Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

Computational design of cephradine synthase in a new scaffold identified from structural databases

Xiaoqiang Huang, [a] Jing Xue, [b] and Yushan Zhu* [a,b]

[a]: Department of Chemical Engineering, Tsinghua University, Beijing 100084, People's Republic of China

[b]: MOE Key lab of Industrial Biocatalysis, Tsinghua University, Beijing 100084, People's Republic of China

*: Corresponding to: yszhu@tsinghua.edu.cn

1. Supplementary Tables

Table S1. Catalytic geometrical constraint parameters for match.

Interacting Pair	Constraint type	Atom 1	Atom2 [a]	Atom3 [a]	Atom4 [a]	Ideal value [b]	Standard deviation [c]
SER1-TS	Distance	OG	#C15			1.5	0.1
	Angle	CB	OG	#C15		110.0	5.0
	Angle	OG	#C15	#O16		110.0	5.0
	Torsion	OG	#O16	#C15	#N14	120.0	0.0
HIS2-SER1	Distance	NE2	#OG			2.8	0.4
	Angle	NE2	#OG	#CB		110.0	40.0
	Angle	CD2	NE2	#OG		120.0	40.0
ASP3-HIS2	Distance	OD1	#ND1			2.8	0.4
	Angle	OD1	#ND1	#CG		120.0	40.0
	Angle	CG	OD1	#ND1		120.0	40.0
	Angle	OD1	#HD1	#ND1		180.0	60.0
ALA4-TS	Distance	N	#O16			2.8	0.4
	Angle	N	#O16	#C15		110.0	40.0
	Angle	CA	N	#O16		120.0	40.0
TYR5-TS	Distance	OH	#O16			2.8	0.4
	Angle	OH	#O16	#C15		110.0	40.0
	Angle	CZ	OH	#O16		120.0	40.0
SER6-ASP3	Distance	OG	#OD2			2.8	0.4
	Angle	OG	#OD2	#CG		120.0	40.0
	Angle	CB	OG	#OD2		110.0	40.0
GLU7-TS	Distance	OE1	#N24			2.8	0.4
	Angle	OE1	#N24	#C17		110.0	40.0
	Angle	CD	OE1	#N24		120.0	40.0

[a]: the symbol '#' means the atom is on the second residue of the interacting pair.

[b]: the ideal values are obtained from atom hybrid state.

[c]: the distances and angles are varied by 0.1 angstroms and 5.0 degrees and torsion angle is fixed for covalent bond. The distances and angles are varied by 0.4 angstroms and 40.0 degrees for hydrogen bond.

Table S2. List of 2,234 protein scaffolds used for match.

12as, 1a05, 1a0c, 1a0e, 1a0i, 1a2o, 1a2z, 1a47, 1a4i, 1a59, 1a5z, 1a7t, 1a8r, 1a8s, 1aa6, 1ad3, 1agy, 1aj8, 1ajk, 1ajo, 1aky, 1al8, 1aln, 1alq, 1amu, 1ao0, 1apx, 1aqj, 1aqm, 1aqu, 1aui, 1auo, 1avq, 1aw1, 1ax4, 1axk, 1ayx, 1azw, 1b0a, 1b1y, 1b48, 1b4p, 1b5p, 1b65, 1b6c, 1b6g, 1b7g, 1b80, 1b8a, 1b8x, 1b9b, 1b9h, 1bbu, 1bd3, 1bdg, 1bf6, 1bg5, 1bg6, 1bi9, 1bif, 1bkp, 1bou, 1bqg, 1bqy, 1bry, 1bs0, 1bs2, 1bsl, 1bue, 1bx1, 1bx4, 1bxb, 1bxc, 1bxk, 1by8, 1byg, 1bzl, 1c24, 1c3c, 1c4k, 1c4x, 1c4z, 1c7g, 1c7q, 1c8x, 1c9w, 1cbf, 1ccw, 1cf2, 1cfr, 1cg2, 1chd, 1ci0, 1ci9, 1cib, 1cjm, 1cl0, 1cm8, 1cnz, 1cru, 1ct9, 1cy9, 1czi, 1d1t, 1d2c, 1d2n, 1d2t, 1d4a, 1d6f, 1d6m, 1d6s, 1d7o, 1d8w, 1d9e, 1d9q, 1dad, 1db3, 1dbt, 1dci, 1ddg, 1ddj, 1dek, 1dfa, 1di1, 1dj0, 1dj2, 1dj3, 1djl, 1dki, 1dl5, 1dlj, 1dmh, 1dos, 1dpg, 1dpj, 1dpo, 1dqs, 1drw, 1dug, 1duv, 1dxq, 1dxy, 1e0c, 1e0t, 1e1h, 1e1m, 1e1o, 1e25, 1e2k, 1e2o, 1e2t, 1e3i, 1e4c, 1e4e, 1e4i, 1e58, 1e5m, 1e5q, 1e6u, 1e6w, 1e7w, 1e89, 1e8g, 1e93, 1e94, 1e9g, 1e9i, 1ebd, 1ebf, 1ec7, 1edo, 1edz, 1eh6, 1ei5, 1ejd, 1ems, 1eov, 1ep3, 1eq2, 1eqr, 1esw, 1eud, 1euh, 1euz, 1ev4, 1evy, 1eye, 1eyq, 1ez0, 1ez4, 1ezr, 1f05, 1f0y, 1f20, 1f28, 1f2d, 1f2j, 1f2v, 1f34, 1f3a, 1f3l, 1f52, 1f5a, 1f5q, 1f60, 1f6d, 1f6w, 1f75, 1f76, 1f82, 1f8m, 1f8w, 1f9v, 1fdr, 1fec, 1fgg, 1fhe, 1fi8, 1fj2, 1fl2, 1fmc, 1fmv, 1fp1, 1fp2, 1fp6, 1frb, 1fur, 1fva, 1fwy, 1fxw, 1fye, 1g0d, 1g0h, 1g0o, 1g3n, 1g55, 1g5a, 1g6a, 1g6c, 1g6q, 1g6s, 1g8m, 1g8p, 1ga6, 1gad, 1gbg, 1gcy, 1gee, 1geg, 1geq, 1ghp, 1gis, 1gk9, 1gkm, 1gm7, 1goj, 1got, 1gp6, 1gpm, 1gpp, 1gq2, 1gs5, 1gsa, 1gsu, 1gtk, 1gv0, 1gvf, 1gwu, 1gx3, 1gxy, 1gyt, 1gz6, 1h0d, 1h1y, 1h3f, 1h3i, 1h5q, 1h5y, 1h6d, 1h72, 1hdh, 1hk8, 1hm9, 1hp1, 1hqs, 1hsk, 1hso, 1hw4, 1hxp, 1hye, 1hyu, 1hzo, 1i0d, 1i24, 1i2k, 1i32, 1i4n, 1i5e, 1i7q, 1i8t, 1iat, 1ibv, 1igw, 1iho, 1ihu, 1ii2, 1iic, 1iip, 1io7, 1is8, 1itx, 1itz, 1iu4, 1iu8, 1iuq, 1ivh, 1iy7, 1iy8, 1iye, 1iyk, 1j0b, 1j0h, 1j1b, 1j20, 1j2p, 1j2q, 1j2z, 1j33, 1j3b, 1j4a, 1j6u, 1j70, 1j8u, 1j93, 1j97, 1j99, 1ja1, 1jbq, 1jcl, 1jdc, 1jdw, 1je0, 1jfb, 1jfl, 1jg1, 1jhd, 1jhf, 1ji1, 1jil, 1jkm, 1jl0, 1jln, 1jnw, 1jof, 1jqi, 1js1, 1jub, 1jvb, 1jzd, 1k0w, 1k28, 1k2w, 1k38, 1k3y, 1k75, 1k89, 1k92, 1k9o, 1ka9, 1kbi, 1kbj, 1kbv, 1kcb, 1kcz, 1kea, 1keq, 1kf6, 1kfg, 1kfi, 1kg2, 1khb, 1kjq, 1kk0, 1kl7, 1kmj, 1kny, 1kol, 1kop, 1kp8, 1krh, 1kux, 1kwp, 1kxq, 1kyq, 1kyz, 1kzl, 1l0o, 1l1f, 1l1l, 1l2u, 1l5w, 1l6j, 1l6s, 1l7d, 1l8k, 1l9x, 1lbq, 1lbv, 1lc0, 1lci, 1lf2, 1lfk, 1lh0, 1li4, 1lii, 1ll7, 1lp8, 1lqa, 1lqt, 1lt3, 1lth, 1ltq, 1lvo, 1lwj, 1lzj, 1m0u, 1m15, 1m2k, 1m2x, 1m40, 1m6j, 1m7s, 1m9h, 1mas, 1mdb, 1mdo, 1meg, 1mhm, 1mj3, 1mj5, 1ml4, 1mla, 1mlw, 1mo0, 1moq, 1mpx, 1ms6, 1muw, 1mvh, 1mvl, 1mw9, 1mx3, 1mxr, 1mza, 1mzh, 1mzy, 1n1b, 1n2t, 1n40, 1n45, 1n7o, 1nb5, 1nd4, 1nec, 1ney, 1nff, 1nio, 1nj1, 1nm8, 1nmm, 1nnl, 1np7, 1npy, 1nr6, 1nrf, 1nrg, 1nsj, 1nsn, 1nu5, 1nu7, 1nuy, 1nv8, 1nvm, 1nvt, 1nw6, 1nw9, 1nxu, 1nyl, 1nyt, 1nzj, 1nzy, 1o4s, 1o5o, 1o5x, 1o60, 1o6e, 1o6i, 1o7q, 1o7x, 1o88, 1o94, 1o98, 1o9j, 1ob3, 1oc4, 1odm, 1odt, 1oe8, 1ogp, 1ohl, 1oht, 1oi7, 1oj4, 1okg, 1omo, 1on3, 1onf, 1onw, 1op8, 1oq9, 1oqf, 1or0, 1otj, 1ovm, 1owl, 1ox6, 1oy5, 1p15, 1p3d, 1p4c, 1p5g, 1p5x, 1p6x, 1p77, 1p7t, 1p9p, 1pa2, 1pam, 1pb1, 1pbg, 1pe9, 1peg, 1pfz, 1pg5, 1phk, 1pix, 1pj3, 1pj9, 1pjr, 1pjs, 1pkh, 1pkx, 1pn2, 1po5, 1ps1, 1ps9, 1pv8, 1pwh, 1pyf, 1pym, 1pz3, 1pzs, 1q1a, 1q1q, 1q20, 1q23, 1q5d, 1q5m, 1q5n, 1q5q, 1q6o, 1q6z, 1q7b, 1q7e, 1q8y, 1qak, 1qam, 1qca, 1qd1, 1qd5, 1qd6, 1qdl, 1qe3, 1qgj, 1qh4, 1qh5, 1qhh, 1qj5, 1qk1, 1qk3, 1qle, 1qmg, 1qo5, 1qqj, 1qre, 1qsg, 1qtw, 1qwl, 1qx4, 1qxy, 1qyr, 1qzu, 1r0v, 1r12, 1r3s, 1r5b, 1r6w, 1r76, 1r9o, 1ra0, 1re0, 1req, 1rf6, 1rg9, 1rgy, 1rgz, 1rii, 1rjw, 1rkd, 1rpm, 1rpx, 1rqb, 1rqq, 1rtf, 1rtq, 1ru4, 1rxt, 1ry2, 1ryb, 1ryd, 1s1p, 1s70, 1sc6, 1scj, 1sg4, 1sgv, 1shz, 1slm, 1smk, 1snn, 1snz, 1sov, 1svu, 1swv, 1syy, 1sz2, 1t10, 1t1u, 1t2a, 1t2f, 1t3b, 1t3i, 1t4b, 1t5c, 1t7q, 1t90, 1tb4, 1te2, 1tel, 1tfx, 1thf, 1tjv, 1tks, 1tqj, 1txg, 1tzj, 1tzs, 1u2g, 1u2k, 1u3d, 1u3i, 1u3u, 1u3w, 1u6r, 1u7t, 1u7z, 1u8f, 1uas, 1ubv, 1uch, 1ued, 1uek, 1uf5, 1ujn, 1ujq, 1umd, 1umk, 1uok, 1uou, 1upi, 1uqt, 1urh, 1us0, 1uuf, 1uxj, 1uyp, 1uzm, 1uzr, 1v25, 1v5i, 1v7z, 1v84, 1v8b, 1v8f, 1v8k, 1v9p, 1vb3, 1vbj, 1vcv, 1vdc, 1vem, 1vf1, 1vfr, 1vhc, 1vhw, 1vi2, 1vl2, 1vlc, 1vlo, 1vlv, 1vm6, 1vpe, 1vpx, 1vqt, 1vr6, 1vrp, 1vs3, 1vzv, 1vzw, 1w0d, 1w0h, 1w23, 1w27, 1w32, 1w5q, 1w6t, 1w6u, 1w85, 1w8o, 1w8s, 1w98, 1w9y, 1wch, 1wdk, 1wdp, 1wl4, 1wle, 1wm1, 1wos, 1wpo, 1wpw, 1wsd, 1wsr, 1wta, 1wvf, 1wvg, 1wy2, 1wyd, 1wzl, 1x0l, 1x0v, 1x12, 1x13, 1x1i, 1x54, 1x7d, 1x9i, 1xa1, 1xah, 1xel, 1xfb, 1xff, 1xi3, 1xk7, 1xk9, 1xkl, 1xlm, 1xm8, 1xmc, 1xp3, 1xql, 1xro, 1xrs, 1xsm, 1xtt, 1xu9, 1xx4, 1xyz, 1y2m, 1y42, 1y6b, 1y79, 1y7o, 1y9a, 1yb1, 1yc0, 1yc5, 1ych, 1ye9, 1yfm, 1yis, 1ykf, 1yl7, 1ylh, 1ym3, 1yoe, 1yqq, 1ytm, 1yvg, 1yvj, 1yw9, 1z41, 1z5h, 1z8o, 1z9h, 1zai, 1zbu, 1zc0, 1zcj, 1zcz, 1zi8, 1zl9, 1zlp, 1zm8, 1zmd, 1zmr, 1zod, 1zoi, 1zq1, 1zrn, 1zuw, 1zws, 1zxx, 1zy4, 1zy8, 1zym, 2a2a, 2a2r, 2a9d, 2a9v, 2aeb, 2aee, 2ag5, 2aio, 2akj, 2akz, 2apo, 2ash, 2aus, 2axq, 2ay1, 2ayh, 2ayq, 2b3o, 2b4g, 2b7n, 2b9h, 2bfd, 2bgs, 2bh9, 2bhg, 2bht, 2bi7, 2bjf, 2bji, 2bju, 2bkr, 2bmb, 2bmw, 2bo9, 2bog, 2bp7, 2bpo, 2btm, 2bty, 2buf, 2buj, 2bur, 2bvr, 2bwp, 2bz6, 2bzl, 2bzr, 2c0a, 2c0r, 2c1h, 2c2n, 2c2x, 2c31, 2c3z, 2c47, 2c4e, 2c4j, 2c53, 2c5a, 2c61, 2c7s, 2c81, 2c8j, 2c8n, 2c9y, 2cb5, 2cby, 2cca, 2cch, 2cdn, 2cev, 2cfo, 2cfv, 2cg8, 2cio, 2cjz, 2clq, 2cm2, 2cmw, 2coi, 2cun, 2cxe, 2cxn, 2czc, 2czh,

