

Supporting Information

Molecular dynamics information improves *cis*-peptide based function annotation of proteins

Sreetama Das¹, Pratiti Bhadra², Suryanarayanarao Ramakumar¹ and Debnath Pal^{2*}

¹Department of Physics

²Department of Computational and Data Sciences

Indian Institute of Science, Bangalore 560012

India

Contact:

* Debnath Pal

Email: dpal@cds.iisc.ac.in

Telephone: +91-80-2293-2901

FAX: +91-80-2360-2648

S. Ramakumar

Email: ramak@physics.iisc.ernet.in

Telephone: +91-80-2293-2312

FAX: +91-80-2360-2602

SUPPORTING INFORMATION**TABLE OF CONTENTS**

Table S1: List of PDB entries containing one or more *cis*-peptides

Table S2: Details of matches in fragment and dynamics between F0 fragments and fragments from test dataset.

Table S3: Suitability of coarse-grained forcefield in reproducing actual protein dynamics

Table S4: Numbers of hits in different classes when searching for F0 fragments in the test dataset using match in *cis*-peptide and dynamics

Table S5: Annotation of the PDB entries with unknown function and containing fragments from the F0 dataset with matching *cis*-peptide and dynamics.

Fig S1: GO directed acyclic graph for PDB ID: 1m56A showing function representation at different GO levels.

Table S1: List of PDB files containing one or more *cis*-peptides. Identifier includes the 4-letter PDB code and chain-id. Resolution is indicated in brackets (NA indicates NMR structures).

1ucsA (0.62),	1gciA (0.78),	1x6zA (0.78),	1w0nA (0.80),	3ui4A (0.80),	1p9gA (0.84),
1m40A (0.85),	3o4pA (0.85),	4ayoA (0.85),	1muwA (0.86),	2ddxA (0.86),	1dy5A (0.87),
4hs1A (0.87),	1gweA (0.88),	1ilwA (0.89),	3ip0A (0.89),	1ix9A (0.90),	1oewA (0.90),
1vyrA (0.90),	3ks3A (0.90),	4ea9A (0.90),	4g78A (0.92),	2bt9A (0.94),	2gudA (0.94),
2uu8A (0.94),	2vhkA (0.94),	1mj5A (0.95),	1rtqA (0.95),	2xomA (0.95),	3dk9A (0.95),
3vlaA (0.95),	3vn3A (0.95),	4a02A (0.95),	4kqpA (0.95),	7a3hA (0.95),	1k5cA (0.96),
1u2hA (0.96),	2e4tA (0.96),	3agnA (0.96),	3pucA (0.96),	1xmkA (0.97),	2bf6A (0.97),
2xfrA (0.97),	3viiA (0.97),	4acjA (0.97),	1k4iA (0.98),	1unqA (0.98),	2pneA (0.98),
2v8tA (0.98),	3ju4A (0.98),	3judA (0.98),	4awtA (0.98),	4bctA (0.98),	1p1xA (0.99),
3q46A (0.99),	1eb6A (1.00),	1ga6A (1.00),	1gkmA (1.00),	1jfbA (1.00),	1mn8A (1.00),
1q6zA (1.00),	2chhA (1.00),	2erlA (1.00),	2ggcA (1.00),	2gkgA (1.00),	2qskA (1.00),
2r3lA (1.00),	2vhaA (1.00),	3l8wA (1.00),	3rwnA (1.00),	3u7qA (1.00),	3u7qB (1.00),
4axoA (1.00),	4gnrA (1.00),	4nliA (1.00),	3nbcA (1.01),	1cc8A (1.02),	1cxqA (1.02),
1rw1A (1.02),	3nvsA (1.02),	3vz9B (1.03),	1d5tA (1.04),	1wuiL (1.04),	1wuiS (1.04),
3a72A (1.04),	3warA (1.04),	1fsgA (1.05),	1kjqa (1.05),	1kmvA (1.05),	1m2dA (1.05),
1yqsA (1.05),	2c7lA (1.05),	2d5mA (1.05),	2oizA (1.05),	2v3iA (1.05),	2wf7A (1.05),
3g5sA (1.05),	3gmxA (1.05),	3iquA (1.05),	3lo8A (1.05),	3pb6X (1.05),	3r2qA (1.05),
3r4lA (1.05),	3vupA (1.05),	4i62A (1.05),	4id4A (1.05),	4ivvA (1.05),	4jfiA (1.05),
4mncA (1.05),	1n40A (1.06),	3cu9A (1.06),	4d8bA (1.06),	4e9sA (1.06),	4fk9A (1.06),
3cijA (1.07),	3vgiA (1.07),	1ds1A (1.08),	1jbeA (1.08),	1uwcA (1.08),	1w23A (1.08),
1w66A (1.08),	2y6hA (1.08),	4avrA (1.08),	4bpsA (1.08),	4m5lA (1.08),	1k5nA (1.09),
1n62B (1.09),	1qlwA (1.09),	2carA (1.09),	4gnuA (1.09),	4nqrA (1.09),	1itxA (1.10),
1l3kA (1.10),	1oh0A (1.10),	1r0rI (1.10),	1su8A (1.10),	1t2dA (1.10),	1yfqA (1.10),
1z2uA (1.10),	2absA (1.10),	2fbaA (1.10),	2nxvA (1.10),	2osxA (1.10),	2w39A (1.10),
2z72A (1.10),	3bvxA (1.10),	3e5tA (1.10),	3kweA (1.10),	3lqbA (1.10),	3mvsA (1.10),
3nznB (1.10),	3rznA (1.10),	3tg2A (1.10),	4a9vA (1.10),	4ateA (1.10),	4bt7A (1.10),
4eirA (1.10),	4ghoA (1.10),	4h7wA (1.10),	4k6bA (1.10),	4l05A (1.10),	4l9dA (1.10),
4lf0A (1.10),	4lqtA (1.10),	4mhpA (1.10),	4mknA (1.10),	4ng5B (1.10),	4nyxA (1.10),
4mf5A (1.11),	1h1nA (1.12),	1ra0A (1.12),	2eabA (1.12),	3g5tA (1.12),	4i4oA (1.12),
4knkA (1.12),	1u07A (1.13),	3einA (1.13),	3jq0A (1.13),	3u8lA (1.13),	3s2rA (1.14),
1odmA (1.15),	1z70X (1.15),	2ciwA (1.15),	2d8dA (1.15),	2dkjA (1.15),	2fhzA (1.15),
2fhzB (1.15),	2uuyB (1.15),	2v7fA (1.15),	2zptX (1.15),	3bfoB (1.15),	3h4tA (1.15),
3wkqA (1.15),	3zzoA (1.15),	4eziA (1.15),	4hk8A (1.15),	4jf5A (1.15),	4jxrA (1.15),
4nx7A (1.15),	1hbnB (1.16),	1x9iA (1.16),	3awuA (1.16),	3c6aA (1.16),	4al0A (1.16),

