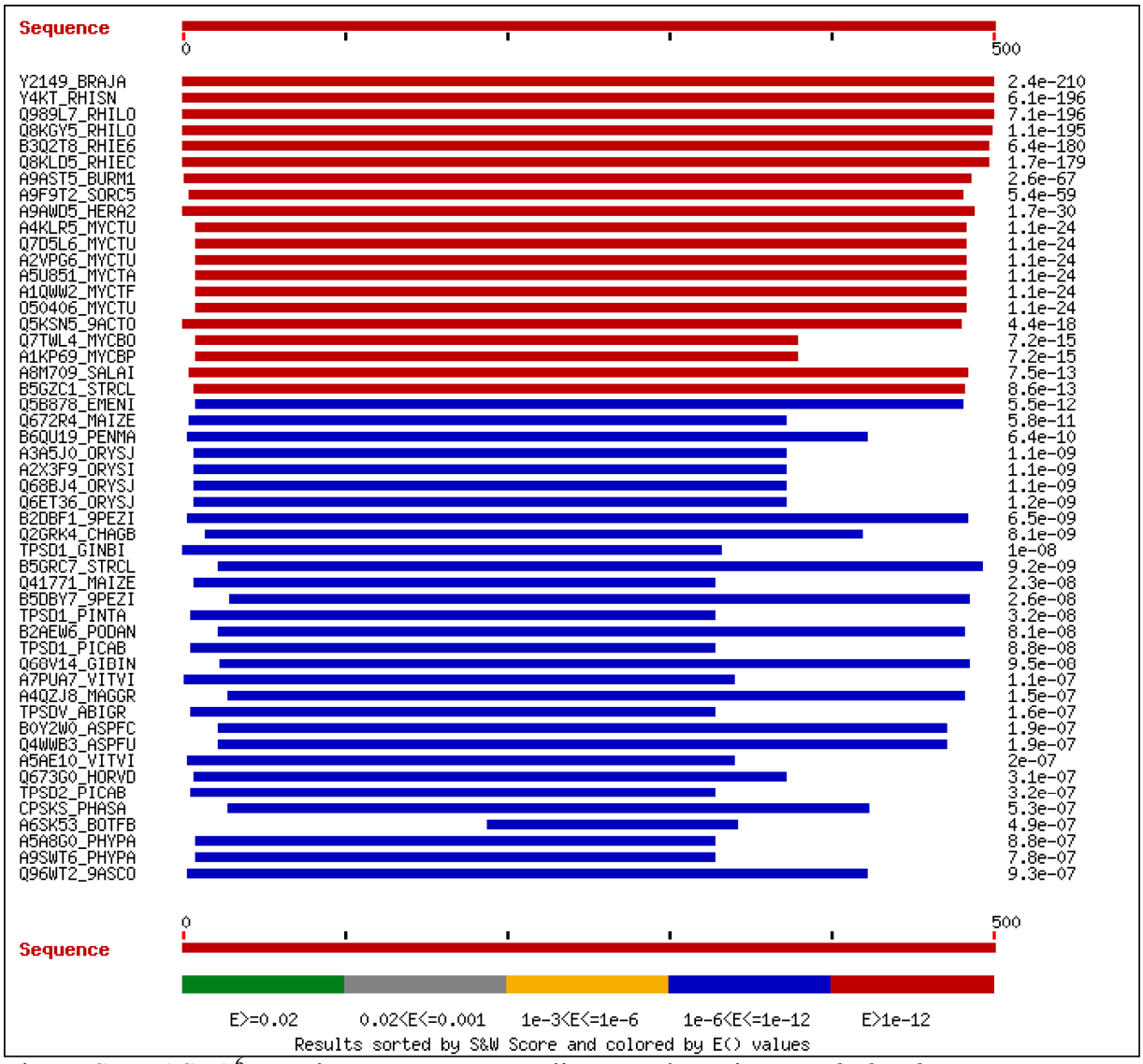


Figure S7: FASTA[6] search program output alignment by using *Bradyrhizobium japonicum* γ-domain only as the query. Note the high similarity to *M. tuberculosis* Rv3377c, rice and maize CPS.