2d0j, 2d1f, 2d1s, 2d4a, 2d7j, 2dbq, 2dfd, 2dfi, 2dfv, 2dho, 2dkv, 2dpl, 2dq0, 2dq3, 2dq4, 2dw4, 2dwc, 2dyu, 2e01, 2e1z, 2e28, 2e5v, 2e8y, 2e9l, 2eer, 2egz, 2ehj, 2ej9, 2eja, 2ejc, 2ep7, 2es4, 2et6, 2eu9, 2ew8, 2ey4, 2ez1, 2f00, 2fbm, 2fdv, 2fh7, 2fok, 2fp3, 2fpq, 2fw2, 2fyf, 2fze, 2g24, 2g36, 2g76, 2g8s, 2gao, 2gcg, 2gdq, 2gdz, 2gf3, 2gi3, 2gjt, 2gk9, 2gke, 2gn0, 2go1, 2gp6, 2gq1, 2gru, 2gry, 2gsq, 2gtr, 2gwh, 2h0v, 2h12, 2h4v, 2h58, 2h63, 2hak, 2hc1, 2hc9, 2he3, 2heh, 2hej, 2hg2, 2hgs, 2hhp, 2hjh, 2hk5, 2hnl, 2hro, 2hsa, 2hw5, 2hw6, 2hwg, 2hxt, 2i1y, 2i4l, 2i5p, 2i6l, 2i6u, 2i87, 2iag, 2ib8, 2ifc, 2ifr, 2ify, 2ija, 2ijm, 2inf, 2inr, 2irp, 2isq, 2iu8, 2iv0, 2iw2, 2iw5, 2iwb, 2iwz, 2ix4, 2ix5, 2ixc, 2iz1, 2izr, 2j07, 2j0i, 2j1q, 2j27, 2j32, 2j4d, 2j5i, 2j66, 2j6i, 2j6l, 2j78, 2j7t, 2j8g, 2j91, 2ja2, 2jan, 2jar, 2jay, 2jbh, 2jbm, 2jbv, 2jc3, 2jc6, 2jf2, 2jfn, 2jfq, 2jgq, 2jhq, 2jif, 2jjk, 2jjn, 2jkb, 2kin, 2nlk, 2nmx, 2no4, 2no7, 2ntp, 2nu8, 2nxw, 2nya, 2o0b, 2o2c, 2o2e, 2o36, 2o3c, 2o3e, 2o3j, 2o4c, 2o5r, 2o7r, 2o7s, 2o8m, 2o9p, 2oat, 2obv, 2oc3, 2ocp, 2ode, 2ody, 2oem, 2ofp, 2oh4, 2ohh, 2olq, 2oo0, 2oo8, 2ooq, 2ort, 2ose, 2otn, 2p0c, 2p0r, 2p3e, 2p4q, 2p6x, 2p91, 2p9t, 2pa6, 2pan, 2pcr, 2pd4, 2pd6, 2pfr, 2pid, 2pnf, 2poc, 2pq8, 2ps1, 2psd, 2psn, 2psq, 2ptr, 2puj, 2q28, 2q3m, 2q3o, 2q3r, 2q3z, 2q45, 2q4e, 2q4h, 2q4w, 2q74, 2q7w, 2q8n, 2qae, 2qcv, 2qep, 2qet, 2qgh, 2qj3, 2qjf, 2qjh, 2qkx, 2qn0, 2qr7, 2quy, 2qvb, 2qy0, 2r0i, 2r0l, 2r11, 2r2j, 2r37, 2r3a, 2r4f, 2r7b, 2r8o, 2r9f, 2ra3, 2rc4, 2rcc, 2rd5, 2rdu, 2reo, 2rhc, 2rhs, 2rkb, 2tod, 2toh, 2uuq, 2uwf, 2uxw, 2uzf, 2v09, 2v1p, 2v1z, 2v2h, 2v3a, 2v3z, 2v40, 2v5b, 2v5h, 2v5j, 2v5q, 2v65, 2v6a, 2v6b, 2v6c, 2v7o, 2v7q, 2v7y, 2v82, 2v8p, 2v8q, 2v9l, 2vag, 2van, 2vba, 2vcv, 2vcy, 2vd4, 2vd5, 2ve3, 2vef, 2vfa, 2vg0, 2vhd, 2vig, 2vjq, 2vk8, 2vp4, 2vp8, 2vre, 2vu1, 2vvm, 2vvt, 2vws, 2vx2, 2vxn, 2vxo, 2vz6, 2w0b, 2w1v, 2w20, 2w2d, 2w2j, 2w2n, 2w37, 2w4i, 2w4j, 2w4o, 2w5a, 2w5f, 2w6r, 2w8n, 2wbf, 2wbi, 2wdq, 2we5, 2wel, 2wf7, 2wfl, 2wfp, 2wgh, 2whz, 2wlr, 2wlt, 2wm5, 2wns, 2woj, 2wq8, 2wqd, 2wqm, 2wrt, 2wsk, 2wtb, 2wu8, 2wu9, 2wug, 2wvg, 2wxu, 2wy8, 2wz1, 2wzb, 2wzm, 2x02, 2x06, 2x0j, 2x55, 2x5o, 2x6t, 2x75, 2x7j, 2xap, 2xb6, 2xdw, 2xfg, 2xgz, 2xhl, 2xij, 2xik, 2xkr, 2xq0, 2xqr, 2xri, 2xrw, 2xsn, 2xsx, 2xtz, 2xu3, 2y3z, 2y5f, 2y65, 2y6o, 2y6t, 2y70, 2y7j, 2y88, 2yab, 2ybx, 2ycf, 2yfh, 2yfq, 2yr1, 2yv1, 2ywb, 2ywg, 2yxe, 2yxn, 2yxx, 2yyu, 2yyy, 2z04, 2z26, 2z3z, 2z61, 2z7f, 2zad, 2zat, 2zbw, 2zbx, 2zd8, 2zdh, 2zfi, 2zgc, 2zj3, 2zkj, 2zmd, 2znc, 2zpt, 2zpu, 2zq7, 2zsi, 2zu2, 2zvi, 2zwu, 3a04, 3a28, 3a2q, 3a32, 3a51, 3a74, 3a99, 3aal, 3aay, 3ab1, 3ado, 3afh, 3afi, 3ag6, 3aii, 3aj7, 3ajx, 3alo, 3ano, 3aqi, 3asa, 3atv, 3ay7, 3b12, 3b1u, 3b2t, 3b3d, 3b4x, 3b6r, 3b6u, 3b6v, 3ba1, 3bac, 3bfn, 3bg9, 3bgs, 3bh7, 3bhg, 3bhh, 3bhy, 3bkb, 3bmx, 3bn9, 3bo5, 3bon, 3bow, 3box, 3bpt, 3bsq, 3bul, 3buv, 3bwy, 3bxy, 3byd, 3c17, 3c2e, 3c7a, 3cbj, 3cc6, 3ce6, 3cg7, 3ckl, 3clh, 3cmc, 3cog, 3cov, 3cq0, 3csu, 3csw, 3ctl, 3cui, 3cw2, 3cx5, 3cx8, 3cyv, 3czh, 3d2f, 3d3l, 3d3w, 3d4p, 3d4u, 3d5t, 3d6e, 3d9a, 3d9d, 3dak, 3dbg, 3dc4, 3dcn, 3ddn, 3ddo, 3dfj, 3dg8, 3dgh, 3dgz, 3dj6, 3djl, 3dk9, 3dko, 3dl0, 3dmp, 3don, 3dqz, 3dt8, 3dtc, 3dva, 3dwb, 3dwg, 3dxb, 3dyd, 3e04, 3e0i, 3e0m, 3e2o, 3e2t, 3e3i, 3e4c, 3e77, 3e7g, 3e7o, 3e7w, 3e9a, 3e9k, 3ea3, 3ea4, 3ecr, 3ed3, 3ed7, 3edh, 3edi, 3eg4, 3ego, 3ehb, 3ejx, 3elf, 3emc, 3ep6, 3err, 3ess, 3etj, 3eua, 3ewb, 3ewm, 3exe, 3eya, 3eyg, 3eyx, 3f1l, 3f3s, 3f7j, 3f9i, 3f9m, 3fce, 3fci, 3fe1, 3fe3, 3fe4, 3ff1, 3fg1, 3fgc, 3fhr, 3fie, 3fjo, 3fju, 3fk4, 3flk, 3fme, 3fmu, 3fpc, 3fpl, 3fpq, 3fq8, 3fr7, 3fs2, 3fsl, 3fst, 3ftd, 3fuc, 3fve, 3fw3, 3fxz, 3fy4, 3fyd, 3g4d, 3g9x, 3gbj, 3gc2, 3gd5, 3gdf, 3gdo, 3gdq, 3ge3, 3gfb, 3ggf, 3giu, 3gl1, 3glc, 3glq, 3gmt, 3goa, 3goq, 3gos, 3gov, 3gp0, 3gp3, 3gqb, 3gr4, 3gr7, 3gru, 3gtd, 3gvi, 3gvp, 3gzd, 3gzy, 3h07, 3h0l, 3h0o, 3h0p, 3h11, 3h1d, 3h42, 3h49, 3h4j, 3h4s, 3h5q, 3h7r, 3h7u, 3h81, 3h8g, 3h9c, 3hbg, 3he2, 3hf1, 3hgr, 3hhp, 3hid, 3hja, 3hjb, 3hm8, 3hmp, 3hna, 3hnc, 3hng, 3ho9, 3hp0, 3hqn, 3hss, 3ht5, 3hul, 3hwc, 3hwn, 3hyh, 3hzn, 3hzu, 3i06, 3i28, 3i2k, 3i33, 3i3t, 3i6a, 3i6u, 3i99, 3ia2, 3iar, 3iau, 3ibd, 3icf, 3ids, 3idv, 3ief, 3ifs, 3ihj, 3ii0, 3iiq, 3ijp, 3ik0, 3ik7, 3im9, 3iml, 3inn, 3io3, 3iof, 3ip4, 3ipl, 3iqi, 3irm, 3ish, 3ist, 3itj, 3iu0, 3iu1, 3ivy, 3ix6, 3jrm, 3jru, 3jsl, 3jtm, 3ju4, 3ju5, 3ju8, 3jvr, 3jxf, 3jxg, 3jyo, 3jz6, 3jze, 3k2u, 3k35, 3k3p, 3k3q, 3k8z, 3k92, 3k96, 3k9v, 3ka0, 3kb9, 3kbb, 3kcg, 3kd9, 3kdn, 3kfa, 3kgb, 3kk8, 3kki, 3kld, 3km3, 3kmu, 3kn6, 3kp1, 3kr6, 3ks3, 3ksk, 3kt2, 3kul, 3kvw, 3kxq, 3ky7, 3kyh, 3kzn, 3l07, 3l24, 3l4e, 3l5k, 3l60, 3l6b, 3l6c, 3l86, 3l8c, 3l9c, 3l9w, 3lac, 3lb8, 3lbf, 3ldo, 3ldv, 3lfu, 3ljk, 3ljq, 3lk7, 3lm5, 3lms, 3ln3, 3lpf, 3lq0, 3lq1, 3lqs, 3lre, 3ls6, 3lvm, 3lwb, 3lxm, 3lzw, 3m00, 3m2w, 3m4y, 3m58, 3m5u, 3m6i, 3m83, 3m9y, 3mbd, 3mca, 3mdm, 3mdy, 3mfr, 3mhp, 3mhs, 3mi0, 3mio, 3ml5, 3mog, 3mpi, 3mt6, 3mtg, 3mtl, 3mue, 3mvi, 3mwd, 3mwq, 3mx6, 3mxt, 3mz0, 3n0g, 3n2b, 3n2l, 3n37, 3n3r, 3n58, 3n5n, 3n75, 3n80, 3n8h, 3nas, 3nc3, 3ncl, 3ndc, 3ndp, 3nea, 3nfy, 3ng0, 3ngx, 3nps, 3nr9, 3nvs, 3ny4, 3nyd, 3nyt, 3nz1, 3o1n, 3o23, 3o47, 3o4p, 3o4r, 3o5s, 3o63, 3o76, 3o8j, 3occ, 3od5, 3oet, 3of5, 3ofm, 3ohs, 3oid, 3oig, 3ojf, 3okf, 3olj, 3olm, 3oml, 3ond, 3ooi, 3orf, 3osu, 3otm, 3ovp, 3oz7, 3ozb, 3ozc, 3ozf, 3ozu, 3p14, 3p1a, 3p26, 3p2l, 3p2o, 3p5p, 3p86, 3p95, 3pao, 3pbh, 3pd6, 3pdk, 3pdx, 3pg8, 3pgx, 3phh, 3piu, 3pkb, 3pki, 3pm0, 3pm6, 3pmo, 3pp8, 3ppm, 3pqa, 3pqe, 3prh, 3px8, 3pyf, 3pzr, 3q12, 3q1k, 3q37, 3q3v, 3q6o, 3q7e, 3q98, 3q9o, 3qan, 3qbe, 3qd2, 3qdl, 3qe2, 3qe3, 3qfa, 3qft, 3qfu, 3qgy, 3qh2, 3qhw, 3qkz, 3qm2,

3qm3, 3qml, 3qn3, 3qne, 3qrv, 3qtg, 3qtp, 3qwd, 3qxh, 3qy9, 3qz1, 3r0q, 3r0s, 3r12, 3r1g, 3r31, 3r38, 3r44, 3r5x, 3r77, 3r7f, 3r7t, 3r8y, 3r9p, 3rde, 3rfw, 3rg9, 3rhj, 3rii, 3rim, 3riy, 3rj5, 3rjl, 3rku, 3rlg, 3rlh, 3rm5, 3rmj, 3rmt, 3ro0, 3ro8, 3rpt, 3rr1, 3rrw, 3rtk, 3ru6, 3rv2, 3rwb, 3rxw, 3s1x, 3s3e, 3s46, 3s5b, 3s6d, 3s82, 3s97, 3s9v, 3s9z, 3scz, 3sds, 3sg1, 3sgv, 3sgz, 3sk0, 3sk3, 3slh, 3sls, 3smq, 3smt, 3so3, 3sr7, 3swe, 3swg, 3sxx, 3sz3, 3sza, 3t4e, 3t57, 3t5a, 3t5c, 3t7b, 3t7v, 3t8b, 3t94, 3ta6, 3tac, 3tbf, 3tc6, 3tcm, 3tde, 3teg, 3tfj, 3tfx, 3tg0, 3tg2, 3thi, 3tjz, 3tk9, 3tl2, 3tlo, 3tm0, 3tnl, 3to7, 3tpf, 3tqi, 3tqp, 3tqs, 3tqt, 3tr2, 3tr7, 3tsm, 3tsr, 3ttv, 3tug, 3tut, 3tvj, 3twk, 3twl, 3tx1, 3tx8, 3ty4, 3ty7, 3tz6, 3tzl, 3u06, 3u0v, 3u28, 3u31, 3u3o, 3u49, 3u4w, 3u62, 3u81, 3u9w, 3uaw, 3ubd, 3uc3, 3uc4, 3udb, 3udu, 3ue9, 3uf7, 3ufm, 3ufx, 3ug7, 3ugs, 3uh0, 3uhf, 3uim, 3uiu, 3uk2, 3uko, 3ulk, 3umo, 3unx, 3uq8, 3utn, 3uto, 3uvj, 3uwd, 3uwl, 3uxo, 3uyq, 3uzu, 3v0s, 3v1y, 3v2h, 3v2i, 3v8d, 3v8e, 3v98, 3vbb, 3vcy, 3vku, 3vl9, 3vmf, 3vmj, 3vn9, 3voc, 3vpg, 3vv2, 3w06, 3zcw, 3zfd, 3zia, 3zoq, 3zq6, 3zr4, 3ztv, 3zv4, 3zxw, 3zzh, 3zzm, 4a14, 4a1n, 4a1x, 4a29, 4a3q, 4a3r, 4a3s, 4a5o, 4a8t, 4aaj, 4acf, 4adm, 4aec, 4af3, 4ag1, 4agu, 4aj9, 4ajj, 4amu, 4amv, 4as4, 4as5, 4asi, 4aw2, 4b9d, 4bbn, 4d8y, 4d9b, 4dd5, 4dfl, 4dg5, 4dgj, 4dmy, 4dng, 4dnx, 4doi, 4dq1, 4dq8, 4dql, 4dr0, 4dv8, 4dyg, 4e01, 4e0b, 4e37, 4e3x, 4e5n, 4e5y, 4e79, 4ead, 4eam, 4ear, 4ec0, 4eca, 4edf, 4eg2, 4egj, 4egq, 4eh1, 4ehi, 4ej0, 4ekf, 4ekn, 4ekz, 4em6, 4emb, 4eo9, 4epi, 4epj, 4eqs, 4eqy, 4esm, 4etp, 4ex4, 4ex5, 4exq, 4eyw, 4ez5, 4ez8, 4f0i, 4f0z, 4f3y, 4f40, 4f9c, 4fb7, 4fc7, 4fce, 4fek, 4fey, 4fg8, 4fgl, 4fgw, 4fk7, 4fkz, 4fmu, 4fsh, 4fuh, 4fvy, 4fxq, 4fzw, 4g38, 4g4s, 4g6z, 4g7l, 4g9b, 4ga6, 4gac, 4gam, 4gbu, 4gcj, 4gcm, 4gdp, 4ge6, 4gev, 4gic, 4gie, 4gj1, 4glw, 4goj, 4gpz, 4gqs, 4gri, 4grz, 4gut, 4gwg, 4gyp, 4h27, 4h2g, 4h31, 4h3d, 4h3z, 4h4f, 4h8a, 4h8e, 4hb7, 4hgh, 4hgv, 4hkt, 4hl2, 4hl6, 4hno, 4ht2, 4htg, 4htr, 4hv4, 4hva, 4hvc, 4hvs, 4i42, 4i90, 4iao, 4ic7, 4ief, 4ig6, 4igd, 4ij5, 4ijn, 4ijz, 4ikp, 4imy, 4ip7, 4iq8, 4is2, 4iu6, 4iv9, 4j15, 4j1y, 4j9t, 4jco, 4jnk, 4jr2, 4jrm, 4jxj, 4k7z, 4k9d, 5mdh, 6gsv, 7odc, 8cgt

Table S3. Matched scaffolds and the number of matches in each scaffold.

Substrate	Active site residues [a]	Selected scaffolds (Number of matches) [b]
Cephradine	SER1, HIS2, ASP3, ALA4, TYR5, SER6, GLU7	1a5z(5), 1a7t(7), 1amu(1), 1aqj(1), 1auo(24), 1axk(5), 1b1y(2), 1bqy(11), 1bry(4), 1bx4(3), 1e1h(3), 1e8g(4), 1ep3(3), 1eud(2), 1frb(8), 1fye(2), 1gsa(6), 1gv0(3), 1hye(26), 1iyk(1), 1j6u(1), 1j93(6), 1jfl(14), 1ji1(4), 1jkm(2), 1jqi(1), 1jvb(53), 1kf6(11), 1m6j(2), 1mas(3), 1mdb(14), 1mla(242), 1mpx(21), 1nj1(2), 1o7x(2), 1odt(2), 1on3(1), 1owl(5), 1pa2(3), 1pj9(2), 1qj5(4), 1shz(3), 1sov(4), 1te2(3), 1tel(1), 1uek(15), 1uf5(1), 1v25(2), 1v7z(4), 1wdk(8), 1wos(2), 1x0l(9), 1yb1(1), 2a9d(10), 2bgs(1), 2dpl(9), 2fpq(7), 2gtr(2), 2hg2(2), 2hjh(2), 2i87(16), 2j07(14), 2jar(1), 2jif(10), 2o4c(3), 2qae(1), 2quy(9), 2uxw(10), 2v3z(4), 2vig(2), 2wfl(1), 2wfp(19), 2wtb(2), 2x0j(2), 2x5o(1), 2xdw(15), 2xhl(1), 2xri(18), 2xrw(5), 2y7j(7), 2yxe(18), 2z3z(37), 3bac(1), 3c2e(2), 3csu(2), 3d2f(1), 3e7g(1), 3ep6(3), 3eyg(1), 3f9m(3), 3fxz(1), 3glc(1), 3hgr(1), 3hja(2), 3ho9(6), 3hss(1), 3i2k(3), 3iar(8), 3ik0(1), 3im9(9), 3k3q(1), 3kgb(1), 3kk8(1), 3kmu(6), 3l9c(2), 3m83(8), 3mio(6), 3ml5(2), 3mvi(155), 3n37(2), 3n3r(4), 3n80(2), 3pao(3), 3qkz(2), 3qm3(1), 3r38(1), 3rrw(1), 3t7v(3), 3to7(4), 3tqp(4), 3tx1(7), 3ue9(8), 3uko(6), 3uwl(6), 3v0s(10), 3vpg(2), 4asi(2), 4edf(36), 4egj(1), 4ex5(6), 4f9c(10), 4fkz(2), 4h3z(9), 4hgh(1)

[a]: The residues are corresponding to residues in the active site model (Fig. 2A).