Supplementary Material

S. Das et al.

3ondA (1.17),	1kq6A (1.18),	3zitA (1.18),	4e2vA (1.18),	2fj8A (1.19),	2xioA (1.19),
3a5fA (1.19),	3essA (1.19),	4b3bA (1.19),	1bx7A (1.20),	1j98A (1.20),	1jetA (1.20),
1jg1A (1.20),	1lc0A (1.20),	1m15A (1.20),	1munA (1.20),	1qw9A (1.20),	1ryoA (1.20),
1v8hA (1.20),	1w6sA (1.20),	1wc2A (1.20),	1wn2A (1.20),	1xdnA (1.20),	1z2nX (1.20),
2blnA (1.20),	2c1vA (1.20),	2fkkA (1.20),	2h1vA (1.20),	2omlA (1.20),	2pgnA (1.20),
2w5qA (1.20),	2znrA (1.20),	3b4uA (1.20),	3bogA (1.20),	3bwzA (1.20),	3dqyA (1.20),
3f0dA (1.20),	3gneA (1.20),	3h5jA (1.20),	3hynA (1.20),	3m0zA (1.20),	3n17A (1.20),
3og2A (1.20),	3oxpA (1.20),	3pxlA (1.20),	3sg0A (1.20),	3uf7A (1.20),	3vqjA (1.20),
3vwnX (1.20),	3znyA (1.20),	4dvcA (1.20),	4fglA (1.20),	4fr9A (1.20),	4g10A (1.20),
4gmuA (1.20),	4i3bB (1.20),	4iiyA (1.20),	4k7yA (1.20),	4kv7A (1.20),	4mjeA (1.20),
4netA (1.20),	4nfhA (1.20),	4nogA (1.20),	4nyqA (1.20),	4nzuL (1.20),	4pq9A (1.20),
1e29A (1.21),	2wnpF (1.21),	3b5mA (1.21),	3bmzA (1.21),	4hqzA (1.22),	1oi7A (1.23),
1w7cA (1.23),	3h4xA (1.23),	4klxA (1.23),	1uscA (1.24),	2bjiA (1.24),	3cp5A (1.24),
3qu5A (1.24),	3s7oA (1.24),	4e3xA (1.24),	4igtA (1.24),	1eaqA (1.25),	1gnlA (1.25),
1m4lA (1.25),	1mexL (1.25),	1mf7A (1.25),	1s3cA (1.25),	1zmaA (1.25),	2b82A (1.25),
2c60A (1.25),	2p2sA (1.25),	2tpsA (1.25),	2w8tA (1.25),	2wlvA (1.25),	3amrA (1.25),
3chbD (1.25),	3gveA (1.25),	3iixA (1.25),	3oyvA (1.25),	3pplA (1.25),	3qzmA (1.25),
3rfeA (1.25),	3sggA (1.25),	3u9wA (1.25),	3zfpA (1.25),	4attA (1.25),	4h4iA (1.25),
4ieuA (1.25),	2e7zA (1.26),	3nheA (1.26),	3rl5A (1.26),	3ur8A (1.26),	3zn4A (1.26),
4c6eA (1.26),	4j4zA (1.26),	4ldcA (1.26),	3ajdA (1.27),	3sxxA (1.27),	3wp4A (1.27),
4gq4A (1.27),	4mumA (1.27),	1rttA (1.28),	2dxuA (1.28),	2pc1A (1.28),	2y7eA (1.28),
4b97A (1.28),	4ep4A (1.28),	4euoA (1.28),	4lttA (1.28),	1gkpA (1.29),	2aebA (1.29),
2z26A (1.29),	3aofA (1.29),	3mu7A (1.29),	4ihmA (1.29),	1gk9B (1.30),	1hdhA (1.30),
1hyoA (1.30),	1jl1A (1.30),	1jndA (1.30),	1k3yA (1.30),	1ka1A (1.30),	1lzlA (1.30),
1o2dA (1.30),	1qddA (1.30),	1qwY (1.30),	1sg4A (1.30),	1vd6A (1.30),	1vlyA (1.30),
1wvfA (1.30),	1z6mA (1.30),	2d5wA (1.30),	2ehpA (1.30),	2gsoA (1.30),	2i8tA (1.30),
2imfA (1.30),	2imqX (1.30),	2mhrA (1.30),	2nqwA (1.30),	2pv2A (1.30),	2vlaA (1.30),
2vokA (1.30),	2x5nA (1.30),	2xhfA (1.30),	2zwuA (1.30),	3aj7A (1.30),	3b0tA (1.30),
3b34A (1.30),	3c2uA (1.30),	3d3bJ (1.30),	3drfA (1.30),	3dx1A (1.30),	3e8tA (1.30),
3eojA (1.30),	3f2zA (1.30),	3fegA (1.30),	3ga4A (1.30),	3hlxA (1.30),	3jtmA (1.30),
3jtzA (1.30),	3lyeA (1.30),	3ognA (1.30),	3pn3A (1.30),	3t2cA (1.30),	3wczA (1.30),
3wdqA (1.30),	3wh2A (1.30),	3znvA (1.30),	4ac1X (1.30),	4batA (1.30),	4bn4A (1.30),
4brcA (1.30),	4gvqA (1.30),	4ht3A (1.30),	4ht3B (1.30),	4i66A (1.30),	4i7wA (1.30),
4ighA (1.30),	4jvoB (1.30),	4kemA (1.30),	4la9A (1.30),	4mq1A (1.30),	4n30A (1.30),
4naxA (1.30),	4oh0A (1.30),	1vmhA (1.31),	2rdqA (1.31),	2xqqA (1.31),	4nbpA (1.31),
1gxmA (1.32),	1mkkA (1.32),	2qt1A (1.32),	2zuxA (1.32),	4edhA (1.32),	4njyA (1.32),