Supporting Experimental Section:

Synthetic Aspects. All reagents used for synthesis were purchased from Aldrich (Milwaukee, WI). The purities of all compounds were routinely monitored by using ^{1}H and ^{31}P NMR spectroscopy at 400 or 500 MHz on Varian (Palo Alto, CA) Unity spectrometers using, in some instances, absolute spin-count quantitative analyses. **BPH-882** was purchased from Biomol international (Plymouth Meeting, PA); **BPH-863** was purchased from EMD Biosciences, Inc. (San Diego, CA). **BPH-891**, **BPH-854** were purchased from Sigma-Aldrich. The syntheses of **BPH-892**, **BPH-904**, **BPH-852**, **BPH-911**, **BPH-902** were reported previously[7-10]. **BPH-895** was a gift from Dr. L.L. Gundersen. **BPH-850** was prepared using the procedure reported[11-13]. **BPH-261 (Minodronate)**, **BPH-661**, **BPH-673** were available from our previous work[12, 13].

General procedure for BPH-554 and BPH-557. An amine (3 mmol), triethyl orthoformate (3.6 mmol), and diethyl phosphite (12 mmol) were heated to 140 °C under N_2 overnight. The mixture was subjected to column chromatography to give the tetraethyl ester, which was then treated with TMSBr (8 equiv) in acetonitrile for 8 h at room temperature. Upon removal of the solvent, the residue was dissolved in ethanol or acetone to precipitate the substituted aminomethylene bisphosphonate.

{[4-(2-Diethylamino-ethylcarbamoyl)-phenylamino]-phosphono-methyl}-phosphonic acid. (BPH-557).

Anal. ($C_{14}H_{27}N_3Na_2O_9P_2$) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 7.62 (d, 2H, J = 8.0Hz), 6.60 (d, 2H, J = 8.0Hz), 4.15-4.38 (m, 2H), 3.70 (t, 1H, J = 22Hz), 3.25-3.38 (m, 2H), 3.13 (t, 4H, J =7.2Hz), 1.11 (t, 4H, J = 7.2Hz). ^{31}P NMR (162 MHz, D_2O): δ 16.1.

4-[(Bis-phosphono-methyl)-amino]-benzoic acid 2-diethylamino-ethyl ester. (BPH-554).

Anal. ($C_{14}H_{27}N_3Na_2O_9P_2$) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 7.47 (d, 2H, *J* = 8.0Hz), 6.64 (d, 2H, *J* = 8.0Hz), 3.89 (t, 1H, *J* = 21Hz), 3.45-3.60 (m, 2H), 3.00-3.19 (m, 4H), 1.15(t, 4H, *J* = 7.2Hz). ^{31}P NMR (162 MHz, D_2O): δ 16.4.

BPH-876 was prepared by refluxing of 3, 4-dichlorobenzyl chloride (1 mmol) and tributyl phosphine (1 mmol) in acetone (5 mL). Removal of solvent gave BPH-876 as a white solid. Yield 89%. Anal. ($C_{19}H_{32}Cl_3P$) C, H. ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.22 (d, 1H, *J* = 8.4Hz), 7.43 (s, 1H), 7.33 (d, 1H, *J* = 8.4 Hz), 4.56 (d, 2H, *J* = 12 Hz), 2.44-2.53 (m, 6H), 1.38-1.42 (m, 12H), 0.94 (t, 3H, *J* = 6.8 Hz).

BPH-916, BPH-922, BPH-901, BPH-899, BPH-905, BPH-904, BPH-913, BPH-918, BPH-919, BPH-920 were prepared using a similar procedure to that previously reported, using alkylbromides and methylallylamine as reactants[14].

16-(N-allyl-N-methylamino)hexadecanoic acid•hydrobromide (BPH-916).

Anal. ($C_{21}H_{42}BrNO_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 6.12-6.21 (m, 1H), 5.50-5.60 (m, 2H), 3.6-3.75 (m, 2H), 2.90-3.05 (m, 2H), 2.73 (s, 3H), 2.35 (t, 2H, *J* = 7.2 Hz), 1.80-1.98 (m, 2H), 1.61-1.69 (m, 2H), 1.20-1.40 (m, 22 H).

Methyl 16-(N-allyl-N-methylamino)hexadecanoate•hydrobromide (BPH-922).