[b]: A total of 1,160 matches were identified in 134 unique scaffolds.

Table S4. Scheme for designing cephradine synthase on scaffold 3i2k.

Scaffold	Catalytic residues	Binding residues		Substrate
		Sequence selection residues	Conformation optimization residues	
3i2k	Tyr44, Ser117, Tyr118, His287, Asp259	Gln55, His82, Val121, Leu146, Ala149, Pro150, Trp151, Ala162, Leu169, Phe261, Leu407, Phe408	Asn42, Trp52, Thr54, Ser56, Val116, Leu119, Ser140, Met141, Ser143, Trp166, Ile170, Trp220, Trp235, Glu264, Ser265, Trp285, Ser286, Ser288, Leu290	Cephradine

Table S5. Catalytic geometrical constraints for design.

Site pair	Type	Atom1	Atom2[a]	Atom3[a]	Atom4[a]	Min[b]	Max[b]
Ser117-TS	Distance	OG	#C15			1.4	1.6
	Angle	CB	OG	#C15		100.0	120.0
	Angle	OG	#C15	#O16		100.0	120.0
	Torsion	OG	#O16	#C15	#N14	100.0	140.0
Tyr118- TS	Distance	N	#O16			2.6	3.0
	Angle	N	HN	#O16		140.0	180.0
	Angle	HN	#O16	#C15		80.0	140.0
Tyr44- TS	Distance	OH	#O16			2.6	3.0
	Angle	OH	HH	#O16		140.0	180.0
	Angle	HH	#O16	#C15		80.0	140.0
His287-Ser117	Distance	NE2	#OG			2.6	3.0
	Angle	NE2	#HG1	#OG		140.0	180.0
	Angle	CE1	NE2	#HG1		90.0	150.0
Asp259-His287 [c]	Distance	OD1	#ND1			2.6	3.0
	Angle	CG	OD1	#HD1		90.0	150.0
	Angle	OD1	#HD1	#ND1		140.0	180.0
Asp259-His287 [c]	Distance	OD2	#ND1			2.6	3.0
	Angle	CG	OD2	#HD2		90.0	150.0
	Angle	OD2	#HD2	#ND1		140.0	180.0
Phe261-Asp259 [d]	Distance	N	#OD2			2.6	3.2
	Angle	HN	#OD2	#CG		90.0	150.0
	Angle	N	HN	#OD2		140.0	180.0
Phe261-Asp259 [d]	Distance	N	#OD1			2.6	3.2
	Angle	HN	#OD1	#CG		90.0	150.0
	Angle	N	HN	#OD1		140.0	180.0

[a] Atoms on the latter residue within a site pair are prefixed with '#'.

[b] Distance measurements are given in Å. Angle and torsion measurements are given in degrees.

[c] Either site pair of Asp259-His287 is needed.

[d] Either site pair of Phe261-Asp259 is needed.

Table S6. Variation rules for generating TS library.

Type	Atom1[a]	Atom2[a]	Atom3[a]	Atom4	Min[b]	Max[b]	Step[b]
Distance	#OG	C15			1.4	1.6	0.1
Angle	#CB	#OG	C15		105.0	115.0	5.0
Torsion	#CA	#CB	#OG	C15	-180.0	179.0	10.0
Angle	#OG	C15	O16		105.0	115.0	5.0
Torsion	#OG	O16	C15	N14	120.0	120.0	0.0
Torsion	#OG	O16	C15	C17	-120.0	-120.0	0.0
Torsion	#CB	#OG	C15	O16	-180.0	179.0	10.0
Torsion	N14	C15	C17	C18	-180.0	179.0	10.0
Torsion	C15	C17	C18	C19	-180.0	179.0	10.0
Torsion	O16	C15	N14	C7	-180.0	179.0	10.0
Torsion	C15	N14	C7	C8	-180.0	179.0	10.0

[a] Atoms on anchor residue Ser117 are prefixed with '#'.

[b] Distance measurements are given in Å. Angle and torsion measurements are given in degrees.

Table S7. Calculated free energy changes for single mutants compared to wild type. The $\Delta\Delta G_{Bind}$ and $\Delta\Delta G_{Fold}$ are in PRODA energy units.

Variants	Mutations	$\Delta\Delta G_{Bind}$ [a]	$\Delta\Delta G_{Fold}$ [b]	Variants	Mutations	$\Delta\Delta G_{Bind}$ [a]	$\Delta\Delta G_{Fold}$ [b]
1	P150L	2.34	-11.54	65	L407T	-0.96	7.64
2	P150M	3.50	-11.70	66	L146T	0.88	5.85
3	P150C	1.50	-8.49	67	Q55S	-1.95	9.09
4	Q55R	-3.00	-3.03	68	A162G	-0.07	7.25
5	P150V	1.20	-6.24	69	F408I	1.00	6.30
6	P150T	0.89	-5.55	70	F408V	0.63	6.71
7	P150A	0.21	-4.09	71	Q55C	0.11	7.32
8	F261T	-6.08	2.81	72	L146V	0.77	6.78
9	F261L	-6.75	5.59	73	F261D	-5.69	13.43
10	P150S	1.15	-2.18	74	L407S	-0.46	8.23
11	L169R	6.00	-6.92	75	F408W	3.60	4.24
12	A162S	0.07	-0.68	76	L146S	1.06	6.82
13	L169H	-2.19	1.76	77	Q55A	0.02	7.92
14	F261N	-4.28	3.89	78	F408M	0.73	7.31
15	F261S	-5.80	5.53	79	F408C	1.03	7.33
16	L169V	-2.46	2.36	80	F408A	0.82	7.58
17	L169I	-1.41	1.75	81	L169D	-2.74	11.15
18	F261H	-0.58	1.01	82	F261E	-6.06	14.49
19	L407I	0.12	0.59	83	L169M	-1.19	9.63
20	A162T	0.78	0.46	84	Q55V	-0.47	9.05
21	F261A	-5.96	7.45	85	F408L	1.20	7.45
22	A162M	0.55	0.75	86	F261G	-5.90	14.56
23	A162V	0.55	0.99	87	F408N	0.84	7.84
24	F261Q	-4.05	5.68	88	Q55T	-2.13	10.89
25	F261C	-6.02	7.66	89	Q55M	-1.46	10.40
26	L407V	0.64	1.06	90	L146I	0.72	8.46
27	L146F	5.32	-3.62	91	F261V	-6.22	15.41
28	L146W	-4.58	6.38	92	L169Q	4.67	4.55
29	H87F	-8.08	10.11	93	F261Y	6.83	2.51
30	A162C	0.54	1.50	94	L169E	-0.78	10.14
31	L169Y	-7.94	10.07	95	L407D	0.46	10.20
32	P150Q	6.55	-4.28	96	F408S	0.94	9.75
33	V121T	-2.12	4.44	97	H87M	0.77	10.27
34	P150N	1.72	0.85	98	Q55I	-1.18	12.24
35	L169C	-3.64	6.23	99	W151F	-4.74	16.03
36	P150H	5.14	-2.45	100	L407G	-1.54	13.34
37	F261M	-7.21	10.02	101	Q55E	1.74	10.33
38	P150I	10.81	-7.92	102	L146G	1.10	11.13

39	V121C	-1.26	4.35	103	L169G	-6.14	18.40
40	P150G	0.39	2.96	104	F408T	1.08	11.26
41	L407H	5.23	-1.84	105	V121G	-1.72	14.48
42	V121A	-2.10	5.91	106	H87Q	0.79	12.57
43	P150D	1.91	1.99	107	H87Y	-0.04	13.43
44	L169A	-4.47	8.42	108	H87V	0.56	13.04
45	Q55L	-1.16	5.14	109	Q55D	0.78	13.22
46	P150E	2.64	1.48	110	F261I	-5.66	19.67
47	V121S	-1.80	6.28	111	A162L	-0.92	15.08
48	L407M	-0.01	4.56	112	F408D	1.45	12.71
49	L407A	-1.24	5.87	113	V121F	-4.46	18.73
50	L146A	1.07	3.60	114	H87L	0.16	14.66
51	L407C	-1.57	6.25	115	H87I	0.32	14.54
52	Q55H	-2.14	6.92	116	P150F	-2.11	17.04
53	L169T	-2.39	7.33	117	L407E	2.63	12.57
54	L407Q	4.42	0.57	118	Q55G	0.14	15.26
55	L169N	-3.12	8.14	119	F408G	0.84	14.78
56	L407N	-2.35	7.44	120	H87C	0.81	15.24
57	L146C	0.89	4.21	121	F408E	1.97	14.51
58	F408H	2.00	3.37	122	F261W	4.72	12.45
59	L169S	-3.24	8.81	123	H87A	0.97	16.32
60	A162I	0.58	5.02	124	Q55N	-1.79	19.86
61	V121I	7.29	-1.51	125	H87N	1.01	17.45
62	F408Y	3.33	2.77	126	A162F	0.43	18.16
63	F408Q	0.87	5.35	127	H87T	0.70	17.94
64	Q55Y	-2.92	9.19				

[a]: $\Delta\Delta G_{Bind}$ is difference of the binding energy of mutant minus that of wild type.

[b]: $\Delta\Delta G_{Fold}$ is difference of the folding energy of mutant minus that of wild type.

Table S8. Calculated free energy changes for double and triple mutants based on F261T compared to wild type. The $\Delta\Delta G_{Bind}$ and $\Delta\Delta G_{Fold}$ are in PRODA energy units.

Variants	Mutations	$\Delta\Delta G_{Bind}$ [a]	$\Delta\Delta G_{Fold}$ [b]	Variants	Mutations	$\Delta\Delta G_{Bind}$ [a]	$\Delta\Delta G_{Fold}$ [b]
1	F261T/Q55R	-0.20	1.00	69	F261T/F408C	4.95	10.65
2	F261T/F408Y	-4.25	6.70	70	F261T/F408W	-3.37	19.06
3	F261T/L146F	3.58	-0.93	71	F261T/F408L	5.07	10.73
4	F261T/P150C	9.26	-5.46	72	F261T/F408N	4.70	11.16
5	F261T/L169R	3.77	1.08	73	F261T/L169D	1.22	14.71
6	F261T/P150A	4.46	0.71	74	F261T/F408M	5.10	11.05
7	F261T/P150W	3.81	1.38	75	F261T/L407M	10.82	5.74
8	F261T/P150V	13.54	-7.75	76	F261T/A162I	4.36	12.29
9	F261T/V121I	4.24	1.81	77	F261T/Q55M	2.37	14.33
10	F261T/V121M	3.17	3.31	78	F261T/Q55V	3.17	13.52
11	F261T/P150Q	2.51	4.15	79	F261T/L169M	7.87	8.90
12	F261T/A162S	3.94	2.87	80	F261T/H87M	3.68	13.11
13	F261T/L407I	4.31	2.54	81	F261T/L146I	3.95	12.88
14	F261T/P150T	11.53	-4.65	82	F261T/Q55W	-10.63	27.77
15	F261T/H87F	3.60	3.70	83	F261T/P150Y	4.22	12.93
16	F261T/P150E	1.20	6.89	84	F261T/P150L	25.80	-8.48
17	F261T/L146W	3.38	4.77	85	F261T/L407D	6.20	11.30
18	F261T/L407V	4.68	3.48	86	F261T/P150I	30.83	-13.16
19	F261T/L407Q	4.88	3.40	87	F261T/F408S	4.57	13.16
20	F261T/L169H	1.98	6.53	88	F261T/L169E	3.41	14.96
21	F261T/H87W	3.30	5.51	89	F261T/L407G	4.89	13.65
22	F261T/P150S	8.66	0.40	90	F261T/F408T	4.78	14.53
23	F261T/P150G	4.96	4.29	91	F261T/L146G	3.96	15.52
24	F261T/L169V	1.23	8.17	92	F261T/Q55E	5.29	14.32
25	F261T/L169I	-0.25	9.81	93	F261T/L169G	-1.68	21.81
26	F261T/L407H	8.60	1.12	94	F261T/W151F	3.00	17.33
27	F261T/P150H	4.87	5.58	95	F261T/H87Q	3.94	16.77
28	F261T/L169C	1.33	9.21	96	F261T/L407E	4.59	16.16
29	F261T/A162T	3.98	6.71	97	F261T/Q55I	2.83	18.37
30	F261T/L169Y	-4.23	15.36	98	F261T/H87V	4.03	17.17
31	F261T/L407A	3.80	7.43	99	F261T/P150F	2.75	18.51
32	F261T/L407C	5.80	5.49	100	F261T/Q55D	3.29	17.97
33	F261T/L169A	-0.14	11.81	101	F261T/F408D	5.32	16.12
34	F261T/L146A	3.96	7.99	102	F261T/V121G	4.34	17.17
35	F261T/V121T	4.38	7.69	103	F261T/P150M	22.88	-1.06
36	F261T/V121C	4.14	8.08	104	F261T/Q55G	2.44	19.41

37	F261T/A162M	4.37	8.06	105	F261T/L169W	15.11	6.78
38	F261T/A162C	3.86	8.62	106	F261T/H87I	3.98	18.53
39	F261T/L146C	3.94	8.60	107	F261T/H87L	4.03	18.63
40	F261T/A162V	4.39	8.22	108	F261T/F408G	4.48	18.23
41	F261T/F408H	6.02	6.60	109	F261T/H87C	4.03	19.29
42	F261T/Q55S	-1.22	14.11	110	F261T/Q55N	-1.46	24.88
43	F261T/L407T	3.90	9.28	111	F261T/F408E	5.96	17.86
44	F261T/L407N	7.25	5.98	112	F261T/H87A	4.03	20.38
45	F261T/V121A	4.12	9.11	113	F261T/H87N	4.12	21.46
46	F261T/F408Q	4.60	8.70	114	F261T/V121F	2.78	22.90
47	F261T/L169N	2.31	11.12	115	F261T/H87T	4.10	22.08
48	F261T/H87Y	4.69	8.80	116	F261T/H87S	4.06	22.79
49	F261T/P150N	9.97	3.58	117	F261T/L146M	3.97	24.59
50	F261T/Q55C	2.21	11.39	118	F261T/V121L	3.87	24.70
51	F261T/L169S	1.53	12.13	119	F261T/W151T	-1.21	30.41
52	F261T/V121S	4.37	9.64	120	F261T/W151Y	3.70	25.55
53	F261T/Q55Y	-1.01	15.02	121	F261T/V121Y	3.78	25.84
54	F261T/Q55H	1.97	12.11	122	F261T/H87G	4.02	28.00
55	F261T/Q55A	2.27	12.02	123	F261T/W151M	3.32	29.18
56	F261T/L146T	4.03	10.26	124	F261T/W151C	-1.29	35.17
57	F261T/P150D	10.44	3.93	125	F261T/V121W	4.14	30.30
58	F261T/L407S	4.55	9.87	126	F261T/W151S	-1.37	35.92
59	F261T/A162G	3.83	10.62	127	F261T/L407F	17.27	17.57
60	F261T/F408V	4.77	9.83	128	F261T/W151L	-1.29	36.80
61	F261T/Q55T	-1.30	15.92	129	F261T/W151A	-1.41	37.49
62	F261T/F408I	5.03	9.59	130	F261T/W151V	-1.29	38.95
63	F261T/L169T	1.73	12.89	131	F261T/A162L	4.13	39.46
64	F261T/L169Q	5.19	9.44	132	F261T/W151I	-1.30	45.35
65	F261T/Q55L	-1.55	16.20	133	F261T/L407W	27.06	17.74
66	F261T/L146V	3.96	11.18	134	F261T/L146Y	3.77	42.05
67	F261T/L146S	4.04	11.20	135	F261T/L407Q/F408Y	0.54	8.06
68	F261T/F408A	4.46	11.03	136	F261T/L407H/F408Y	7.50	11.77

[a]: $\Delta\Delta G_{Bind}$ is difference of the binding energy of mutant minus that of wild type.

[b]: $\Delta\Delta G_{Fold}$ is difference of the folding energy of mutant minus that of wild type.

Table S9. Steady-state kinetic parameters of the wild-type cocaine esterase and designed mutants for hydrolysis of cephradine. Reaction conditions are: pH 7.0 at 22 °C.

Variants	Mutations	k_{cat} ($\times 10^{-3}$ s^{-1})	K_m (mM)	k_{cat}/K_m ($\times 10^{-3}$ $M^{-1}s^{-1}$)
WT	-	0.61	33.90	18.11
M1	F261T	2.33	12.10	192.72
M2	Q55R	0.02	25.16	0.61
M3	P150E	2.11	151.34	13.93
M4	F408L	0.65	70.03	9.26
M5	F261T/P150E	0.44	51.16	8.63
M6	F261T/L407Q	0.04	7.46	5.61
M7	F261T/L407H	0.32	77.98	4.16
M8	F261T/F408Y	0.19	19.47	9.84
M9	F261T/L407Q/F408Y	0.26	23.68	10.99
M10	F261T/L407H/F408Y	0.07	5.02	13.41

2. Supplementary Figures

Fig. S1 Computational enzyme design framework in PRODA.