Supplementary Material

S. Das et al.

2uv4A (1.33),	2xtsA (1.33),	3k05A (1.33),	3nggA (1.33),	4ippA (1.33),	3sukA (1.34),
4bgbA (1.34),	1ijyA (1.35),	1v0wA (1.35),	1wddA (1.35),	2gkeA (1.35),	2i49A (1.35),
2ovgA (1.35),	2p9wA (1.35),	2phnA (1.35),	2qe8A (1.35),	2vovA (1.35),	2wqfA (1.35),
2xr6A (1.35),	2yfoA (1.35),	2z51A (1.35),	2z5wA (1.35),	3bc9A (1.35),	3ckmA (1.35),
3fghA (1.35),	3gnzP (1.35),	3i4gA (1.35),	3ib5A (1.35),	3no2A (1.35),	3p0bA (1.35),
3piuA (1.35),	3qxzA (1.35),	3szyA (1.35),	3v7pA (1.35),	3vvvA (1.35),	3w0tB (1.35),
3znxA (1.35),	4a6rA (1.35),	4ak2A (1.35),	4bazA (1.35),	4bmbA (1.35),	4di9A (1.35),
4dknA (1.35),	4efiA (1.35),	4gvfA (1.35),	4h4dA (1.35),	4h6cA (1.35),	4ipuA (1.35),
4ix3A (1.35),	4mv2A (1.35),	4nmuA (1.35),	2p8gA (1.36),	3b0xA (1.36),	3l84A (1.36),
3vyvB (1.36),	4az6A (1.36),	4h5iA (1.36),	4muqA (1.36),	1rgzA (1.37),	3b79A (1.37),
3cjsB (1.37),	3ct5A (1.37),	4fioA (1.37),	4hwvA (1.37),	4l6gA (1.37),	1hx0A (1.38),
1s9uA (1.38),	2jekA (1.38),	4l77A (1.38),	4l7gA (1.38),	4psrA (1.38),	2bfdB (1.39),
2c3vA (1.39),	2vwsA (1.39),	2y24A (1.39),	3cp7A (1.39),	3vk5A (1.39),	3vxjA (1.39),
4b7hA (1.39),	4gwqA (1.39),	1ecaA (1.40),	1es5A (1.40),	1f8eA (1.40),	1ifrA (1.40),
1k0mA (1.40),	1kqrA (1.40),	1ks8A (1.40),	1llfA (1.40),	1mlfA (1.40),	1m22A (1.40),
1m55A (1.40),	1q7lA (1.40),	1qnrA (1.40),	1r26A (1.40),	1r6xA (1.40),	1s0pA (1.40),
1u7gA (1.40),	1u7iA (1.40),	1w5qA (1.40),	1y43B (1.40),	1ye8A (1.40),	1zi8A (1.40),
1zk5A (1.40),	2cxnA (1.40),	2e3nA (1.40),	2fcoA (1.40),	2filA (1.40),	2fsqA (1.40),
2gu9A (1.40),	2hoxA (1.40),	2hzlA (1.40),	2j73A (1.40),	2octA (1.40),	2pofA (1.40),
2q7wA (1.40),	2qpxA (1.40),	2vngA (1.40),	2vptA (1.40),	2w7aA (1.40),	2w9lA (1.40),
2whlA (1.40),	2yc3A (1.40),	2ycdA (1.40),	2zwsA (1.40),	3a09A (1.40),	3achA (1.40),
3aiaA (1.40),	3glpA (1.40),	3ghaA (1.40),	3h9cA (1.40),	3itqA (1.40),	3kgkA (1.40),
3mvcA (1.40),	3oajB (1.40),	3obqA (1.40),	3om0A (1.40),	3pzwA (1.40),	3q62A (1.40),
3qnsA (1.40),	3v75A (1.40),	3vypA (1.40),	3w88A (1.40),	3wmyA (1.40),	4alzA (1.40),
4avsA (1.40),	4b6gA (1.40),	4do4A (1.40),	4e2xA (1.40),	4f0lA (1.40),	4fh0A (1.40),
4fp5D (1.40),	4h4rA (1.40),	4h4vA (1.40),	4hakA (1.40),	4halA (1.40),	4hbqA (1.40),
4i3gA (1.40),	4kn9L (1.40),	4kn9S (1.40),	4knnB (1.40),	4lxqA (1.40),	4m0kA (1.40),
4maiA (1.40),	4msoA (1.40),	4mxtA (1.40),	4n5wA (1.40),	4n6pA (1.40),	4o5oA (1.40),
2bwqA (1.41),	2yimA (1.41),	3aslA (1.41),	3ctpA (1.41),	3fsoA (1.41),	3lzcA (1.41),
1pkhA (1.42),	2f22A (1.42),	2oztA (1.42),	2pqxA (1.42),	2wuuA (1.42),	3a35A (1.42),
3cbpA (1.42),	3mabA (1.42),	4lerA (1.42),	1v05A (1.43),	2dplA (1.43),	3fssA (1.43),
3wp6A (1.43),	3zmrA (1.43),	4kavA (1.43),	2pr7A (1.44),	2qjlA (1.44),	3k7iB (1.44),
3lmzA (1.44),	3mbrX (1.44),	3w6zA (1.44),	3zxkA (1.44),	4e4rA (1.44),	4kt3A (1.44),
4n4bA (1.44),	1e6uA (1.45),	1i7lA (1.45),	1q0rA (1.45),	1qh5B (1.45),	1qqfA (1.45),
1wlhA (1.45),	1x54A (1.45),	1xbiA (1.45),	2cvdA (1.45),	2egvA (1.45),	2f5tX (1.45),
2jl1vA (1.45),	2pq8A (1.45),	2vchA (1.45),	2wlrA (1.45),	2wolA (1.45),	2xziA (1.45),