Anal. ($C_{21}H_{41}NO_2Br•0.3H_2O$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 11.5 (s, 1H), 6.14-6.19 (m, 1H), 5.51-5.60 (m, 4H), 3.66 (s, 3H), 3.55-3.70 (m, 3H), 2.81-3.09 (m, 2H), 2.30 (s, 3H), 2.30 (t, 2H, J = 6.5Hz), 1.79-2.00 (m, 2H), 1.59 (t, 2H, J = 6.5Hz), 1.24-1.32 (m, 24H).

(4-(4-(N-allyl-N-methylamino)butoxy)phenyl)(phenyl)methanone•hydrobromide (BPH-901).

Anal. ($C_{21}H_{26}BrNO$). C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.85 (d, 2H, *J* = 8 Hz), 7.69 (d, 2H, *J* = 7.2Hz), 7.47 (t, 1H, *J* = 7.2Hz), 7.45 (t, 3H, *J* =7.2Hz), 6.10-6.21 (m, 2H), 5.52-5.64 (m, 2H), 3.90 (t, 2H, *J* = 6.4 Hz), 3.60-3.67 (m, 2H), 2.80-3.18 (m, 2H), 2.77 (s, 3H), 1.68-2.00 (m, 4H).

(4-(8-(N-allyl-N-methylamino)octyloxy)phenyl)(phenyl)methanone (BPH-899).

Anal. ($C_{25}H_{34}BrNO•0.3H_2O$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.83 (d, 2H, *J* = 8.8Hz), 7.74 (d, 2H, *J* = 7.2Hz), 7.57 (t, 1H, *J* = 7.2Hz), 7.46 (t, 3H, *J* =7.2Hz), 6.14-6.21 (m, 2H), 5.51-5.61 (m, 2H), 3.62-3.68 (m, 2H), 2.80-3.18 (m, 2H), 2.73 (s, 3H), 1.78-2.00 (m, 4H), 1.32-1.49 (m, 8H).

(4-(3-(methylallylamino)propoxy)phenyl)(phenyl)methanone•hydrobromide (BPH-905).

Anal. ($C_{20}H_{24}BrNO_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 11.8 (s, 1H), 8.04 (d, 2H, *J* = 9.0 Hz), 7.7 (d, 2H, *J* = 9.0Hz), 7.59 (t, 1H, *J* = 7.5Hz), 7.56 (t, 2H, *J* = 7.5Hz), 6.92 (t, 2H, *J* = 9.0Hz), 6.19-6.23 (m, 1H), 5.56-5.65 (m, 2H), 4.18 (t, 3H, *J* = 4.8 Hz), 3.60-3.82 (m, 2H), 3.10-3.42 (m, 2H), 2.81 (s, 3H), 2.4-2.60 (m, 2H).

(E)-N-allyl-5-((4aR,5R)-decahydro-1,1,4a-trimethylnaphthalen-5-yl)-N,3-dimethylpent-2-en-1-amine•hydrobromide (BPH-904). Anal. ($C_{24}H_{44}BrNO$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 6.12-6.20 (m, 1H), 5.42-5.80 (m, 4H), 3.4-3.75 (m, 4H), 2.65 (s, 3H), 0.6-2.4

6-(m-tolyloxy)-N-allyl-N-methylhexan-1-amine (BPH-913). Anal. ($C_{17}H_{27}NO$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.18 (t, 2H, J = Hz), 6.82-6.87 (m, 3H), 5.8-6.0 (m, 1H), 5.11-5.21 (m, 2H), 3.96 (t, 2H, J = 6.8Hz), 3.00 (d, 2H, J = 6.4Hz), 2.31-2.35 (m, 5H), 2.21 (s, 3H), 1.76-1.80 (m, 2H), 1.39-4.58 (m, 6H).

10-(N-allyl-N-methylamino)decanoic acid•hydrobromide) (BPH-918). Anal. ($C_{14}H_{28}BrNO_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 6.12-6.20 (m, 1H), 5.52-5.60 (m, 2H), 3.6-3.75 (m, 2H),

2.90-3.05 (m, 2H), 2.75 (s, 3H), 2.35 (t, 2H, *J* = 7.2 Hz), 1.80-1.98 (m, 2H), 1.61-1.69 (m, 2H), 1.20-1.40 (m, 10H).

N1, N6-diallyl-N1,N6-dimethylhexane-1,6-diamine•Hydrobromide (BPH-919).