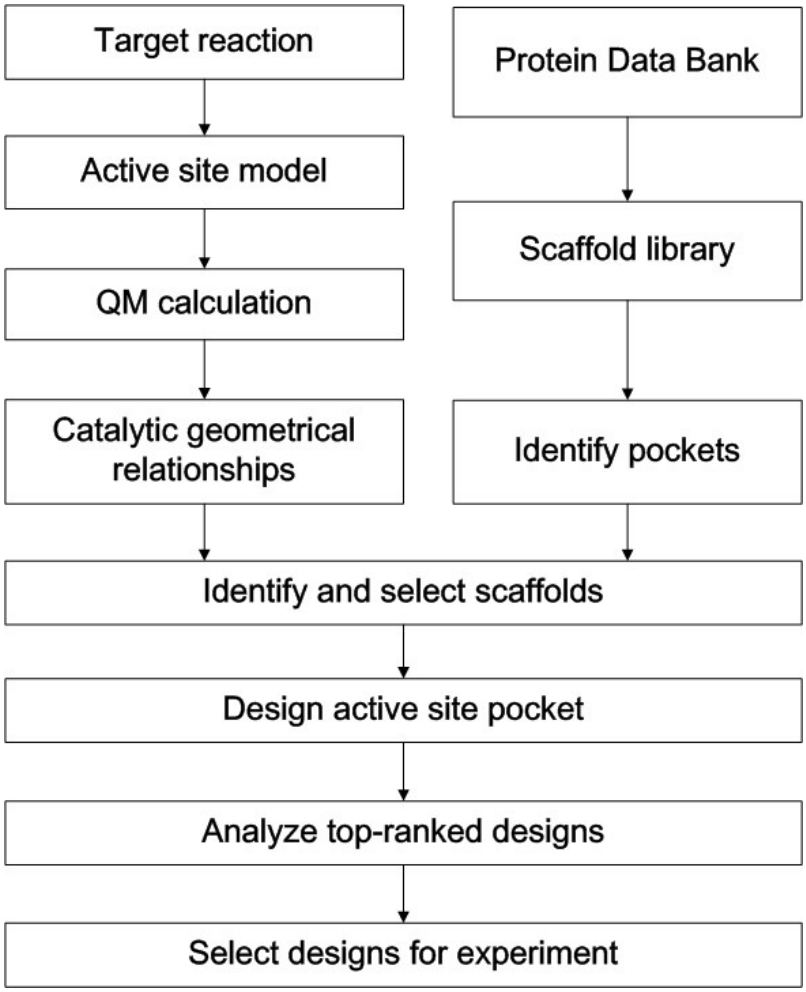


Fig. S2 Chemical routines of cephradine synthesis.

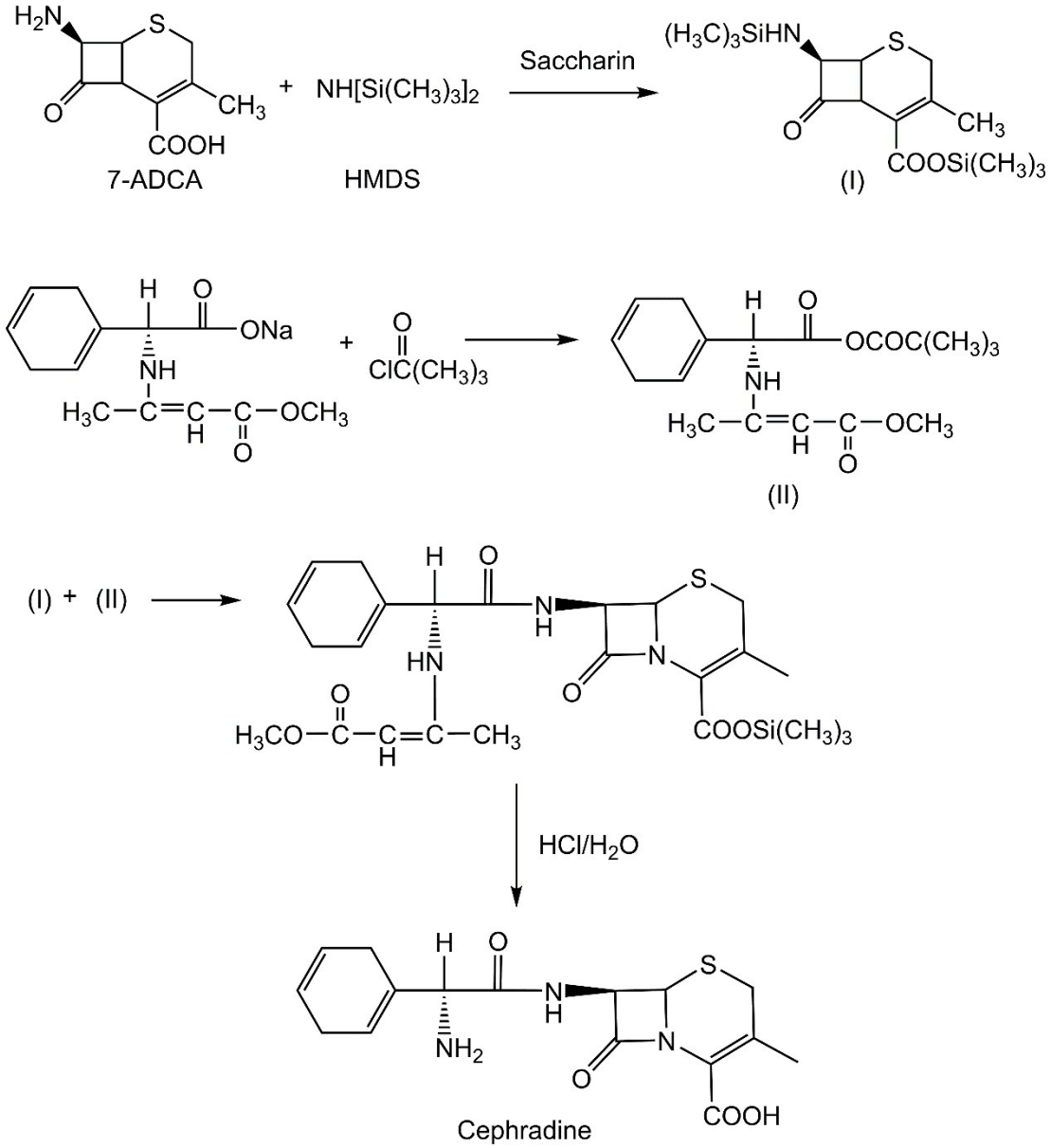


Fig. S3 Structural formula of cephradine marked with heavy atom names.

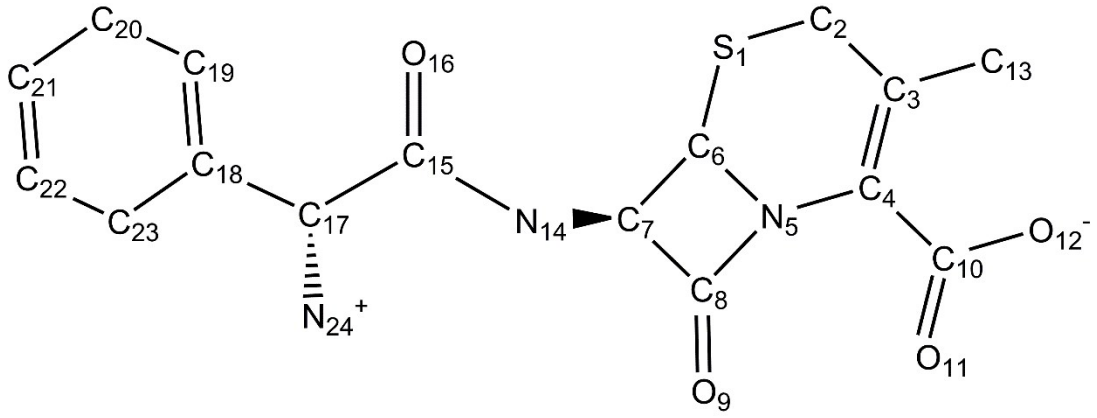


Fig. S4 Sequence alignment between *Rh.* cocE (574 AAs) and *E.coli* PGA (766 AAs). A red box indicates catalytic residues of *Rh.* cocE and a green box the catalytic residues of *E.coli* PGA.

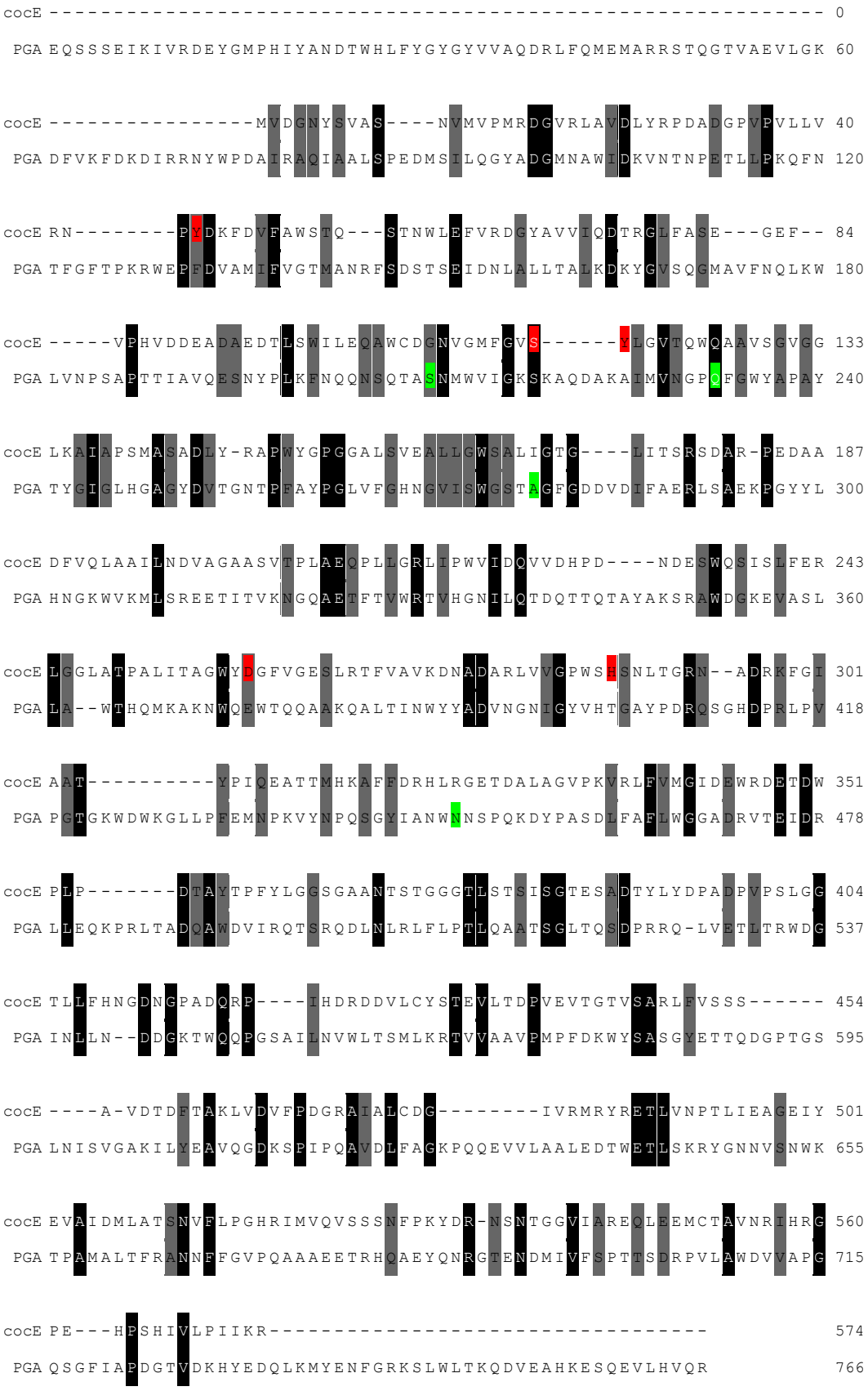


Fig. S5 Sequence alignment between *Rh. cocE* (574 AAs) and *X.citri* AEH (637 AAs). A red box indicates catalytic residues of *Rh.* cocE and a green box the catalytic residues of *X.citri* AEH, and a cyan box the residues of *X.citri* AEH involved in binding the α-moiety.