Supplementary Material

S. Das et al.

2zouA (1.45),	3b4nA (1.45),	3fucA (1.45),	3giwA (1.45),	3ie4A (1.45),	3ledA (1.45),
3m0mA (1.45),	3mwxA (1.45),	3n2wA (1.45),	3o7bA (1.45),	3o9zA (1.45),	3pojA (1.45),
3r5tA (1.45),	3rtlA (1.45),	3t94A (1.45),	3u2uA (1.45),	3uueA (1.45),	3v39A (1.45),
4b21A (1.45),	4bf5A (1.45),	4c4cA (1.45),	4j5hA (1.45),	4kmrA (1.45),	4l2iA (1.45),
4l2iB (1.45),	4l6dA (1.45),	4l9pB (1.45),	4ldvA (1.45),	4nzcA (1.45),	lg3pA (1.46),
1lc5A (1.46),	1tkeA (1.46),	3w8mA (1.46),	4hvyA (1.46),	1smoB (1.47),	2wzbA (1.47),
3gjyA (1.47),	3mjfA (1.47),	3n4jA (1.47),	3zdbA (1.47),	4bb9A (1.47),	4jduA (1.47),
4mamA (1.47),	1gqiA (1.48),	2olmA (1.48),	3aj6B (1.48),	3robA (1.48),	3tjmA (1.48),
3vmkA (1.48),	3vsvA (1.48),	3wdfA (1.48),	4eu9A (1.48),	4ftfA (1.48),	4hqsA (1.48),
4lgyA (1.48),	4mb4A (1.48),	2ovjA (1.49),	2vxtI (1.49),	2w40A (1.49),	2wiY (1.49),
2y7lA (1.49),	3fw9A (1.49),	3h87A (1.49),	3icvA (1.49),	3mcxA (1.49),	3ts3A (1.49),
4gaiA (1.49),	4htuB (1.49),	4ixlA (1.49),	4j7aA (1.49),	4oq4A (1.49),	1a4iB (1.50),
1ah7A (1.50),	1brtA (1.50),	1dj0A (1.50),	1f46A (1.50),	1i0rA (1.50),	1jl0A (1.50),
1jz8B (1.50),	1l7aA (1.50),	1lmiA (1.50),	1lucA (1.50),	1lyqA (1.50),	1m7jA (1.50),
1nkgA (1.50),	1nlqA (1.50),	1od6A (1.50),	1oi0A (1.50),	1plmA (1.50),	1p3cA (1.50),
1r5lA (1.50),	1rocA (1.50),	1sx5A (1.50),	1uuqA (1.50),	1v5vA (1.50),	1wmhB (1.50),
1wvhA (1.50),	1z21A (1.50),	1z3xA (1.50),	1z6nA (1.50),	1zhxA (1.50),	2asbA (1.50),
2bwrA (1.50),	2c61A (1.50),	2cbzA (1.50),	2cibA (1.50),	2czsA (1.50),	2dpfA (1.50),
2dskA (1.50),	2g8sA (1.50),	2gs5A (1.50),	2gs8A (1.50),	2hp0A (1.50),	2imjA (1.50),
2inwA (1.50),	2iw1A (1.50),	2j9oA (1.50),	2jc9A (1.50),	2jg0A (1.50),	2mcmA (1.50),
2o1qA (1.50),	2o6pA (1.50),	2oaaA (1.50),	2pxrC (1.50),	2q0sA (1.50),	2r6jA (1.50),
2rkqA (1.50),	2uvkA (1.50),	2uzcB (1.50),	2v8iA (1.50),	2vv6B (1.50),	2wh6A (1.50),
2whmA (1.50),	2wqkA (1.50),	2xbgA (1.50),	2xepA (1.50),	2xhgA (1.50),	2xtpA (1.50),
2yn0A (1.50),	2yogA (1.50),	2z0jA (1.50),	2z6rA (1.50),	2zm9A (1.50),	3a2zA (1.50),
3a57A (1.50),	3bb0A (1.50),	3c8eA (1.50),	3c9qA (1.50),	3cqlA (1.50),	3cz1A (1.50),
3d59B (1.50),	3dgtA (1.50),	3ejnA (1.50),	3fvsA (1.50),	3g02B (1.50),	3g89A (1.50),
3gg7A (1.50),	3gmG (1.50),	3h0uA (1.50),	3h79A (1.50),	3h8gA (1.50),	3irpX (1.50),
3iv4A (1.50),	3l4nA (1.50),	3llxA (1.50),	3md9A (1.50),	3mdqA (1.50),	3n0xA (1.50),
3o7iA (1.50),	3pfeA (1.50),	3q2iA (1.50),	3q64A (1.50),	3rjuA (1.50),	3sc7X (1.50),
3scyA (1.50),	3sz3A (1.50),	3t9wA (1.50),	3uplA (1.50),	3vnyA (1.50),	3vv1A (1.50),
3vwcA (1.50),	3vx0A (1.50),	3vykA (1.50),	3wc3A (1.50),	3zquA (1.50),	3zypA (1.50),
4aiwA (1.50),	4asmB (1.50),	4b89A (1.50),	4bq1A (1.50),	4bu0A (1.50),	4dq6A (1.50),
4e15B (1.50),	4ecqA (1.50),	4g4pA (1.50),	4gcyA (1.50),	4h2qA (1.50),	4h4pA (1.50),
4h4tA (1.50),	4h4xA (1.50),	4hjlB (1.50),	4ic4A (1.50),	4iedA (1.50),	4ihzA (1.50),
4isvA (1.50),	4isvB (1.50),	4j7nA (1.50),	4kp6A (1.50),	4lr2A (1.50),	4mlnA (1.50),
4mckA (1.50),	4mdxA (1.50),	4mk3A (1.50),	4morA (1.50),	4o0kA (1.50),	3nftA (1.51),

Supplementary Material

S. Das et al.

4hhrA (1.51),	4j9yB (1.51),	4ljiA (1.51),	1qq5A (1.52),	2a2kA (1.52),	2fsrA (1.52),
2p17A (1.52),	3futA (1.52),	3ge3A (1.52),	3pv8A (1.52),	4ay0A (1.52),	4ew1A (1.52),
4mykA (1.52),	1h16A (1.53),	1t6cA (1.53),	1usgA (1.53),	3axdA (1.53),	3puiA (1.53),
3t41A (1.53),	3v5cA (1.53),	3wjpa (1.53),	3zn6A (1.53),	4gd6A (1.53),	4gobB (1.53),
4obbA (1.53),	2hhcA (1.54),	3d7fA (1.54),	3va4B (1.54),	3vmvA (1.54),	3zy2A (1.54),
1ejdA (1.55),	1eluA (1.55),	1iuqA (1.55),	1jkeA (1.55),	1nthA (1.55),	1oqjA (1.55),
1tluA (1.55),	1ugiA (1.55),	1uteA (1.55),	1v11A (1.55),	1wbhA (1.55),	2elcA (1.55),
2q6kA (1.55),	2qmcB (1.55),	2uytA (1.55),	2wi8A (1.55),	2wnkA (1.55),	2yh6A (1.55),
2z1eA (1.55),	3cxnA (1.55),	3jq1A (1.55),	3k1zA (1.55),	3mxnA (1.55),	3mxnB (1.55),
3nr5A (1.55),	3nzmA (1.55),	3pohA (1.55),	3qguA (1.55),	3r4zA (1.55),	3tnyA (1.55),
3tosA (1.55),	3tt9A (1.55),	3vwpA (1.55),	3wjtA (1.55),	4bwjA (1.55),	4c0rA (1.55),
4g54A (1.55),	4gb5A (1.55),	4h14A (1.55),	4i7uA (1.55),	4ik8A (1.55),	4itcA (1.55),
4iyaA (1.55),	4j8sA (1.55),	4jemA (1.55),	4lixA (1.55),	4mcdA (1.55),	4n6cA (1.55),
4o6gA (1.55),	2ooaA (1.56),	3aowA (1.56),	3bpkA (1.56),	3hj4A (1.56),	4c81A (1.56),
4chbB (1.56),	4g38A (1.56),	4ikcA (1.56),	4inkA (1.56),	1jovA (1.57),	1zz1A (1.57),
2f9hA (1.57),	2imhA (1.57),	2pw0A (1.57),	3h9mA (1.57),	3nsxA (1.57),	3on9A (1.57),
3pmsA (1.57),	3tg7A (1.57),	3uidA (1.57),	4hstB (1.57),	4ly1A (1.57),	4nzjA (1.57),
2wcwA (1.58),	2ybyA (1.58),	2za4B (1.58),	3b9tA (1.58),	3dl0B (1.58),	3ho7A (1.58),
3tutA (1.58),	3whrA (1.58),	3zboA (1.58),	4ci9A (1.58),	4es8A (1.58),	4ilfA (1.58),
1r91A (1.59),	2nnuA (1.59),	3h31A (1.59),	3nreA (1.59),	3qwbA (1.59),	3tqlA (1.59),
4a8tA (1.59),	4b8vA (1.59),	1btka (1.60),	1dd9A (1.60),	1dj7A (1.60),	1e2wA (1.60),
1fcqA (1.60),	1ft5A (1.60),	1g12A (1.60),	1ii5A (1.60),	1in4A (1.60),	1iupA (1.60),
1jhjA (1.60),	1juhA (1.60),	1kqfA (1.60),	1kqfC (1.60),	1lqvA (1.60),	1n08A (1.60),
1nc5A (1.60),	1po5A (1.60),	1q7eA (1.60),	1qgiA (1.60),	1ro2A (1.60),	1rylA (1.60),
1s1dA (1.60),	1s9rA (1.60),	1t1vA (1.60),	1t92A (1.60),	1ukka (1.60),	1v7wA (1.60),
1wnyA (1.60),	1yb3A (1.60),	1yt3A (1.60),	2apjA (1.60),	2b4hA (1.60),	2bo9B (1.60),
2cxaA (1.60),	2dg5A (1.60),	2dvmA (1.60),	2dwuA (1.60),	2eb4A (1.60),	2ev1A (1.60),
2f23A (1.60),	2fctA (1.60),	2g30A (1.60),	2hlyA (1.60),	2icuA (1.60),	2ilkA (1.60),
2j43A (1.60),	2jdcA (1.60),	2je6B (1.60),	2nw8A (1.60),	2o2kA (1.60),	2okfA (1.60),
2oqzA (1.60),	2ou5A (1.60),	2p6wA (1.60),	2qb7A (1.60),	2qnkA (1.60),	2qrlA (1.60),
2rdgA (1.60),	2rk9A (1.60),	2wawA (1.60),	2wsdA (1.60),	2wzoA (1.60),	2xwsA (1.60),
2yvtA (1.60),	2ywiA (1.60),	2yzyA (1.60),	2z6oA (1.60),	2zblA (1.60),	2zq0A (1.60),
3a9sA (1.60),	3bmwA (1.60),	3c8wA (1.60),	3ctzA (1.60),	3d9nA (1.60),	3dffA (1.60),
3dg6A (1.60),	3e48A (1.60),	3fpwA (1.60),	3gaeA (1.60),	3gkrA (1.60),	3gybA (1.60),
3gzaA (1.60),	3h75A (1.60),	3hidA (1.60),	3irvA (1.60),	3l81A (1.60),	3lkmA (1.60),
3lyhA (1.60),	3m11A (1.60),	3n10A (1.60),	3nnbA (1.60),	3npdA (1.60),	3nvwc (1.60),