Anal.($C_{14}H_{30}Br_2N_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.81 (s, 2H), 6.08-6.17 (m, 2H), 5.54-5.59 (m, 4H), 3.43-3.78 (m, 4H), 3.16-3.18 (m, 4H), 3.02-3.04 (m, 4H), 2.7 (s, 6H), 1.87-1.98 (m, 2H), 1.50-1.53 (m, 4H).

N1, N8-diallyl-N1, N8-dimethyloctane-1, 8-diamine•hydrobromide (1:2) (BPH-920).

Anal.($C_{16}H_{34}Br_2N_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.96 (s, 2H), 6.09-6.18 (m, 2H), 5.53-5.70 (m, 4H), 3.45-3.79 (m, 4H), 3.16-3.18 (m, 4H), 3.02-3.04 (m, 4H), 2.7 (s, 6H), 1.83-1.98 (m, 2H), 1.34-1.53 (m, 8H).

BPH-890, **BPH-921**, **BPH-917**, **BPH-907**, **BPH-909**, **BPH-908**, **BPH-924**, **BPH-896**, **BPH-923** were prepared by refluxing different alkyl bromides in pyridine followed by removal of solvent and recrystallization from acetone.

[(4-phenylcarbonyl)-phenyoxyoctyl]pyridinium bromide (BPH-890). Anal. ($C_{26}H_{30}BrNO_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.54 (d, 2H, *J* = 5.6 Hz), 8.50 (t, 1H, *J* = 8 Hz), 8.10 (t, 2H, *J* = 6.8 Hz), 7.80 (d, 2H, *J* = 6.8Hz), 7.60 (d, 2H, *J* = 6.8Hz), 7.55 (t, 1H, *J* = 7.6Hz), 7.50 (t, 2H, *J* = 7.6 Hz), 6.95 (d, 2H, *J* = 8 Hz), 5.05 (t, 2H, *J* = 7.6 Hz), 4.02 (t, 2H, *J* = 6.4 Hz), 2.04-2.08 (m, 2H), 1.74-1.82 (m, 2H), 1.36-1.46 (m, 8H).

[6-(4-(formylphenyl)(phenoxy))hexyl]pyridinium bromide (BPH-921). Anal. ($C_{24}H_{26}BrNO_2$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.60 (d, 2H, *J* = 5.6 Hz), 8.46 (t, 1H, *J* = 7.6Hz), 8.07 (t, 2H, *J* = 7.2Hz), 7.64-7.17 (m, 4H), 7.50 (t, 1H, *J* = 7.8Hz), 7.42 (t, 2H, *J* = 7.8Hz), 6.90 (t, 2H, *J* = 8.8Hz), 5.00 (t, 2H, *J* = 7.2 Hz), 3.95 (t, 2H, *J* = 6.0 Hz), 2.00- 2.07 (m, 2H), 1.73-1.78 (m, 2H), 1.43-1.47 (m, 4H).

[1-carbamoylhexadecanyl]pyridinium chloride (BPH-917). Anal. ($C_{21}H_{36}ClNO_2$•0.3H_2O) C, H, N. ^{1}H NMR (400 MHz, CD3OD): δ 9.56 (d, 2H, *J* = 5.6 Hz), 8.42 (t, 1H, *J* = 7.6 Hz), 8.01 (t, 2H, *J* = 6.8 Hz), 4.9 (t, 2H, *J* = 7.6 Hz), 2.60 (t, 2H, *J* = 7.2Hz), 2.2-2.4 (m, 2H), 1.30-2.4 (m, 24H).

[6-(4-(methoxycarbonyl)-3-hydroxyphenoxy)hexyl]-pyridinium bromide (BPH-907). Anal . ($C_{19}H_{24}BrNO4$•1.4H_2O) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 8.74 (d, 2H, *J* = 6Hz), 8.35 (t, 1H, J = 7 Hz), 7.82 (t, 2H, J = 7Hz), 7.54 (d, 1H, J = 9Hz), 6.23 (d, 1H, J = 9Hz), 6.20 (s, 1H), 4.31 (t, 2H, J = 7.5Hz), 3.80 (t, 2H, J= 6Hz), 3.70 (s, 3H), 1.82-1.85 (m, 2H), 1.50-1.80 (m, 2H), 1.29-1.51 (m, 4H).