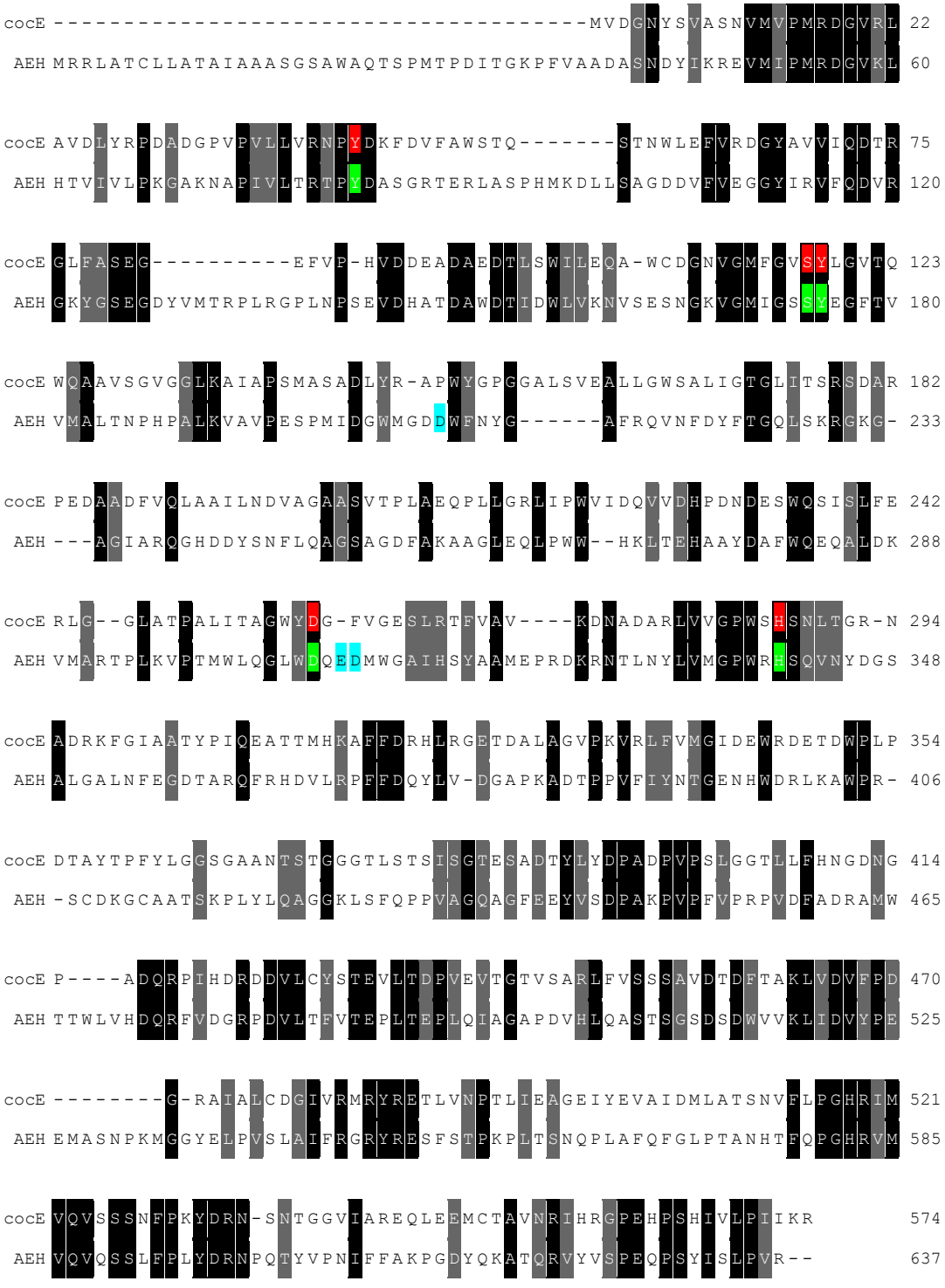
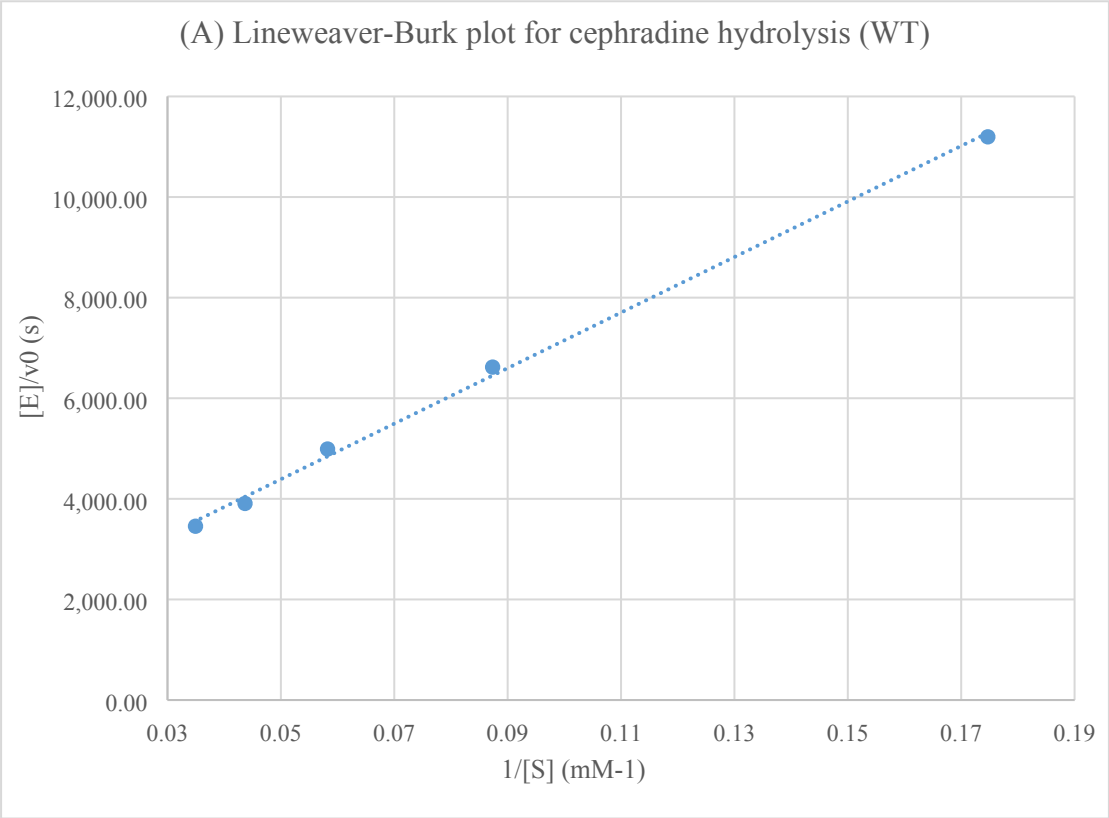
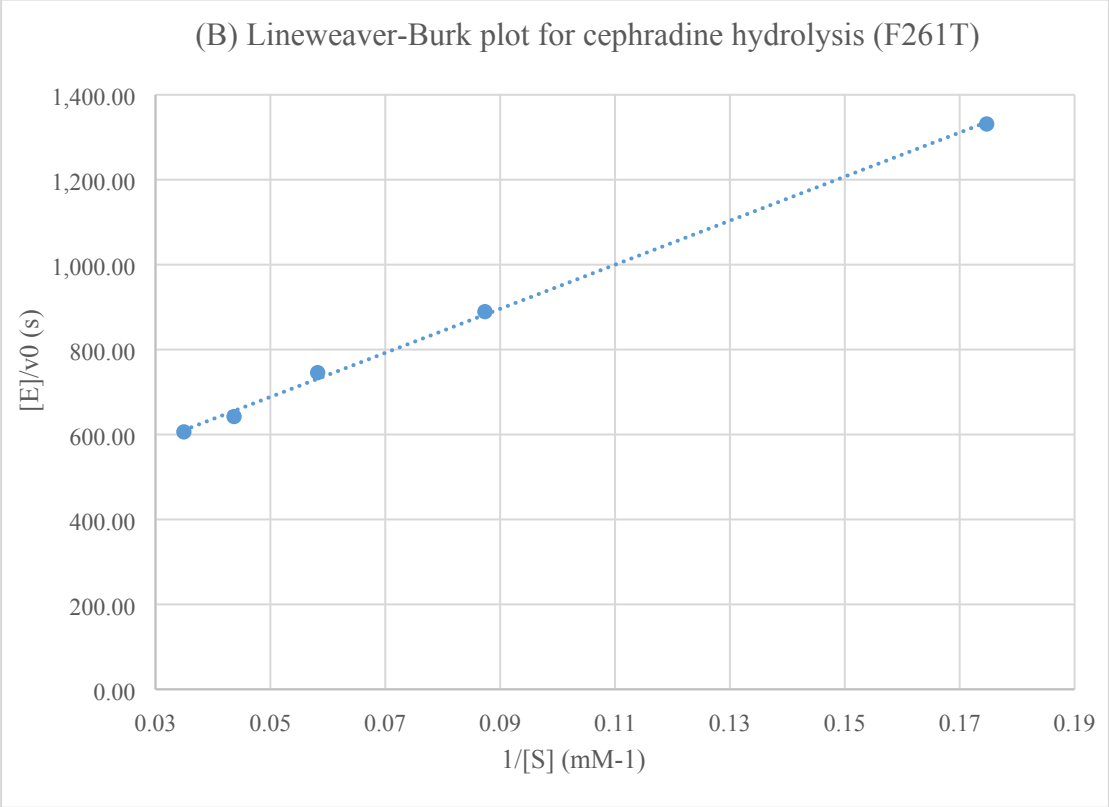
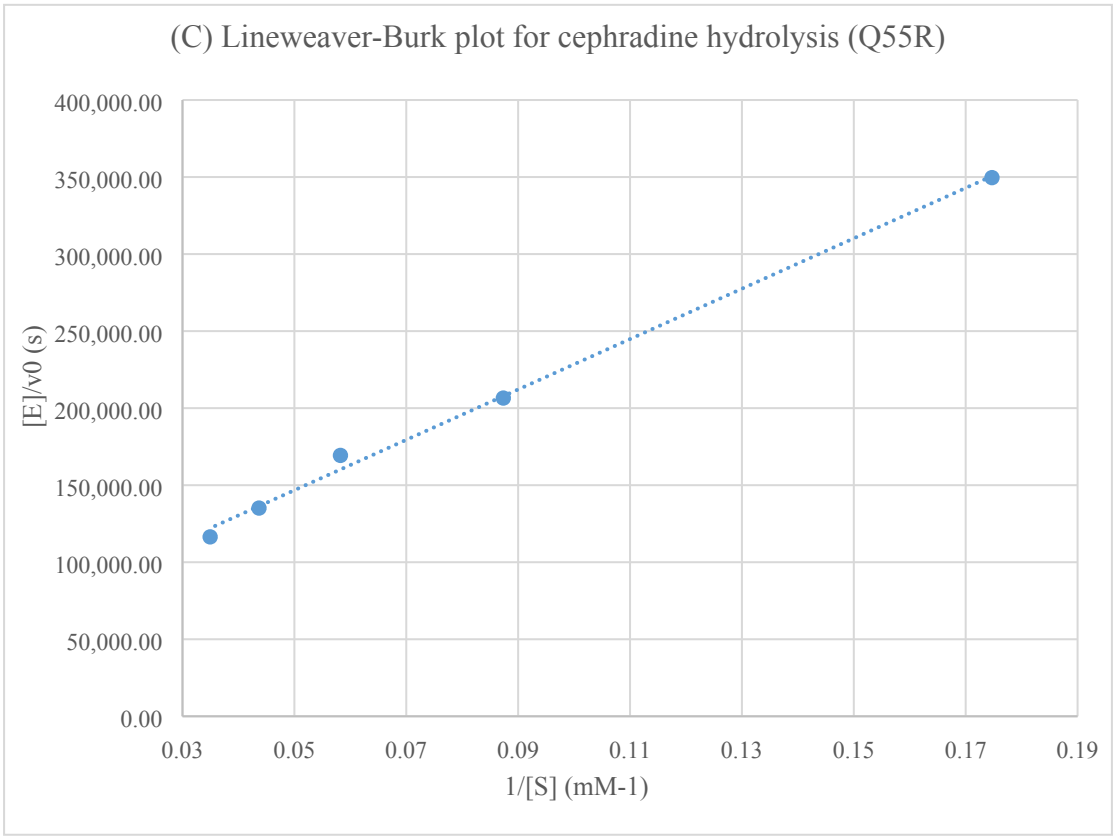
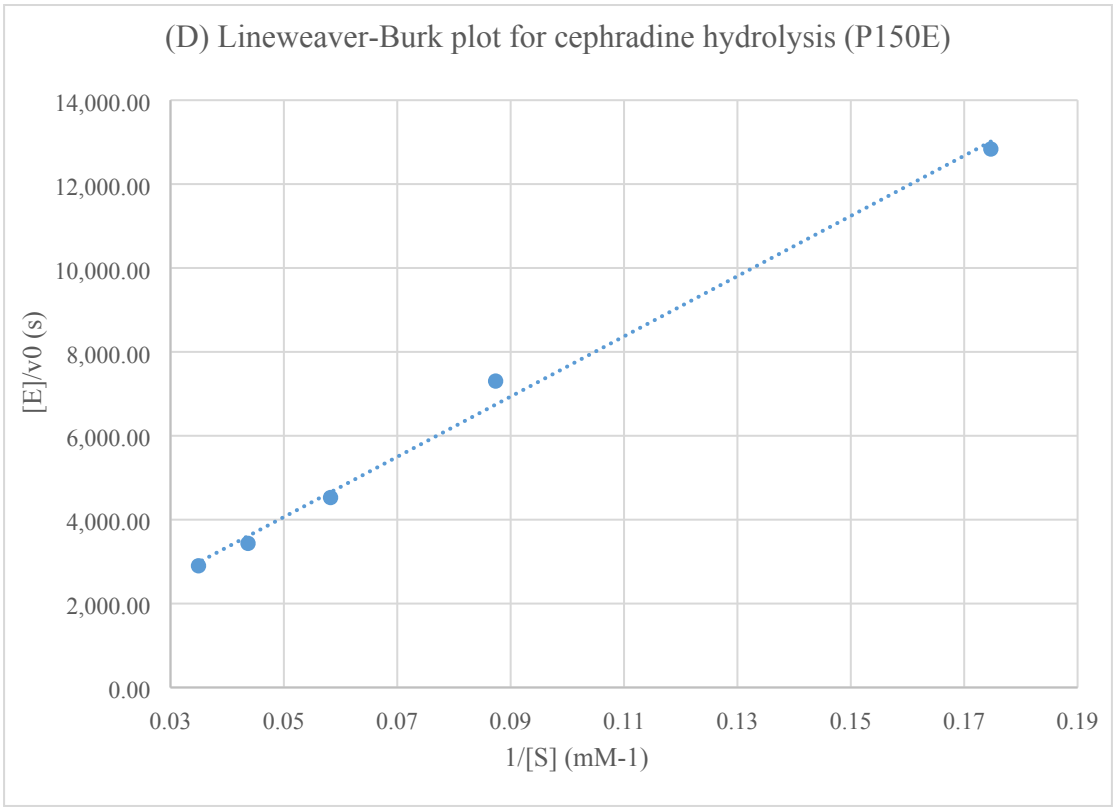


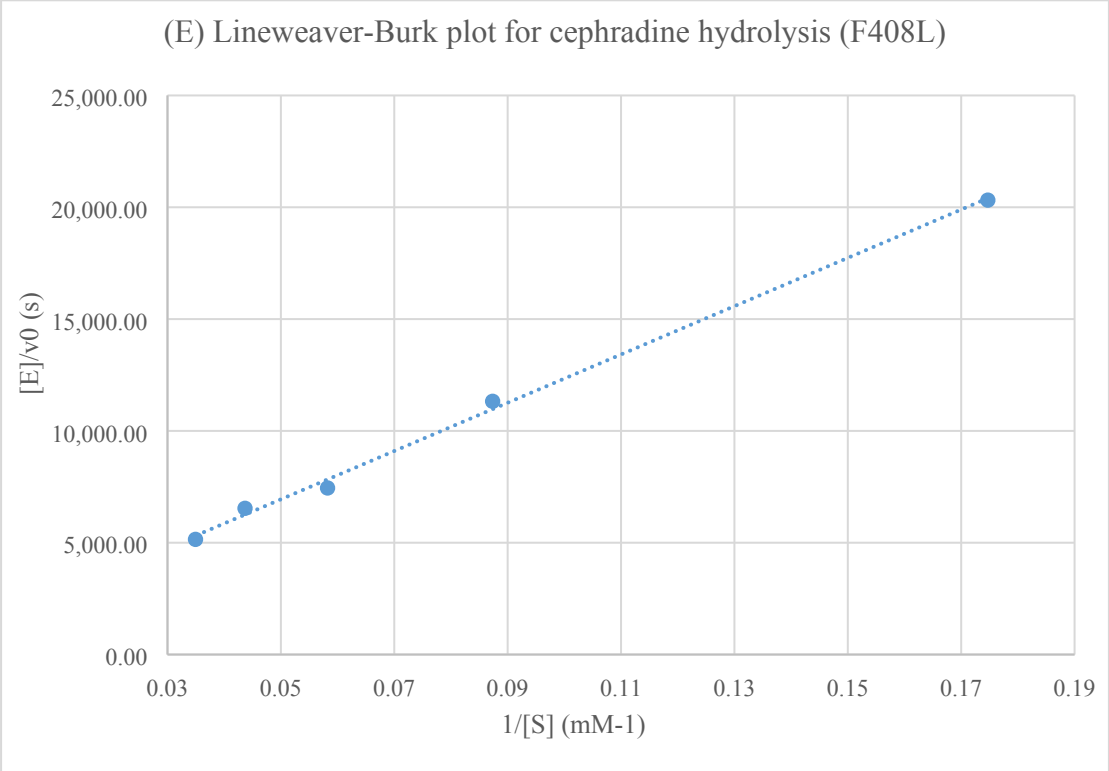
Fig. S6 Lineweaver-Burk plots for determination of steady-state kinetic parameters of the wild-type cocaine esterase and designed mutants for hydrolysis of cephradine and DHME.