Supplementary Material

S. Das et al.

3o94A (1.60),	3p2cA (1.60),	3pftA (1.60),	3q7rA (1.60),	3qx1A (1.60),	3r8jA (1.60),
3suvA (1.60),	3t7hA (1.60),	3tfjA (1.60),	3uxfA (1.60),	3uzqB (1.60),	3vwlA (1.60),
3vwmA (1.60),	3w0kA (1.60),	3zw5A (1.60),	4ao6A (1.60),	4dziA (1.60),	4eqaC (1.60),
4h7uA (1.60),	4h9jA (1.60),	4h9kA (1.60),	4hizA (1.60),	4hzzA (1.60),	4j0eA (1.60),
4j6oA (1.60),	4k3fA (1.60),	4lc3A (1.60),	4lk9A (1.60),	4lzgA (1.60),	4mlqA (1.60),
4m65A (1.60),	4maeA (1.60),	4moqA (1.60),	4mupA (1.60),	4nsnA (1.60),	4o7hA (1.60),
2e8bA (1.61),	3hdoA (1.61),	3qlnA (1.61),	3rf3A (1.61),	4bjtA (1.61),	4l7tA (1.61),
1vjfA (1.62),	1wolA (1.62),	2w6lA (1.62),	2zkmX (1.62),	3gd0A (1.62),	3go9A (1.62),
3mltA (1.62),	4a5sA (1.62),	4b4cA (1.62),	4bdxX (1.62),	4gbmX (1.62),	4h0cA (1.62),
4hz4A (1.62),	4kq7A (1.62),	1gnyA (1.63),	1k77A (1.63),	2wz1A (1.63),	2zb4A (1.63),
3uuwA (1.63),	3zgoB (1.63),	3zpiB (1.63),	4limA (1.63),	1g8kA (1.64),	2x1dA (1.64),
3f0pA (1.64),	3qeeA (1.64),	4c3sA (1.64),	4hikA (1.64),	4ho4A (1.64),	4ipyB (1.64),
4lswA (1.64),	1ezwA (1.65),	1gxra (1.65),	1hxrA (1.65),	1ogoX (1.65),	1u5hA (1.65),
1ubpC (1.65),	1vkfA (1.65),	2bbaA (1.65),	2fmpA (1.65),	2j66A (1.65),	2nooA (1.65),
2nrkA (1.65),	2qeeA (1.65),	2qruA (1.65),	2uz1A (1.65),	2vvpA (1.65),	2yleA (1.65),
3clsC (1.65),	3dalA (1.65),	3dczA (1.65),	3esmA (1.65),	3ff1A (1.65),	3h05A (1.65),
3kgwA (1.65),	3ksnA (1.65),	3oisB (1.65),	3oj0A (1.65),	3popA (1.65),	3q7mA (1.65),
3qhoA (1.65),	3vwrA (1.65),	3w5fA (1.65),	3w90A (1.65),	3zeuB (1.65),	3zk4A (1.65),
4ak4A (1.65),	4bz7A (1.65),	4eg0A (1.65),	4eojD (1.65),	4exqA (1.65),	4j8lA (1.65),
4jcmA (1.65),	4jqpA (1.65),	4ka5A (1.65),	4md5A (1.65),	4md5B (1.65),	4mmwA (1.65),
4mnrA (1.65),	4mooA (1.65),	4murA (1.65),	4mzaA (1.65),	4nq1A (1.65),	4nv0A (1.65),
4oenA (1.65),	4onwA (1.650),	1e4cP (1.66),	2ft0A (1.66),	2isbA (1.66),	2oebA (1.66),
3kbrA (1.66),	4h60A (1.66),	4l2cA (1.66),	2aj7A (1.67),	2bjfA (1.67),	2jlqA (1.67),
2r7gA (1.67),	2wfpA (1.67),	3atsA (1.67),	3dr2A (1.67),	3ikbA (1.67),	4d9bA (1.67),
4jmdA (1.67),	4k4kA (1.67),	4k8wA (1.67),	4m98A (1.67),	4mo4A (1.67),	2a15A (1.68),
2xfgB (1.68),	2z14A (1.68),	3p2eA (1.68),	4ae2A (1.68),	4bc3A (1.68),	4h3vB (1.68),
4kk7A (1.68),	4lq6A (1.68),	1jy1A (1.69),	2hftA (1.69),	2ykyA (1.69),	3gf6A (1.69),
3hsuA (1.69),	3vpzA (1.69),	4acyA (1.69),	4awuA (1.69),	4fkzA (1.69),	4gzkA (1.69),
4j7hA (1.69),	4mlhA (1.69),	1ctfA (1.70),	1eexG (1.70),	1erzA (1.70),	1f5nA (1.70),
1h6fA (1.70),	1hplA (1.70),	1kaeA (1.70),	1kzqA (1.70),	1lvwA (1.70),	1mtYG (1.70),
1nepA (1.70),	1njhA (1.70),	1nrjA (1.70),	1of1A (1.70),	1ogqA (1.70),	1oqlA (1.70),
1osyA (1.70),	1pbyA (1.70),	1pbyC (1.70),	1rh6A (1.70),	1rwjA (1.70),	1sg6A (1.70),
1t0bA (1.70),	1tuoA (1.70),	1uekA (1.70),	1wy2A (1.70),	1xcrA (1.70),	1xkpB (1.70),
1yrkA (1.70),	1yumA (1.70),	1z9lA (1.70),	2a14A (1.70),	2a9iA (1.70),	2b8mA (1.70),
2bjnA (1.70),	2c2qA (1.70),	2cwrA (1.70),	2d4pA (1.70),	2d80A (1.70),	2dvtA (1.70),
2e8eA (1.70),	2fphX (1.70),	2fsuA (1.70),	2g0wA (1.70),	2g40A (1.70),	2i71A (1.70),