[8-(4-carbonyl)phenoxy)octyl] pyridinium bromide (BPH-909). Anal. ($C_{20}H_{26}BrNO_3$) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 8.63 (d, 2H, *J* = 5.5Hz), 8.33 (t, 1H, *J* = 8 Hz), 7.82 (t, 2H, *J* = 6.5Hz), 7.78 (d, 1H, *J* = 9Hz), 6.84 (d, 1H, *J* = 8.5Hz), 4.37 (t, 2H, *J* = 7.5Hz), 3.92 (t, 2H, *J* = 6.5Hz), 1.75-1.82 (m, 2H), 1.55-1.80 (m, 2H), 1.29-1.51 (m, 8H).

1-[8-(4-(methoxycarbonyl)phenoxy)octyl]-pyridinium bromide (BPH-908). Anal . ($C_{21}H_{28}BrNO3$•H_2O) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 8.66 (d, 2H, *J* = 5.5Hz), 8.33 (t, 1H, *J* = 8 Hz), 7.86 (t, 2H, *J* = 6.5 Hz), 7.52 (d, 1H, *J* = 8.5 Hz), 6.52 (d, 1H, *J* = 8.5 Hz), 4.37 (t, 2H, *J* = 7.5 Hz), 3.66 (s, 3H), 3.51 (t, 2H, *J* = 6.5 Hz), 1.71-1.75 (m, 2H), 1.30-1.55 (m, 2H), 0.98-1.20 (m, 8H).

[6-(4-(phenylcarbonyl)(phenoxy))propyl]pyridinium bromide (BPH-924).

Anal. ($C_{21}H_{20}NO_2Br$) C, H, N. ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.62 (d, 2H, *J* = 6 Hz), 8.48 (t, 1H, *J* = 8.0 Hz), 8.04 (t, 2H, *J* = 8.0 Hz), 7.60-7.63 (m, 4H), 7.48 (t, 1H, *J* = 7.5 Hz), 7.36 (t, 2H, *J* = 7.5 Hz), 6.73 (d, 2H, *J* = 7.5 Hz), 5.27 (t, 2H, *J* = 7.0 Hz), 4.23 (t, 2H, *J* = 5.5 Hz), 2.57-2.62 (m, 2H).

[6-(4-(carboxy)-3-hydroxyphenoxy)hexyl]-pyridinium chloride (BPH-923A). Anal. ($C_{18}H_{22}ClNO_4$) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 8.64 (d, 2H, *J* = 6Hz), 8.32 (t, 1H, J = 8 Hz), 7.82 (t, 2H, J = 7Hz), 7.50 (d, 1H, J = 9Hz), 6.26(d, 1H, J = 9Hz), 6.18 (s, 1H), 4.41 (t, 2H, J = 7.5Hz), 3.80 (t, 2H, J= 6Hz), 1.80-1.84 (m, 2H), 1.49-1.80 (m, 2H), 1.19-1.51 (m, 4H).

General procedure for BPH-888 and BPH-914. To a solution of malonic acid, monoethyl ester (1 mmol) were added N-ethyl-N-(3-dimethylaminopropyl)-carbodiimide (EDC) (1.5 mmol) and 1-hydroxybenzotriazole (1 mmol). After stirring for 2h at room temperature, 50 mL of ethyl acetate was added and the reaction mixture washed successively with 1N HCl (5 mL), water (5 mL) and saturated $NaHCO_3$ (5 mL), dried, and evaporated. The amide was purified using flash chromatography.

N-[3-(3-Phenoxy-phenyl)-propyl]-melonic acid monoamide (BPH-888). 3-(3-phenoxy-phenyl)-propylamine was coupled with monoethyl ester of malonic acid to give the ethyl ester of BPH-888, which was then hydrolyzed with 3 equiv. of KOH in MeOH/H_2O for 1h. The reaction mixture was acidified, extracted with ethyl acetate, and the organic layer evaporated. The oily residue was dissolved in methanol, neutralized with KOH and evaporated to give **BPH-888** as a white powder (250 mg, 66% overall yield). Anal . ($C_{18}H_{18}KNO4\bullet0.5H_2O\bullet0.25KCl$) C, H, N. ^{1}H NMR (400 MHz, D_2O): δ 6.70-7.40 (m, 9H), 3.33-3.40 (m, 2H), 3.32 (s, 2H,), 2.62 (t, 2H, *J* = 7.6 Hz), 1.80-1.90 (m, 2H).