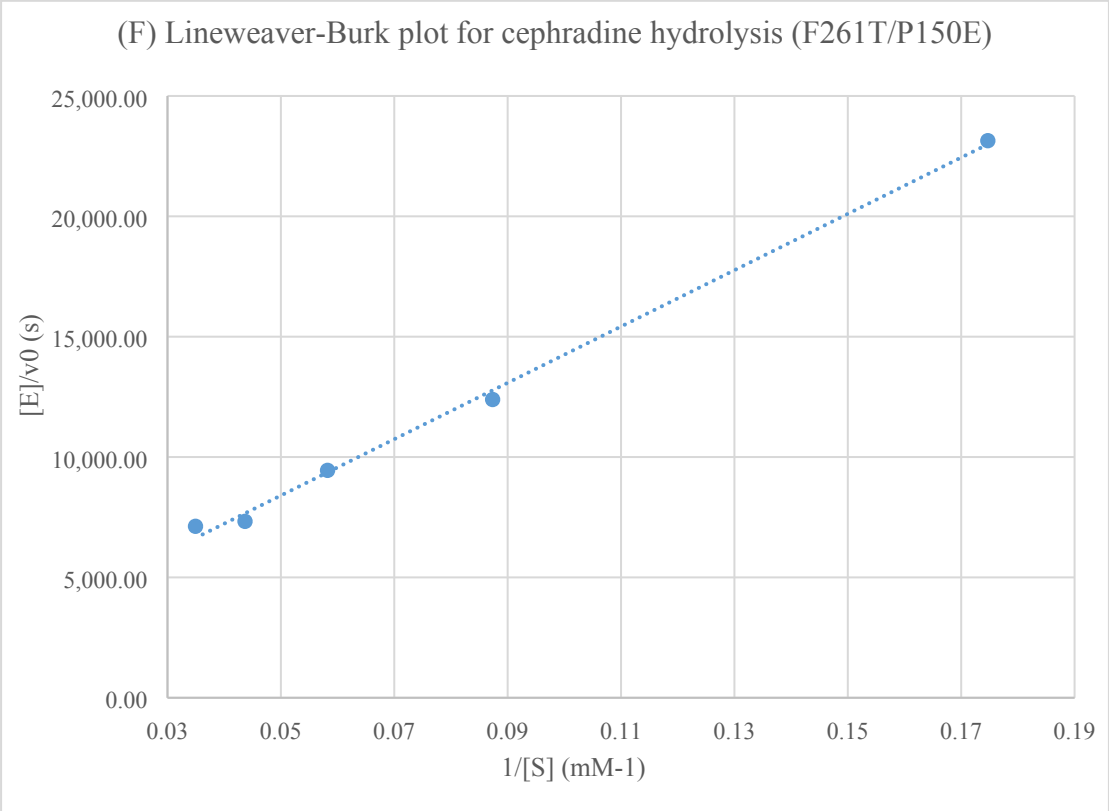


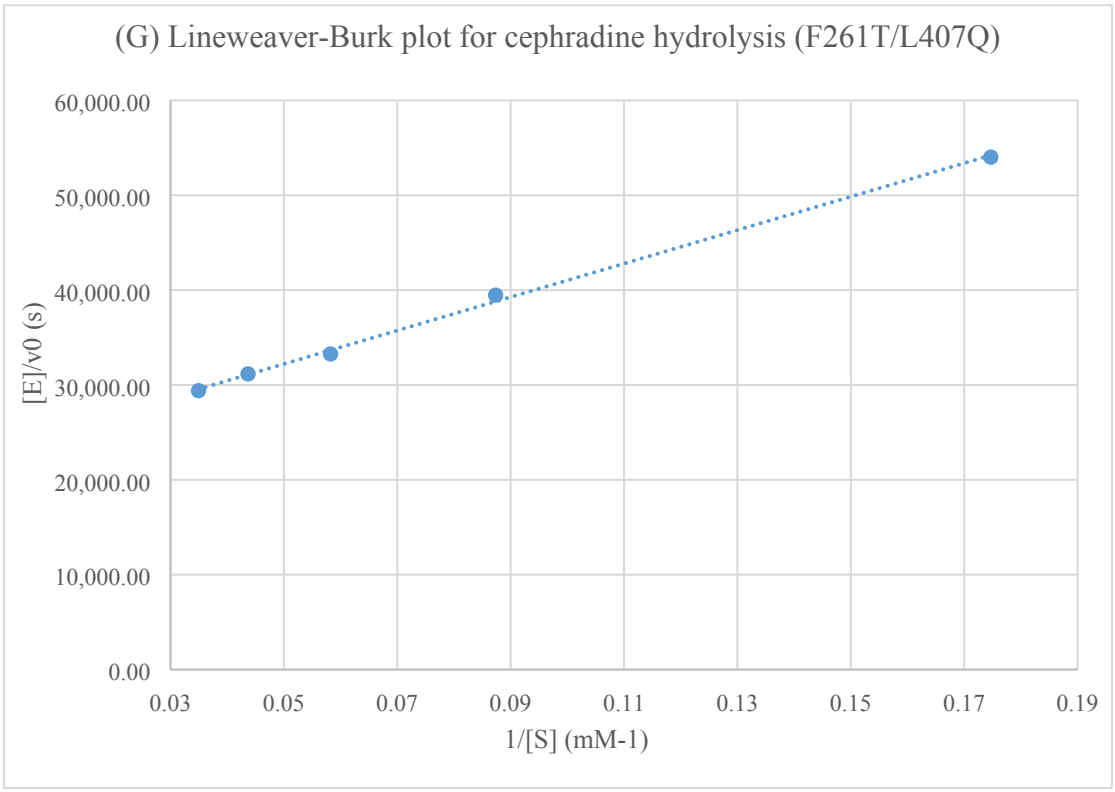


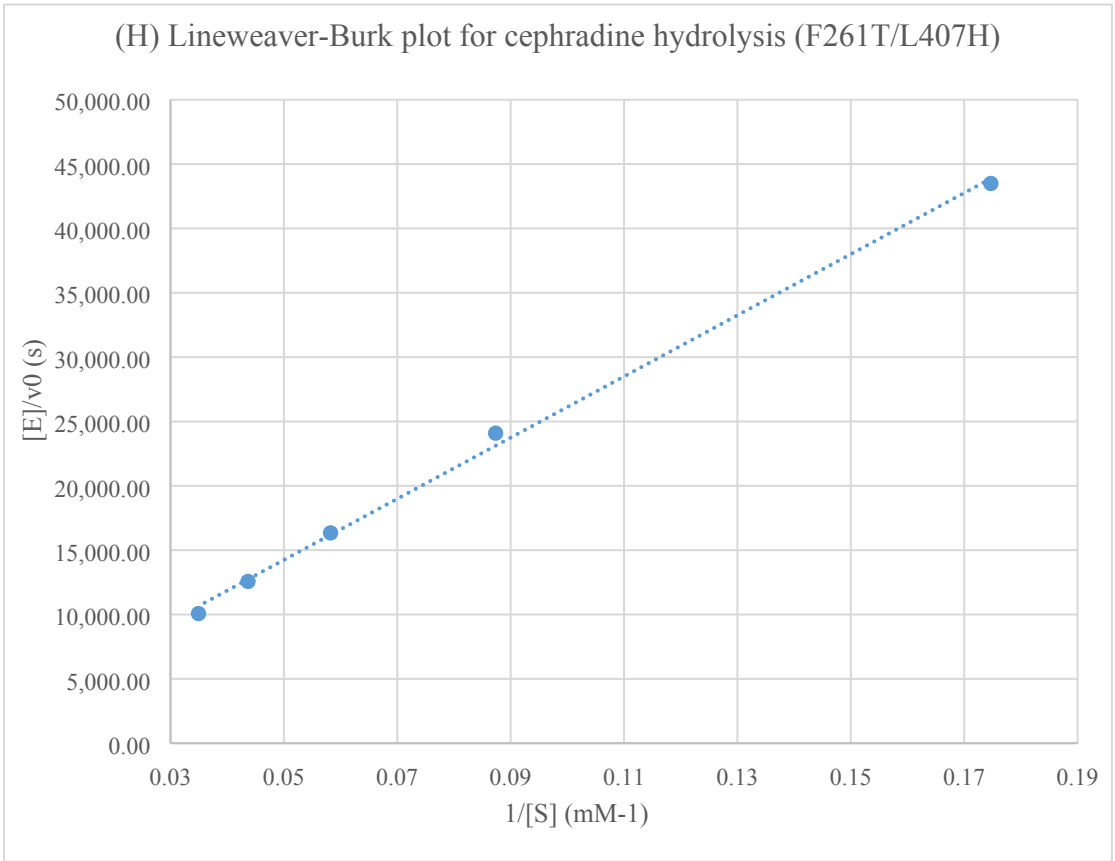

(C) Lineweaver-Burk plot for cephradine hydrolysis (Q55R)
400,000.00
350,000.00
300,000.00
250,000.00
200,000.00
150,000.00
100,000.00
50,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)

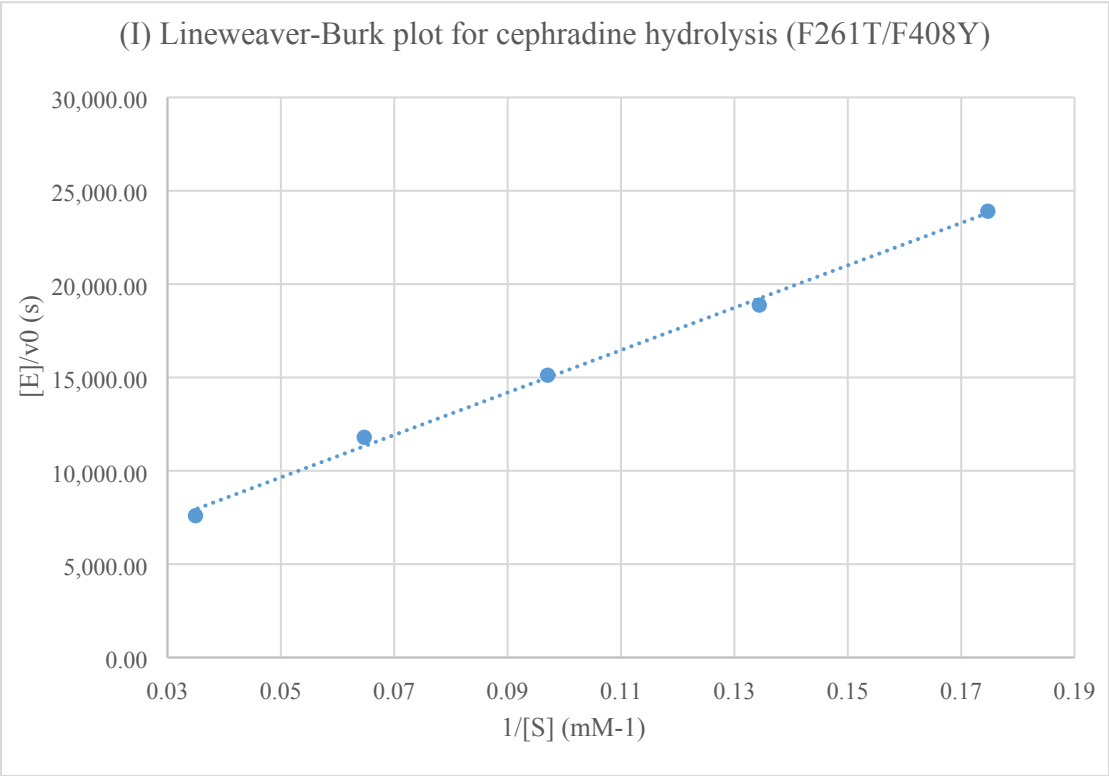

(D) Lineweaver-Burk plot for cephradine hydrolysis (P150E)
14,000.00
12,000.00
10,000.00
8,000.00
6,000.00
4,000.00
2,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)

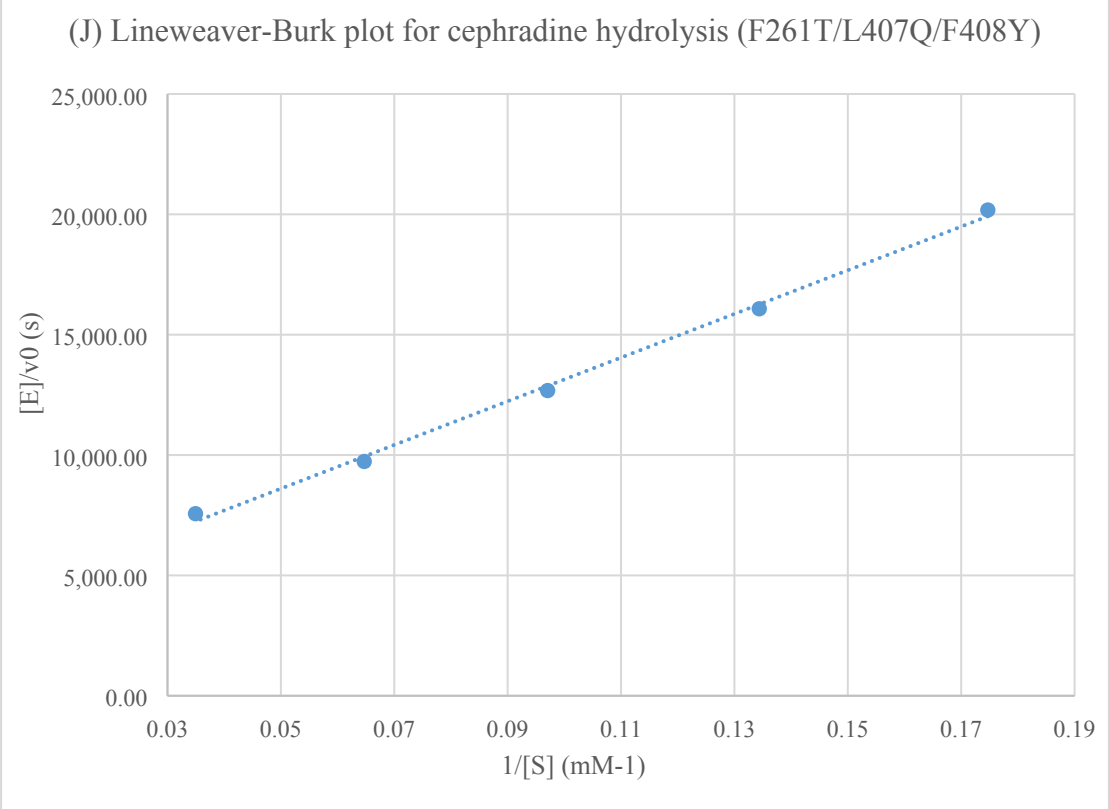

(E) Lineweaver-Burk plot for cephradine hydrolysis (F408L)
25,000.00
20,000.00
15,000.00
10,000.00
5,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)


(F) Lineweaver-Burk plot for cephradine hydrolysis (F261T/P150E)
25,000.00
20,000.00
15,000.00
10,000.00
5,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)

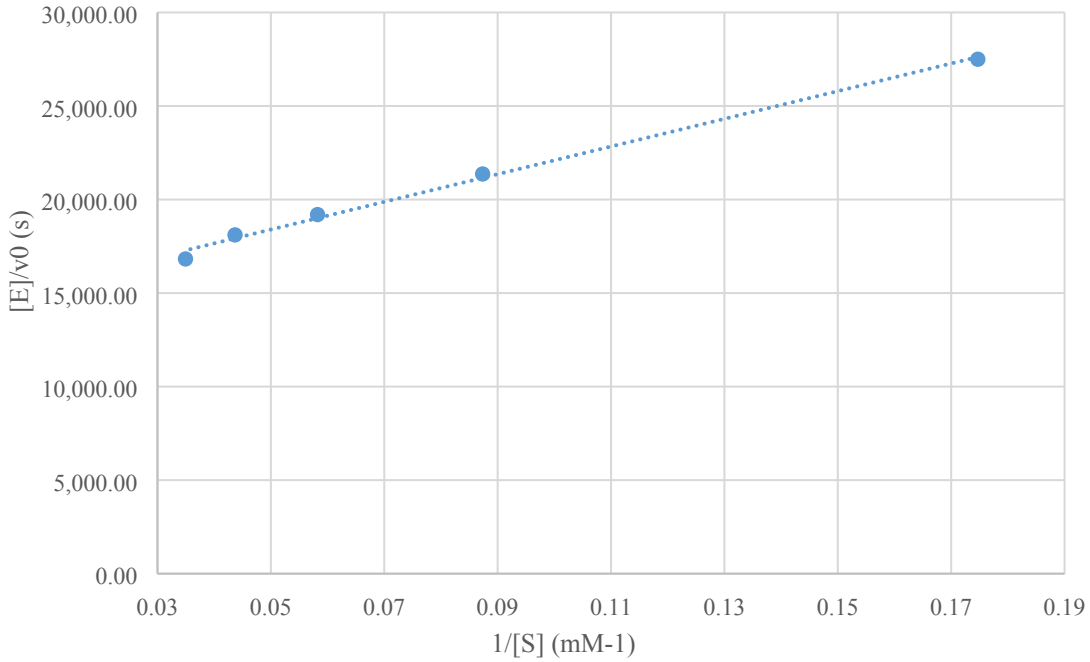

(G) Lineweaver-Burk plot for cephradine hydrolysis (F261T/L407Q)
60,000.00
50,000.00
40,000.00
30,000.00
20,000.00
10,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)


(H) Lineweaver-Burk plot for cephradine hydrolysis (F261T/L407H)
50,000.00
45,000.00
40,000.00
35,000.00
30,000.00
25,000.00
20,000.00
15,000.00
10,000.00
5,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)

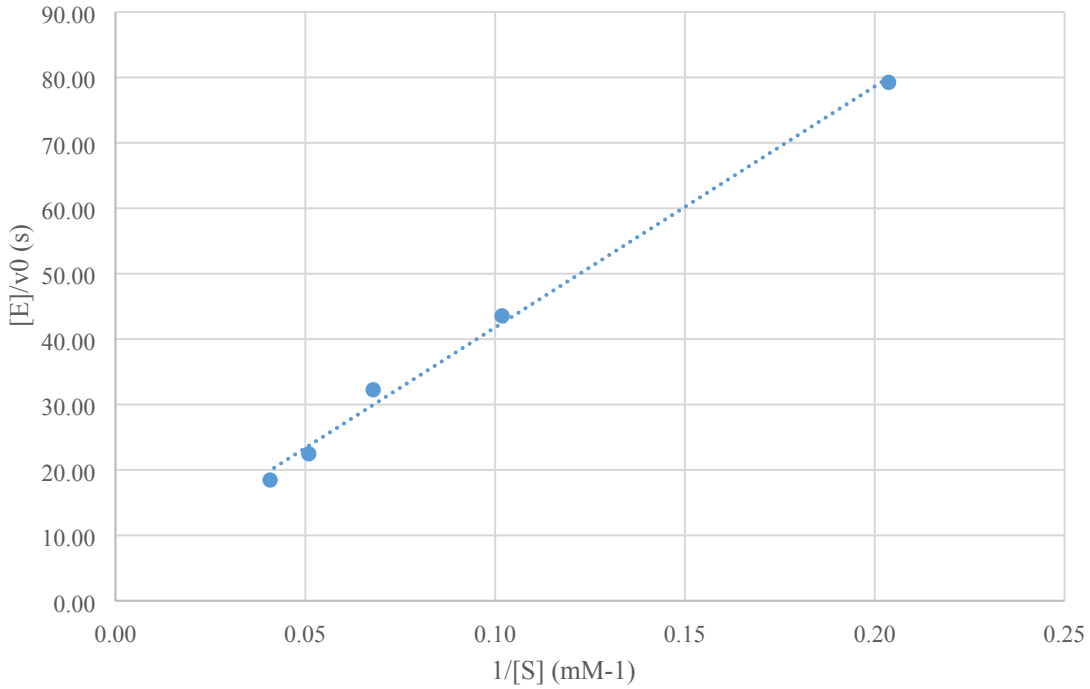

(I) Lineweaver-Burk plot for cephradine hydrolysis (F261T/F408Y)
30,000.00
25,000.00
20,000.00
15,000.00
10,000.00
5,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)


(J) Lineweaver-Burk plot for cephradine hydrolysis (F261T/L407Q/F408Y)
25,000.00
20,000.00
15,000.00
10,000.00
5,000.00
0.00
[E]/v0 (s)
0.03
0.05
0.07
0.09
0.11
0.13
0.15
0.17
0.19
1/[S] (mM-1)

(K) Lineweaver-Burk plot for cephradine hydrolysis (F261T/L407H/F408Y)



(L) Lineweaver-Burk plot for DHME hydrolysis (WT)



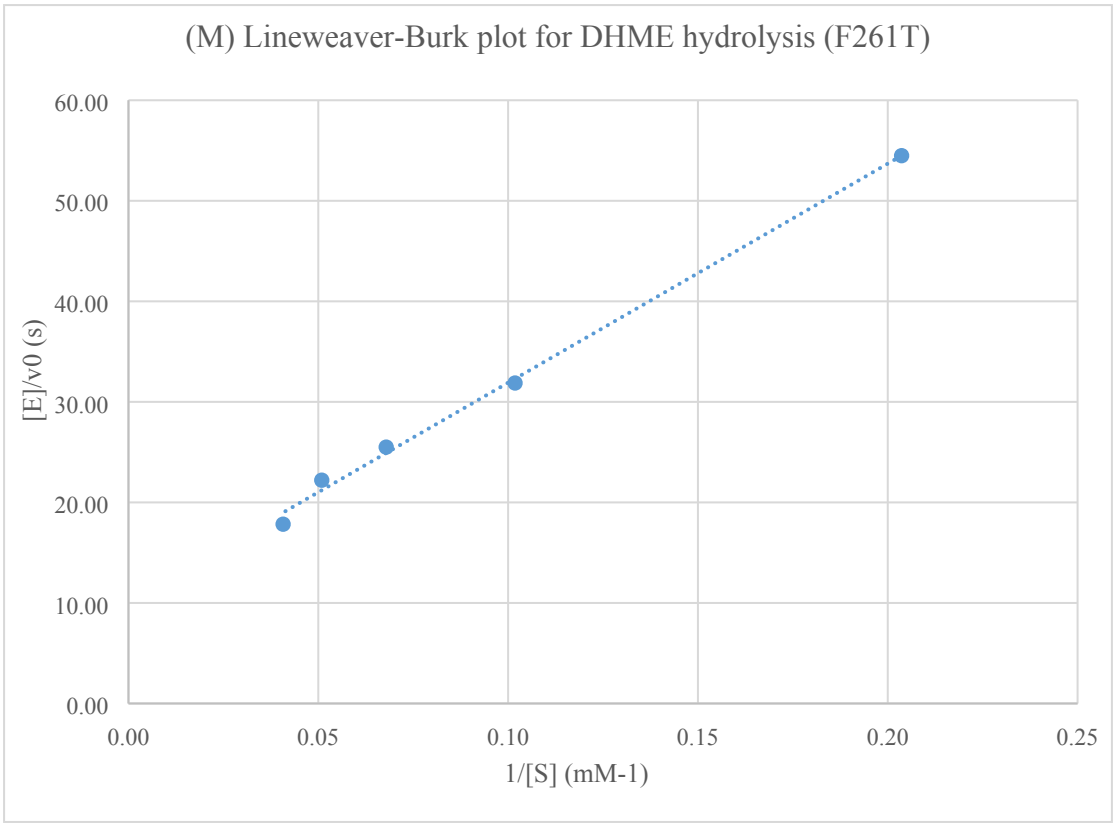
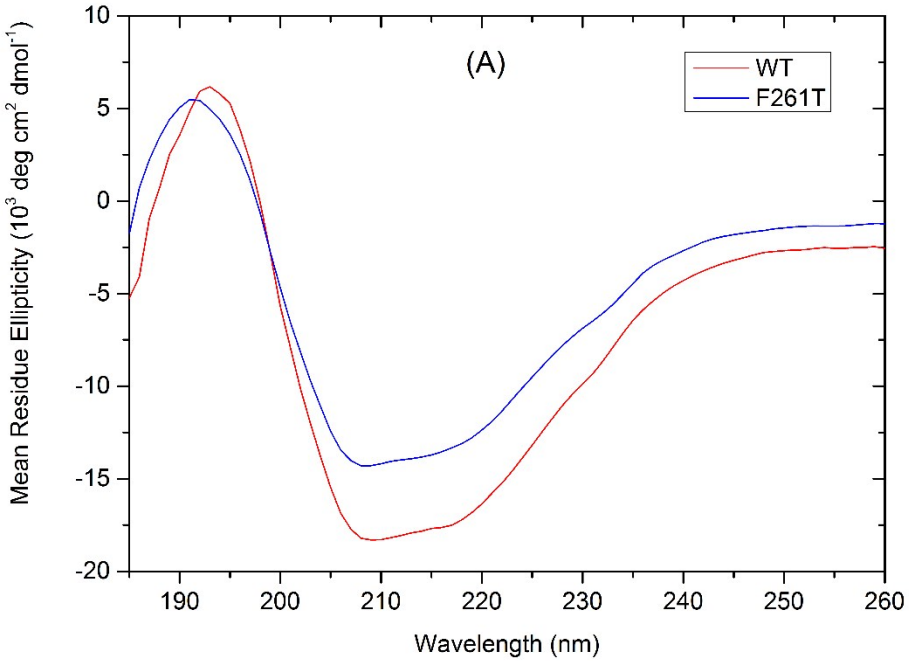

(M) Lineweaver-Burk plot for DHME hydrolysis (F261T)
60.00
50.00
40.00
30.00
20.00
10.00
0.00
[E]/v0 (s)
0.00
0.05
0.10
0.15
0.20
0.25
1/[S] (mM-1)

Fig. S7 Circular dichroism spectra and melting curves of the wild-type cocaine esterase and F261T.