Supplementary Material

S. Das et al.

2idlA (1.70),	2j6aA (1.70),	2je8A (1.70),	2qk1A (1.70),	2r4fA (1.70),	2r85A (1.70),
2uvpA (1.70),	2wyaA (1.70),	2xskA (1.70),	2xvyA (1.70),	2yavA (1.70),	2z84A (1.70),
2z9wA (1.70),	2zwaA (1.70),	2zyzB (1.70),	3aa0A (1.70),	3aa0B (1.70),	3biyA (1.70),
3bm3A (1.70),	3bofA (1.70),	3byqA (1.70),	3cinA (1.70),	3d33A (1.70),	3ecbB (1.70),
3eziA (1.70),	3h5lA (1.70),	3hq1A (1.70),	3ilsA (1.70),	3it4A (1.70),	3ixsA (1.70),
3kg9A (1.70),	3lewA (1.70),	3mhxA (1.70),	3n1eA (1.70),	3nuqA (1.70),	3o0yA (1.70),
3p2tA (1.70),	3pbqA (1.70),	3pg6A (1.70),	3q18A (1.70),	3ql6A (1.70),	3rlsA (1.70),
3rt4B (1.70),	3sgH (1.70),	3sgwA (1.70),	3t1fA (1.70),	3tawA (1.70),	3twlA (1.70),
3uq8A (1.70),	3v9wA (1.70),	3vwqA (1.70),	3w4rA (1.70),	3w9sA (1.70),	3wa7A (1.70),
3wd0A (1.70),	3wdpP (1.70),	3webA (1.70),	4anoA (1.70),	4bhrA (1.70),	4blpA (1.70),
4bolA (1.70),	4c43A (1.70),	4c8lA (1.70),	4c97A (1.70),	4cl3D (1.70),	4cllA (1.70),
4dl8A (1.70),	4esqA (1.70),	4g22A (1.70),	4gs1A (1.70),	4gv5A (1.70),	4h5sB (1.70),
4hvfA (1.70),	4hw6A (1.70),	4hwsA (1.70),	4inzA (1.70),	4iy7A (1.70),	4jclA (1.70),
4jhlA (1.70),	4k5uA (1.70),	4ka6A (1.70),	4kieA (1.70),	4kmdA (1.70),	4kmfA (1.70),
4l83A (1.70),	4l8eA (1.70),	4lb0A (1.70),	4ls3A (1.70),	4m68A (1.70),	4mg4A (1.70),
4mo3M (1.70),	4mvnB (1.70),	4mwiA (1.70),	4n0vA (1.70),	4namA (1.70),	4ndaA (1.70),
4o6yA (1.70),	4oj0A (1.70),	4pobA (1.70),	2r4gA (1.71),	3sonA (1.71),	3zuiA (1.71),
4el6A (1.71),	4k2mA (1.71),	1rwrA (1.72),	2dg1A (1.72),	3vygD (1.72),	4lv8A (1.72),
4m9pA (1.72),	4mlzA (1.72),	4natA (1.72),	2i7gA (1.73),	2pbkA (1.73),	3logA (1.73),
4df3B (1.73),	1jyaA (1.74),	1px5A (1.74),	3bsoA (1.74),	3fw2A (1.74),	3v7bA (1.74),
4a35A (1.74),	4nf2A (1.74),	1d2nA (1.75),	1dqpA (1.75),	1f32A (1.75),	1l1lA (1.75),
1nszA (1.75),	1o66A (1.75),	1oi2A (1.75),	1s5dA (1.75),	1u7lA (1.75),	1w2wA (1.75),
1w2wB (1.75),	1w99A (1.75),	1wd3A (1.75),	1xe7A (1.75),	1y0kA (1.75),	2a9sA (1.75),
2c0gA (1.75),	2o62A (1.75),	2oj5A (1.75),	2p26A (1.75),	2ps1A (1.75),	2qcuA (1.75),
2quoA (1.75),	2uxqA (1.75),	2vm9A (1.75),	2vrsA (1.75),	2y8nA (1.75),	2yncC (1.75),
3f47A (1.75),	3ffrA (1.75),	3fncA (1.75),	3fotA (1.75),	3hn0A (1.75),	3k40A (1.75),
3mxmB (1.75),	3nswA (1.75),	3oamA (1.75),	3oepA (1.75),	3qtcA (1.75),	3qy3A (1.75),
3rnlA (1.75),	3s9jA (1.75),	3tcvA (1.75),	3uoaB (1.75),	3vfiA (1.75),	3vnrA (1.75),
3vu9A (1.75),	4a0dA (1.75),	4bqhA (1.75),	4br5A (1.75),	4c2lA (1.75),	4ccwA (1.75),
4cg8A (1.75),	4cnmA (1.75),	4fusA (1.75),	4fvqA (1.75),	4g26A (1.75),	4g32A (1.75),
4gf0A (1.75),	4ggjA (1.75),	4i79A (1.75),	4imeA (1.75),	4ipiA (1.75),	4ivgA (1.75),
4kc7A (1.75),	4lanA (1.75),	4mpqA (1.75),	4mubA (1.75),	4n0pA (1.75),	4n34A (1.75),
1n93X (1.76),	1te2A (1.76),	2p3pA (1.76),	2rdmA (1.76),	3bwuF (1.76),	3gykA (1.76),
3lidA (1.76),	3lxzA (1.76),	3or1A (1.76),	3r6aA (1.76),	3vs8A (1.76),	3vv3A (1.76),
4egoB (1.76),	4hobA (1.76),	4jwjA (1.76),	4m88A (1.76),	1r7aA (1.77),	2qtcA (1.77),
3kt7A (1.77),	3mkcA (1.77),	4jisA (1.77),	2hu9A (1.78),	2imrA (1.78),	2qw5A (1.78),