(4-(6-bromohexyloxy)phenylcarbamoyl)methylphosphonic acid (BPH-914). 4-(6-bromohexyl)aniline was coupled with dibenzyl phosphonoacetic acid, according to the general method, to give the dibenzyl ester of **BPH-914**. The benzyl groups were removed by hydrogenation for 1hr, catalyzed with 5% Pd/C in methanol, followed by neutralization with NaOH to give **BPH-914** as a white powder (289 mg, 60%). Anal. ($C_{14}H_{19}NaNO_5\bullet2.5H_2O$) C, H,

N. ^{1}H NMR (400 MHz, D_2O): δ 7.22 (d, 2H, *J* = 9Hz), 3.90 (t, 3H, *J* = 7Hz), 3.45 (t, 3H, *J* = 7Hz), 2.51 (d, 2H, *J* = 18.5Hz), 1.56-1.62 (m, 4H,), 1.28-1.32 (m, 4H).

References:

1. Fuchs, P. F.; Alix, A. J., High accuracy prediction of beta-turns and their types using propensities and multiple alignments. *Proteins* **2005,** 59, (4), 828-39.

2. Krissinel, E.; Henrick, K., Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 12 Pt 1), 2256-68.

3. Shindyalov, I. N.; Bourne, P. E., Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* **1998,** 11, (9), 739-747.

4. Cuff, J. A.; Clamp, M. E.; Siddiqui, A. S.; Finlay, M.; Barton, G. J., JPred: a consensus secondary structure prediction server. *Bioinformatics* **1998,** 14, (10), 892-893.

5. Vogel, B. S.; Wildung, M. R.; Vogel, G.; Croteau, R., Abietadiene synthase from grand fir (*Abies grandis*). cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *J Biol Chem* **1996,** 271, (38), 23262-8.

6. Pearson, W. R.; Lipman, D. J., Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* **1988,** 85, (8), 2444-8.

7. Roy, A.; Roberts, F. G.; Wilderman, P. R.; Zhou, K.; Peters, R. J.; Coates, R. M., 16-Aza-*ent*-beyerane and 16-Aza-*ent*-trachylobane: potent mechanism-based inhibitors of recombinant ent-kaurene synthase from *Arabidopsis thaliana*. *J Am Chem Soc* **2007,** 129, (41), 12453-60.

8. Ravn, M. M.; Jin, Q.; Coates, R. M., Synthesis of allylic isoprenoid diphosphates by SN_2 displacement of diethyl phosphate. *Eur. J. Org. Chem.* **2000,** 2000, (8), 1401-10.

9. Peters, R. J.; Ravn, M. M.; Coates, R. M.; Croteau, R. B., Bifunctional abietadiene synthase: free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites. *J Am Chem Soc* **2001,** 123, (37), 8974-8.

10. Prisic, S.; Xu, J.; Coates, R. M.; Peters, R. J., Probing the role of the DXDD motif in Class II diterpene cyclases. *Chembiochem* **2007,** 8, (8), 869-874.

11. Wiemer, A. J.; Yu, J. S.; Shull, L. W.; Barney, R. J.; Wasko, B. M.; Lamb, K. M.; Hohl, R. J.; Wiemer, D. F., Pivaloyloxymethyl-modified isoprenoid bisphosphonates display enhanced inhibition of cellular geranylgeranylation. *Bioorg Med Chem* **2008,** 16, (7), 3652-60.

12. Szabo, C. M.; Martin, M. B.; Oldfield, E., An investigation of bone resorption and *Dictyostelium discoideum* growth inhibition by bisphosphonate drugs. *J Med Chem* **2002,** 45, (14), 2894-903.

13. Liu, C. I.; Liu, G. Y.; Song, Y.; Yin, F.; Hensler, M. E.; Jeng, W. Y.; Nizet, V.; Wang, A. H.; Oldfield, E., A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science* **2008,** 319, (5868), 1391-1394.

14. Dehmlow, H.; Aebi, J. D.; Jolidon, S.; Ji, Y. H.; von der Mark, E. M.; Himber, J.; Morand, O. H., Synthesis and structure-activity studies of novel orally active non-terpenoic 2,3-oxidosqualene cyclase inhibitors. *J Med Chem* **2003,** 46, (15), 3354-70.