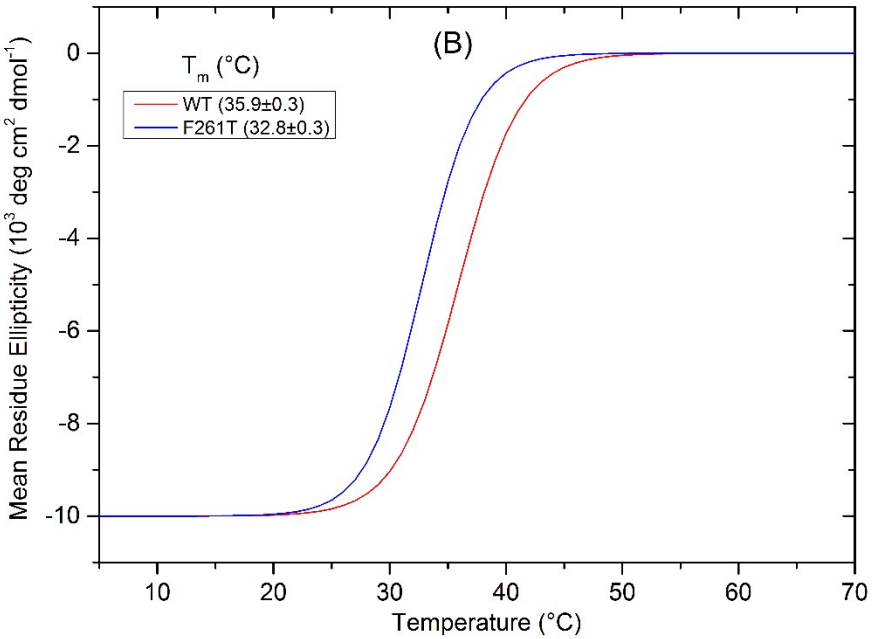
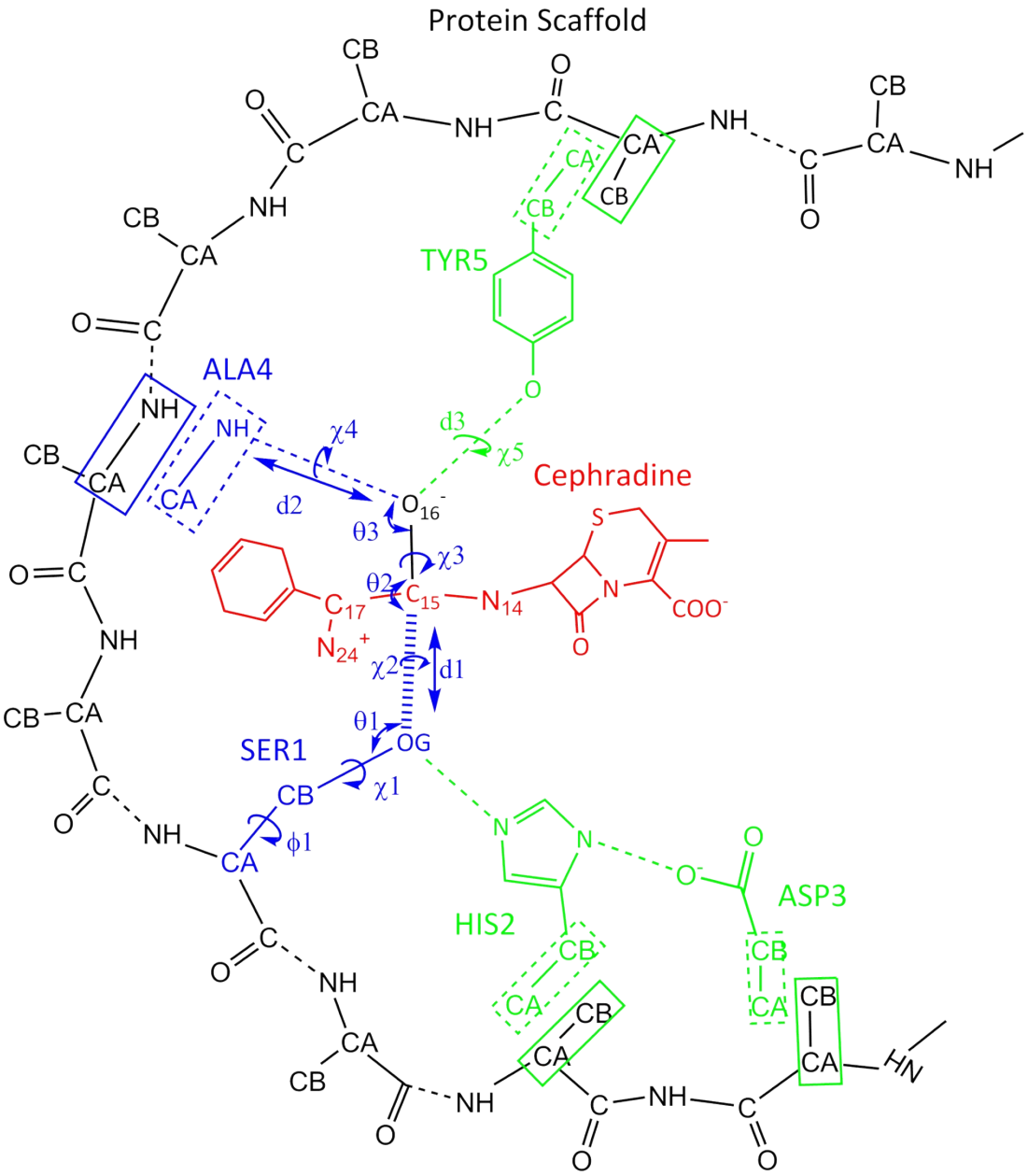


Fig. S8 Overview of the matching process in ProdaMatch.



3. Computational Methods

Active site matching

A scaffold library of 2,234 protein entries were constructed by the intersection of PDB database and CSA enzyme database using the criteria described in our earlier work. Firstly, a set of 15,174 protein structures is selected from the Protein Data Bank (PDB) according to the following criteria,[1] as: (i) The protein is expressible in *Escherichia coli*; (ii) The protein structure was measured by X-ray crystallography and its structural resolution is better than 3.0Å; (iii) The number of amino acids of the protein lies between 200 and 800; (iv) The sequence identity in this set is lower than 95%. The scaffold library was finally generated by overlapping this set with a set of 25,358 protein entries from the enzyme structural database CSA (Catalytic Site Atlas).[2] For each scaffold in the library, the anchoring positions were selected to be the sites in the vicinity of the primary catalytic sites provided by the CSA database. In detail, the centroid coordinates of the CA atoms of the native catalytic residues are calculated for each scaffold, and the residues that lie within 15 Å to the centroid are selected as anchoring positions for matching.

A revised version of ProdaMatch was used to search the scaffold library based on the complex active site model and the unified catalytic geometrical constraint parameters (Table S1).[3, 4] In the complex active site model, the transition state (TS) structure of cephradine was constructed based on its crystal structure, but the planar form of the central amide bond is converted to a tetrahedral intermediate form. Generally, the distances and angles of the catalytic constraint parameters are varied around the ideal values (Table S1), which are coincident with the values statistically obtained from a series of crystal structures of hydrolyases (Table 1 in Ref. 5).[5] An overview of the matching process in ProdaMatch is shown in Fig. S8. Generally speaking, the whole matching problem was decomposed into a series of loop-closure problems. First of all, a main loop, comprising two residues and the transition state of cephradine (SER1-cephradine-ALA4), was chosen. The target of the main loop closure problem is to anchor the calculated terminal N and CA atoms onto the backbone N and CA atoms by adjusting the variables (ϕ, d, χ and θ) using a quasi-Newton direction-based optimization algorithm. Once the main loop is successfully matched, other residues, considered as side loops in the algorithm, were anchored to potential residue positions in the protein scaffold one after another. The order of matching side loops was TYR5->HIS2->ASP3->SER6->GLU7, and the order of matching all residues and transition state of cephradine was SER1->cephradine->ALA4->TYR5-> HIS2->ASP3->SER6->GLU7. To make the figure clear and easy to understand, the residues SER6 and GLU7 are not shown in Fig. S8. Several matches may be obtained on one scaffold and the position of the TS in each match is checked by visual inspection. If the TS is located outside of the pocket of the scaffold, the match is discarded. If no feasible match is identified by ProdaMatch, the scaffold is also discarded.

The Energy function in PRODA

A combined free-energy function was used to calculate enzyme-ligand interactions in PRODA, with the reference state referring to the template of the scaffold in solvent and the isolated ligand in solvent. The total energy is a linear combination of seven energy terms listed in Eq. 1:

$$\Delta G = E_{LJVDWAttr} + E_{LinearVDWRep} + E_{HB} + E_{desolv} + E_{SC} + E_{hydrophobic} + E_{entropy} \quad (1)$$

The Lennard-Jones potential for van der Waals attractive interactions ($E_{LJVDWAttr}$), linear van der Waals repulsive interaction ($E_{LinearVDWRep}$), orientation-dependent terms for hydrogen bonds (E_{HB}), desolvation energy of polar atoms (E_{desolv}), screened Coulombic interaction (E_{SC}), hydrophobic contribution of nonpolar atom burial ($E_{hydrophobic}$), and the side chain entropy contribution term ($E_{entropy}$) were the same as those previously described.[6] The readers can refer to Ref. 6 for the details of each energy term.

Computational design and sequence selection

Scaffold 3i2k was selected because its native scaffold contains a catalytic triad based motif and a large hydrophobic active site pocket, which may be appropriate to accommodate the bulky cephradine. Water molecules in native scaffold 3i2k were removed and the amino-acid hydrogen atoms were added using PRODA based on the topology parameters of the all-atom CHARMM 22 force field.[7] The atomic coordinates of cephradine and DHME were taken from the PDB database, and hydrogen atoms were added using the molecular modeling software Discovery Studio. Twelve sequence selection positions (Gln55, His82, Val121, Leu146, Ala149, Pro150, Trp151, Ala162, Leu169, Phe261, Leu407 and Phe408) which contact with the TS directly were chosen to vary amino acid types. Five catalytic residues and another 19 residues that surrounds the sequence selection residues were chosen to alter rotameric states. A backbone-independent rotamer library compiled by Xiang and Honig,[8] which contains 11,810 original rotamers, was used to model side-chain conformations of design sites. The crystal conformation of the native amino acid at each design position was also considered as a rotamer. The rotamers for serine, threonine, and tyrosine were expanded as previously described due to the diversity configurations that the hydroxyl hydrogen atom could adopt[4]. A library of 5,470 tetrahedral TS conformers for cephradine was generated and screened using the previously proposed small-molecule placement approach.[9, 10] The atomic van der Waals parameters for cephradine were obtained from the model molecules of the CHARMM 22 force field, and the atomic partial charges were assigned based on the PARSE models.[11]

Free energy of the enzyme-TS complex system was calculated using the energy functions mentioned in our earlier work and sequence selection optimization problems were solved by the previously developed deterministic algorithm.[6, 10] To adequately search the conformational and sequence space, a large number of sequences around the global minimum energy sequence were generated by restricting that a predefined number of rotamer types were different for any two sequences.[12] The binding energy (ΔG_{Bind}) was calculated as the energy difference between the bound enzyme-TS system and the unbound enzyme-TS system,[4] as shown in Eq. 2:

$$\Delta G_{Bind} = \Delta G_{bound} - \Delta G_{unbound} \quad (2)$$

where ΔG_{bound} and $\Delta G_{unbound}$ are the free energies of the complex and the unbound enzyme-TS systems.

The folding energies are the free energies of the apo-form enzyme, equal to $\Delta G_{unbound}$, as shown in Eq. 3:

$$\Delta G_{Fold} = \Delta G_{unbound} \quad (3)$$

The binding energy change upon mutation ($\Delta\Delta G_{Bind,Mut}$) is computed as the difference of the

binding energies of the wild-type ($\Delta G_{Bind,WT}$) and mutant enzyme-TS systems ($\Delta G_{Bind,Mut}$).

$$\Delta\Delta G_{Bind,Mut} = \Delta G_{Bind,Mut} - \Delta G_{Bind,WT} \quad (4)$$

The folding energy change upon mutation ($\Delta\Delta G_{Fold,Mut}$) is computed as the difference of the folding energies of the wild-type ($\Delta G_{Fold,WT}$) and mutant enzymes ($\Delta G_{Fold,Mut}$).

$$\Delta\Delta G_{Fold,Mut} = \Delta G_{Fold,Mut} - \Delta G_{Fold,WT} \quad (5)$$

Mutations were selected for experimental characterization if one of the following criteria is satisfied: (1) the predicted binding energy of the enzyme-TS complex not significantly higher than wild type, (2) hydrogen-bonding binding energy term lower than wild type, (3) van der Waals repulsive binding energy term lower than wild type, and (4) buried nonpolar atoms binding energy term lower than wild type, where higher means more positive or less negative and lower the opposite. All calculations are performed on a computer cluster with 256 cores, where each core represents a 2.1 GHz CPU from a sub-cluster with 64 cores sharing 128 GB memory.

Example input lines for the computational programs

Example input lines for ProdaMatch

```
!/* Read parameters */
ATOMPARAM    ../data/AllResidueAtomParameter.txt
ATOMPARAM    ../batch_match/cephradine/CephradineAtomParameter.txt
TOPOLOGY     ../data/top_all22_prot.inp
TOPOLOGY     ../data/SupplementaryICForProtein.txt
TOPOLOGY     ../batch_match/cephradine/CephradineCharmmFromCephalexin.txt
SEQUENCE     ../batch_match/sca2234/1MPX/sequence.txt
SEQUENCE     ../batch_match/cephradine/Sequence_Cephradine.txt
COORDINATE   CRD          ../batch_match/cephradine/Cephradine.pdb
COORDINATE   CHAINA       ../batch_match/sca2234/1MPX/chainA.pdb

!/* Read candidate matching positions */
!MATCH_SITES              ../batch_match/pos2234/1MPX.pos
!MATCH_SITES_EXPANSION
!MATCH_SITES_WRITE        ../batch_match/pos2234_15A/1MPX.pos
MATCH_SITES               ../batch_match/pos2234_15A/1MPX.pos

!/* Match process */
MATCH        SER1   ../batch_match/cephradine/Motif1/SER1_PlacingRule.txt
             ALA4   ../batch_match/cephradine/Motif1/ALA4_PlacingRule_OH1.txt
             TYR5   ../batch_match/cephradine/Motif1/TYR5_PlacingRule_OH2.txt
             HSD2   ../batch_match/cephradine/Motif1/HSD2_PlacingRule.txt
             ASP3   ../batch_match/cephradine/Motif1/ASP3_PlacingRule.txt
             SER6   ../batch_match/cephradine/Motif1/SER6_PlacingRule.txt
             GLU7   ../batch_match/cephradine/Motif1/GLU7_PlacingRule.txt
```

These are the example input lines for the matching algorithm ProdaMatch. In each line, the

first capitalized word (i.e., ATOMPARAM, TOPOLOGY, etc) is the keyword defined in the computational program. Once the keyword is identified, the program will perform corresponding duties. The files after the keyword contain necessary data that the program needs to perform the matching process. These data are the atom parameters (ATOMPARAM), topologies of residue and small molecule (TOPOLOGY), protein and small molecule sequence (SEQUENCE), coordinates of protein scaffold and small molecule (COORDINATE), candidate sites for matching (MATCH_SITES), and placing rules for matching residues (MATCH). 'ATOMPARAM' reads atom parameters for residue and small molecule. 'TOPOLOGY' reads topologies for residue and small molecule. 'SEQUENCE' reads sequence information for protein and small molecule. 'COORDINATE' reads coordinate information for protein scaffold and small molecule. 'CRD' and 'CHAINA' are the chain name for small molecule and protein sequence, respectively. 'MATCH_SITES' reads the candidate site positions for matching. 'MATCH_SITES_EXPANSION' expand the range for selecting candidate site positions. 'MATCH_SITES_WRITE' output candidate site positions into specified file. 'MATCH' reads the placing rules for matching residues and performs the matching process. The four-character names (i.e., SER1, HSD2, etc) are the matching residues defined in the complex active site model, and the file after each residue contains the placing rules of how to match the corresponding residue. The placing rule files specify the catalytic geometrical constraints for match (see Table S1).