Supplementary Material

S. Das et al.

3a6fA (1.78),	3ppmA (1.78),	3zlbA (1.78),	4a20A (1.78),	4hatC (1.78),	4ic3A (1.78),
4mdcA (1.78),	4nhdA (1.78),	4o5aA (1.78),	2fg0A (1.79),	2jk9A (1.79),	2nt0A (1.79),
3ft1A (1.79),	3ilvA (1.79),	3lvuA (1.79),	3mh9A (1.79),	3qfwA (1.79),	3rkcA (1.79),
3s9zA (1.79),	4cftB (1.79),	4f67A (1.79),	4ja2A (1.79),	4l3hA (1.79),	4l3hF (1.79),
4me3A (1.79),	4n6qA (1.79),	1a73A (1.80),	1a9xA (1.80),	1aylA (1.80),	1bu8A (1.80),
1eokA (1.80),	1flmA (1.80),	1h0hB (1.80),	1h4rA (1.80),	1h72C (1.80),	1kzfA (1.80),
1lbvA (1.80),	1m0wA (1.80),	1q9uA (1.80),	1qwrA (1.80),	1ryiB (1.80),	1rzhH (1.80),
1rzhM (1.80),	1tifA (1.80),	1vcaA (1.80),	1vdkA (1.80),	1ve2A (1.80),	1vknB (1.80),
1vprA (1.80),	1vptA (1.80),	1w96A (1.80),	1wlgA (1.80),	1xauA (1.80),	1xffaA (1.80),
1xkiA (1.80),	1xttA (1.80),	1xuoA (1.80),	1yacA (1.80),	1yj7A (1.80),	1zjcA (1.80),
2alhA (1.80),	2ag4A (1.80),	2axoA (1.80),	2blyA (1.80),	2b4vA (1.80),	2de6A (1.80),
2hltA (1.80),	2h98A (1.80),	2hfnA (1.80),	2if6A (1.80),	2inuA (1.80),	2j0aA (1.80),
2j7qA (1.80),	2je3A (1.80),	2jhnB (1.80),	2ns9A (1.80),	2o70A (1.80),	2pc8A (1.80),
2q66A (1.80),	2qisA (1.80),	2qsiA (1.80),	2r6oA (1.80),	2r7dA (1.80),	2rinA (1.80),
2sicI (1.80),	2tgiA (1.80),	2uvjA (1.80),	2uwaA (1.80),	2vhjA (1.80),	2vsmA (1.80),
2waoA (1.80),	2wteA (1.80),	2x5fA (1.80),	2xvsA (1.80),	2yjaA (1.80),	2ypvA (1.80),
2zzjA (1.80),	3ag3A (1.80),	3ag3B (1.80),	3ag3C (1.80),	3avrA (1.80),	3bamA (1.80),
3bl9A (1.80),	3bqkA (1.80),	3c0fB (1.80),	3c5nA (1.80),	3cl5A (1.80),	3d34A (1.80),
3danA (1.80),	3e05B (1.80),	3e96A (1.80),	3egaA (1.80),	3erbA (1.80),	3ew8A (1.80),
3exnA (1.80),	3f95A (1.80),	3fgrB (1.80),	3fkrA (1.80),	3fo8D (1.80),	3gn6A (1.80),
3greA (1.80),	3h8tA (1.80),	3irbA (1.80),	3jsrA (1.80),	3ju2A (1.80),	3k3vA (1.80),
3kbyA (1.80),	3lfuA (1.80),	3ligA (1.80),	3lm4A (1.80),	3ly1A (1.80),	3m6nA (1.80),
3mvuA (1.80),	3mwcA (1.80),	3ngqA (1.80),	3nrsA (1.80),	3o0qA (1.80),	3oc9A (1.80),
3ohgA (1.80),	3oskA (1.80),	3pbfA (1.80),	3pijB (1.80),	3pywA (1.80),	3qmqA (1.80),
3qsqA (1.80),	3qy9A (1.80),	3rppA (1.80),	3s8gA (1.80),	3sx6A (1.80),	3t9aA (1.80),
3tbdA (1.80),	3tiwA (1.80),	3tvkA (1.80),	3uf6A (1.80),	3ufbA (1.80),	3vbcA (1.80),
3vpbA (1.80),	3vxsD (1.80),	3w0eA (1.80),	3w5nA (1.80),	3w8zA (1.80),	3waqA (1.80),
3wbdA (1.80),	3wbpA (1.80),	3whtA (1.80),	3zn2A (1.80),	3zssA (1.80),	4adiA (1.80),
4ajsA (1.80),	4aq4A (1.80),	4b2nA (1.80),	4c2oA (1.80),	4c7gA (1.80),	4c9lB (1.80),
4c9pA (1.80),	4cheA (1.80),	4cm2A (1.80),	4e0aA (1.80),	4fdzA (1.80),	4fo7A (1.80),
4g68A (1.80),	4gv4A (1.80),	4gxbA (1.80),	4h04A (1.80),	4h4lA (1.80),	4h7yA (1.80),
4i0nA (1.80),	4i4tA (1.80),	4ia6A (1.80),	4ibmA (1.80),	4iyoB (1.80),	4j4hA (1.80),
4jj0A (1.80),	4jj6A (1.80),	4jkzA (1.80),	4jyuH (1.80),	4klvA (1.80),	4k90A (1.80),
4kb1B (1.80),	4kdeA (1.80),	4knuA (1.80),	4krxA (1.80),	4kyvA (1.80),	4l00A (1.80),
4lfyA (1.80),	4lhdA (1.80),	4ljrA (1.80),	4lw2A (1.80),	4mljC (1.80),	4mlrA (1.80),
4mf7A (1.80),	4mixA (1.80),	4mnA (1.80),	4mwjA (1.80),	4mw1A (1.80),	4mynA (1.80),