Example input lines for PRODA

```
!/* Read parameters */
ATOMPARAM    ../data/AllResidueAtomParameter.txt
ATOMPARAM    ../3i2k/CephradineAtomParameter.txt
TOPOLOGY     ../data/top_all22_prot.inp
TOPOLOGY     ../data/SupplementaryICForProtein.txt
TOPOLOGY     ../3i2k/CephradineCharmm.txt
SEQUENCE     ../3i2k/Sequence.txt
SEQUENCE     ../3i2k/Sequence_cephradine.txt
COORDINATE CRD          ../3i2k/Cephradine.pdb
COORDINATE CHAINA       ../3i2k/3i2k_ChainA.pdb

!/* Generate protein rotamers by rotamer library and crystal structure */
ROTAMER      ../data/rotlib11810.txt    ../3i2k/DesignScheme_Catalytic.txt
ROTAMER      ../data/rotlib11810.txt    ../3i2k/DesignScheme_NonCatalytic.txt
CRYSTAL_ROTAMER
ROTAMER_EXPANSION

!/* Generate/read small molecule rotamers by small molecule placement method */
!SMALLMOL_ROTAMER       ../3i2k/CatalyticCons.txt       ../3i2k/CephradinePlacingRule.txt
!SMALLMOL_ROTAMER_WRITE                ../3i2k/3i2k_SmallMolRotamer.pdb
SMALLMOL_ROTAMER_READ                  ../3i2k/3i2k_SmallMolRotamer.pdb

!/* Calculate energy matrix */
```

```
!ROTAMER_AREAS ../3i2k/temp/3i2kRP.txt ../3i2k/temp/3i2kRA.txt 1 1
!BORN_RADIUS ../3i2k/temp/3i2kRP.txt ../3i2k/temp/3i2kRA.txt ../3i2k/temp/3i2kBR 1 1
!SELF_ENERGY ../3i2k/temp/3i2kRP.txt ../3i2k/temp/3i2kRA.txt ../3i2k/temp/3i2kBR.txt
../3i2k/temp/3i2kSE.txt 1 1
!PARTIAL_ENERGY ../3i2k/temp/3i2kRP.txt ../3i2k/temp/3i2kRA.txt
../3i2k/temp/3i2kBR.txt ../3i2k/temp/3i2kSE.txt ../3i2k/temp/RotamerListSEC.txt
../3i2k/temp/3i2kEM4T 1 1
!ROTAMER_DELETE ../3i2k/temp/3i2kEM4T.txt ../3i2k/temp/RotamerListSEC.txt
../3i2k/CatalyticConst.txt ../3i2k/temp/RotamerListRF.txt
!ENERGY ../3i2k/temp/3i2kRP.txt ../3i2k/temp/3i2kRA.txt ../3i2k/temp/3i2kBR.txt
../3i2k/temp/RotamerListRF.txt ../3i2k/EnergyMatrix/3i2kEM 1 1

!/* Design: Amino acid sequence selection and analysis */
MILP ../3i2k/EnergyMatrix/3i2kEM.txt ../3i2k/temp/RotamerListRF.txt
../3i2k/DesignConstraint-Catalytic.txt ../3i2k/DesignConstraint-Binding.txt
../3i2k/result/3i2k_DesignResults.pdb
!POST_ANALYSIS ../3i2k/result/3i2k_DesignResults.pdb_sequence.txt
EXIT
```

These are the example input lines for the computational protein design program, PRODA. In each line, the first capitalized word (i.e., ATOMPARAM, TOPOLOGY, etc) is the keyword defined in the computational program. The files after the keyword contain necessary data that the program needs to perform the matching process. The keywords of ATOMPARAM, TOPOLOGY, SEQUENCE, COORDINATE are the same as those in ProdaMatch. 'ROTAMER' is used to generate protein rotamers for the selected amino acid design positions based on the rotamer library. 'CRYSTAL_ROTAMER' is used to add the crystal conformer of the native amino acids of the design positions into the rotamer sets. 'ROTAMER_EXPANSION' is used to expand the rotamers for serine, threonine, and tyrosine due to the diversity configurations that the hydroxyl hydrogen atom could adopt. 'SMALLMOL_ROTAMER' is used to generate rotamers for small molecule (i.e., ,ligand and TS) by the catalytic geometrical constraint parameters and placing rules. 'SMALLMOL_ROTAMER_WRITTE' outputs the structures of small molecule rotamers into a file. 'SMALLMOL_ROTAMER_READ' reads the structures of small molecule rotamers from file. The keyword from 'ROTAMER_AREAS' to 'ENERGY' are used to calculate the energy terms for the enzyme design system. 'ROTAMER_AREAS' is used to calculate the residue position for each design site and the exposed and buried areas in each rotamer. 'BORN_RADIUS' is used to calculate the born radii for all polar atoms in each rotamer and the solvation energy of the whole rotamer. 'SELF_ENERGY' is used to calculate the intrinsic energy between rotamer and the protein template. 'PARTIAL_ENERGY' is used to rapidly calculate the four-term energetics for the design system. 'ROTAMER_DELETE' is used to delete and reserve rotamers in each design position based on the four-term energies. 'ENERGY' is used to calculate the complete seven-term energetics for the reserved rotamers of the design system. 'MILP' is used to select amino acid sequences by optimizing design system energies, output design structures into '3i2k_DesignResults.pdb' and sequences into '3i2k_DesignResults.pdb_sequences.txt'. 'POST_ANALYSIS' is used to analyze the design results based on the energy terms. 'EXIT' exits the

computational program.

4. Experimental Methods

Chemicals

Cephradine was obtained from North China Pharmaceutical Co., Ltd. (Shijiazhuang, China). DHME, D-dihydrophenylglycine (DHPG) and 7-ADCA were obtained from China Shijiazhuang Pharmaceutical Group Co., Ltd. (Shijiazhuang, China).

Cloning, plasmid, strain, and mutations

The original gene (GeneBank: AF173165) was obtained from the nucleotide sequence of *Rhodococcus* sp. cocaine esterase,[13, 14] and was synthesized by Takara Biotechnology Co. (Dalian, China). The synthesized *Rh*. cocE gene was subcloned into the pET28a(+) vector with resistance to kanamycin. The gene fragment was ligated into the *NcoI* and *XhoI* sites on the vector, and the wild-type enzymes and the mutants were expressed as C-terminal hexahistidine-tagged proteins. Mutants were created by site-directed mutagenesis and the mutations were validated by automated DNA sequencing techniques.

Enzyme expression and purification

Eschericha coli BL21(DE3) cells was used as a host for the cloning and expression of wild-type and mutant enzymes, and the strains were grown overnight (over 12 hours) at 37 °C in solid LB medium containing 50 μg/mL kanamycin. Single colony growing well was transferred into 100 mL liquid LB medium with 50 μg/mL kanamycin, and incubated at 37 °C and 200 rpm for 7 hours. Then 5 mL liquid bacterial germ each was inoculated into 500 mL × 2 LB fresh liquid LB medium with 50 μg/mL kanamycin and incubated at 37 °C and 200 rpm to an OD600 of 0.6. Protein expression was induced with 0.5 mM isoproyl-β-thiogalactopyranoside and cells were harvested after a further incubation at 28 °C and 120 rpm for 20 hours. After centrifugation, the cell pellet was resuspended in the potassium phosphate buffer (100 mM, pH 7.0) and subjected to ultrasonication by 20 min cycles with 4 sec on and 6 sec off at 60% amplitude. The periplasmic extracts were centrifuged at 13,500g for 10min and the insoluble cell debris was removed. The soluble lysate was purified by Ni-NTA agarose gel chromatography in the following steps: (1) the lysate was loaded into the Ni-NTA column equilibrated with binding buffer I (100 mM potassium phosphate, 500 mM NaCl and 20 mM imidazole, pH 8.0); (2) the column was washed using binding buffer II (100 mM potassium phosphate, 500 mM NaCl and 50 mM imidazole, pH 8.0); (3) The target enzyme was eluted by the elution buffer (100 mM potassium phosphate, 500 mM NaCl and 200 mM imidazole, pH 8.0). After concentration using 10k Amicon Ultra-0.5 centrifugal filter (Millipore, Billerica, MA, USA), the purified enzymes were desalted twice using the potassium phosphate buffer (100 mM, pH 7.0) to remove NaCl and imidazole. The enzymes were >90% pure as determined by gel electrophoresis stained with Coomassie blue G250. The enzyme concentration was determined by Bradford method with the absorbance measured at 595 nm.

Activity measurement of DHME hydrolysis

The catalytic activities and kinetic parameters were determined via high-performance liquid chromatography (LC-20AT; Shimadzu, Japan) with a reversed phase Inertsil C18 column (GL Sciences, 5 μm, 150 × 4.6 mm). In detail, 0.5 mL of enzyme (in 100 mM phosphate buffer, pH 7.0) was incubated with 0.5 mL DHME solution (pH 7.0, a linear gradient with the maximum 1% w/v) in 10 mL tubes for 2 hours at 22 °C in the experimental group. The final DHME concentrations were as follows: 24.55, 19.64, 14.73, 9.82 and 4.91 mM. The reactions were terminated by mixing with 1mL chromatographic pure methanol. In the control group, 0.5 mL of enzyme in each tube (10 mL) was first deactivated at 22 °C with 1mL chromatographic pure methanol and then corresponding amounts of DHME solution was added into each tube. The final volumn of reaction mixture in each volumn is 2 mL (0.5 mL enzyme, 0.5 mL DHME solution and 1 mL chromatographic pure methanol). The reaction mixtures were kept still for 24 hours. 5 μL supernatant liquid of the reaction mixture was analyzed with high-performance liquid chromatography at 230 nm with a flow rate of 0.8 mL/min. The column temperature was kept at 25 °C. The elution buffer comprised 75% phosphate buffer (30 mM, pH 4.5) and 25% methanol (v/v). The retention times were 2.8 and 6.0 min for DHPG and DHME, respectively. The determined concentration of each target product was calculated by the difference of target signals between experimental group and control group. Steady-state kinetic parameters k_{cat} and K_m for DHME hydrolysis were obtained by fitting the data to the classic Michaelis-Menten equation via the Lineweaver-Burk method.

Activity measurement of cephradine hydrolysis

The catalytic activities and kinetic parameters were determined via high-performance liquid chromatography (LC-20AT; Shimadzu, Japan) with a reversed phase Inertsil C18 column (GL Sciences, 5 μm, 150 × 4.6 mm). In detail, 0.5 mL of enzyme (in 100 mM phosphate buffer, pH 7.0) was incubated with 0.5 mL cephradine solution (pH 7.0, a linear gradient with the maximum 2% w/v) in 10 mL tubes for 2 hours at 22 °C in the experimental group. The final cephradine concentrations were as follows: 28.62, 22.90, 17.17, 11.45 and 5.72 mM. The reactions were terminated by mixing with 1mL chromatographic pure methanol. In the control group, 0.5 mL of enzyme in each tube (10 mL) was first deactivated at 22 °C with 1mL chromatographic pure methanol and then corresponding amounts of cephradine solution was added into each tube. The final volumn of reaction mixture in each volumn is 2 mL (0.5 mL enzyme, 0.5 mL cephradine solution and 1 mL chromatographic pure methanol). The reaction mixtures were kept still for 24 hours. 5 μL supernatant liquid of the reaction mixture was analyzed with high-performance liquid chromatography at 254 nm with a flow rate of 0.8 mL/min. The column temperature was kept at 25 °C. The elution buffer comprised 75% phosphate buffer (30 mM, pH 4.5) and 25% methanol (v/v). The retention times were 3.1 and 7.2 min for 7-ADCA and cephradine, respectively. The determined concentration of each target product was calculated by the difference of target signals between experimental group and the control group. Steady-state kinetic parameters k_{cat} and K_m for cephradine hydrolysis were obtained by fitting the data to the classic Michaelis-Menten equation via the Lineweaver-Burk method.

Kinetics measurement of cephradine synthesis

The kinetic parameters V_s/V_h of cephradine synthsis were determined via high-performance liquid chromatography (LC-20AT; Shimadzu, Japan) with a reversed phase Hypersil C18 column (Dalian Elite Analytical Instruments Co., Ltd., 5 μm, 250 × 4.6 mm). In detail, 0.5 mL of enzyme (in 100mM phosphate buffer, pH 7.0) was incubated with 0.5 mL substrate solution (pH 7.0, 30 mM DHME and 60 mM 7-ADCA) in 10 mL tubes for 1 to 24 hours at 22 °C in the experimental group. The final concentrations of DHME and 7-ADCA were 15 and 30 mM, respectively. The final concentrations of wild-type and F261T enzyme were 2.37 and 2.41 μM, respectively. The reactions were terminated by mixing with 1mL chromatographic pure methanol. In the control group, 0.5 mL of enzyme in each tube (10 mL) was first deactivated at 22 °C with 1mL chromatographic pure methanol and then 0.5 mL substrate solution was added into each tube. The final volume of reaction mixture in each tube is 2 mL (0.5 mL enzyme, 0.5 mL substrate solution and 1 mL chromatographic pure methanol). The reaction mixtures were kept still for 24 hours. 5 μL supernatant liquid of the reaction mixture was analyzed with high-performance liquid chromatography at 230 nm with a flow rate of 0.8 mL/min. The column temperature was kept at 25 °C. The elution buffer comprised 75% phosphate buffer (30 mM, pH 4.5) and 25% methanol (v/v). The retention times were 3.8, 4.7, 12.5 and 14.8 min for 7-ADCA, DHPG, DHME and cephradine, respectively. The determined concentration of each target product was calculated by the difference of target signals between experimental group and control group. The parameter V_s/V_h was calculated from the 1-hour rates of production of cephradine and DHPG.

Circular dichroism measurements

To determine the circular dichroism (CD) spectra, the enzyme was firstly purified using Tris-HCl buffer (100 mM Tris, pH 8.0). The purified enzymes of the wild-type cocaine esterase and the single mutant F261T in Tris-HCl buffer (100 mM Tris, pH 8.0) were initially eluted twice with pure water (to minimize the influence on CD signals caused by Tris) and then concentrated to 0.5-0.6 mg/mL using 10k Amicon Ultra-0.5 centrifugal filter (Millipore, Billerica, MA, USA). Circular dichroism (CD) spectra (185-260 nm) were then scanned for each sample of the eluted enzyme using 0.1-cm quartz cuvette. The data were collected using a Chirascan plus CD Spectrometer equipped with a Peltier thermostat (Applied Photophysics). Mean residue ellipticity ([θ]) was calculated according to Eq. 6:

$$[\theta] = \theta \cdot M/(10 \cdot n_{resi} \cdot l \cdot c) \quad (6)$$

where θ is the ellipticity (mdeg), M is the molecular weight of the enzyme, n_{resi} is the residue number of the enzyme, l is the cell path length, and c is the mass concentration (mg/mL) of the puried enzyme. A factor of 1/10 is used to give the standard unit of $deg \cdot cm^2 \cdot dmol^{-1}$. Thermal unfolding of purified enzyme was determined at 220 nm with a temperature range of 5-70 °C (a resolution of 1 °C). The heating rate was 1 °C/min, and data were collected after an additional minute when the sample reached thermal equilibrium. Sigmoid curve was used to fit denaturation curve of each sample by Pro-Data Viewer (Applied Photophysics, version 4.2.5). The T_m of each sample was determined as the midpoint (x_0) of the normalized thermal unfolding transition. The final T_m of each enzyme was determined as the average x_0 of three scans.

5. References

1 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, Nucleic Acids Res., 2000, 28, 235.

2 N. Furnham, G. L. Holliday, T. A. de Beer, J. O. Jacobsen, W. R. Pearson, J. M. Thornton, Nucleic Acids Res., 2014, 42, 485.

3 J. Xue, X. Q. Huang, M. Lin, Y. S. Zhu, J. Mol. Model., 2016, 22, 49.

4 X. Q. Huang, J. Xue, M. Lin, Y. S. Zhu, PloS ONE, 2016, 11, e0156559.

5 A. J. Smith, R. Müller, M. D. Toscano, P. Kast, H. W. Hellinga, D. Hilvert, K. Houk, J. Am. Chem. Soc., 2008, 130, 15361.

6 Y. Tian, X. Q. Huang, Y. S. Zhu, J. Mol. Model., 2015, 21, 191.

7 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B, 1998, 102, 3586.

8 Z. Xiang, B. Honig, J. Mol. Biol., 2001, 311, 421.

9 J. K. Lassila, H. K. Privett, B. D. Allen, S. L. Mayo, Proc. Natl. Acad. Sci. USA, 2006, 103, 16710.

10 X. Q. Huang, K. H. Han, Y. S. Zhu, Protein Sci., 2013, 22, 929.

11 D. Sitkoff, K. A. Sharp, B. Honig, J. Phys. Chem., 1994, 98, 1978.

12 Y. Tian, X. Q. Huang, Q. Li, Y. S. Zhu, Appl. Microbiol. Biotechnol., 2016, 621.

13 N. A. Larsen, J. M. Turner, J. Stevens, S. J. Rosser, A. Basran, R. A. Lerner, N. C. Bruce, I. A. Wilson, Nat. Struct. Mol. Biol., 2002, 9, 17.

14 D. Narasimhan, M. R. Nance, D. Gao, M.-C. Ko, J. Macdonald, P. Tamburi, D. Yoon, D. M. Landry, J. H. Woods, C.-G. Zhan, Protein Eng. Des. Sel., 2010, 23, 537.