Supplementary Material

S. Das et al.

4mzeA (1.80),	4n4jA (1.80),	4n6hA (1.80),	4n8mA (1.80),	4ns0A (1.80),	4ofkA (1.80),
4ovyA (1.80),	1c96A (1.81),	2x1kA (1.81),	3lurA (1.81),	3txsA (1.81),	4gwrA (1.81),
4kd1B (1.81),	4lcqA (1.81),	4mb6A (1.81),	1h8pA (1.82),	1zsqA (1.82),	2pwwA (1.82),
3a5pA (1.82),	3kepA (1.82),	3o3mA (1.82),	3o3mB (1.82),	4ca7A (1.82),	4gxtA (1.82),
4lgnA (1.82),	4ln9A (1.82),	4m9dA (1.82),	1cl1A (1.83),	1hq0A (1.83),	1xwvA (1.83),
1yqeA (1.83),	2bsjA (1.83),	2zyhB (1.83),	3ph9A (1.83),	3rhtA (1.83),	4ilyA (1.83),
4o1kA (1.83),	1ss4B (1.84),	2wuxA (1.84),	3g85A (1.84),	3uanA (1.84),	3zjbA (1.84),
4bzcA (1.84),	4lzaA (1.84),	1b25A (1.85),	1t61A (1.85),	1vkhA (1.85),	2auwA (1.85),
2cwyA (1.85),	2cxkB (1.85),	2e4mC (1.85),	2ez2A (1.85),	2gwnA (1.85),	2hboA (1.85),
2i2cA (1.85),	2in3A (1.85),	2od6A (1.85),	2w2kA (1.85),	2y7pA (1.85),	2yyyA (1.85),
2zfxA (1.85),	2zuvA (1.85),	3cjpA (1.85),	3el6A (1.85),	3f2kA (1.85),	3fxqA (1.85),
3g3tA (1.85),	3gycA (1.85),	3iukA (1.85),	3kogA (1.85),	3m8uA (1.85),	3ndnB (1.85),
3njeA (1.85),	3nn1A (1.85),	3p61A (1.85),	3q2uA (1.85),	3qxfA (1.85),	3sxuA (1.85),
4ba0A (1.85),	4bluA (1.85),	4c7dA (1.85),	4dlqA (1.85),	4egdA (1.85),	4fwwA (1.85),
4g2uA (1.85),	4gmfA (1.85),	4hpmB (1.85),	4imcA (1.85),	4iuwA (1.85),	4jnpA (1.85),
4kxqA (1.85),	4mdrA (1.85),	4mgeA (1.85),	4n36A (1.85),	4nmxB (1.85),	4ohcA (1.85),
4oieA (1.85),	4ol9A (1.85),	4onzA (1.85),	1jb7B (1.86),	2c0nA (1.86),	2wjnC (1.86),
3enuA (1.86),	3qi7A (1.86),	4ksnA (1.86),	4nbiB (1.86),	4nf8B (1.86),	4o6aA (1.86),
3d3sA (1.87),	3o6cA (1.87),	3sumA (1.87),	3zhgA (1.87),	3zhgD (1.87),	4hsqA (1.87),
4hzdA (1.87),	1y6zA (1.88),	3n0wA (1.88),	3qp9A (1.88),	4fhaA (1.88),	4l3kA (1.88),
4ncrB (1.88),	3dvoA (1.89),	3k7xA (1.89),	3no3A (1.89),	3zbuA (1.89),	4bwcB (1.89),
4mu9A (1.89),	1a81A (1.90),	1ayoA (1.90),	1dzfA (1.90),	1g3kA (1.90),	1guiA (1.90),
1hn0A (1.90),	1jdwA (1.90),	1kczA (1.90),	1lshA (1.90),	1me8A (1.90),	1nnwA (1.90),
1oaoC (1.90),	1obbA (1.90),	1q16A (1.90),	1qqp1 (1.90),	1qqp2 (1.90),	1qr0A (1.90),
1rqbA (1.90),	1sfpA (1.90),	1sq9A (1.90),	1u0sA (1.90),	1ua4A (1.90),	1ut7A (1.90),
1uvjA (1.90),	1ux6A (1.90),	1wteA (1.90),	1wzoA (1.90),	1y60A (1.90),	1ynfA (1.90),
1z40A (1.90),	1zx8A (1.90),	2bkrA (1.90),	2c51C (1.90),	2dp9A (1.90),	2fefA (1.90),
2ftxA (1.90),	2g5gX (1.90),	2hy7A (1.90),	2ig8A (1.90),	2j4dA (1.90),	2jdiD (1.90),
2jdiH (1.90),	2odaA (1.90),	2odfA (1.90),	2opwA (1.90),	2osoA (1.90),	2p0aA (1.90),
2p0wA (1.90),	2q88A (1.90),	2qjvA (1.90),	2qtsA (1.90),	2qzqA (1.90),	2wmkA (1.90),
2wvxA (1.90),	2wyhA (1.90),	2y6uA (1.90),	2ymoA (1.90),	3arcU (1.90),	3b9oA (1.90),
3bjkB (1.90),	3c8nA (1.90),	3c9fA (1.90),	3cptA (1.90),	3cq1A (1.90),	3cx5C (1.90),
3d30A (1.90),	3d3mB (1.90),	3d85C (1.90),	3dmlA (1.90),	3e99A (1.90),	3ecfA (1.90),
3er6A (1.90),	3fidA (1.90),	3fvyA (1.90),	3g40A (1.90),	3g8yA (1.90),	3hxlA (1.90),
3ijwA (1.90),	3isrB (1.90),	3k1tA (1.90),	3k8uA (1.90),	3ks9A (1.90),	3lywA (1.90),
3m6jA (1.90),	3mdpA (1.90),	3ntxA (1.90),	3o10A (1.90),	3o83A (1.90),	3onpA (1.90),

Supplementary Material

S. Das et al.

3picA (1.90),	3qfmA (1.90),	3rhzb (1.90),	3smpA (1.90),	3sy6A (1.90),	3tmgA (1.90),
3twdA (1.90),	3ty1A (1.90),	3typA (1.90),	3vk6A (1.90),	3vptA (1.90),	3vv5A (1.90),
3wbaA (1.90),	3we0A (1.90),	3wisA (1.90),	3wkpA (1.90),	3znnB (1.90),	4afkA (1.90),
4amwA (1.90),	4at7B (1.90),	4b0bB (1.90),	4blrB (1.90),	4bq2A (1.90),	4cj6A (1.90),
4cotA (1.90),	4dplA (1.90),	4dx5A (1.90),	4ehxA (1.90),	4eiuA (1.90),	4h0aA (1.90),
4hjlA (1.90),	4hnrA (1.90),	4i0oA (1.90),	4i2aA (1.90),	4it6A (1.90),	4iv8A (1.90),
4j6cA (1.90),	4j6pB (1.90),	4j6vA (1.90),	4jdeB (1.90),	4jkoA (1.90),	4jmqA (1.90),
4k2dA (1.90),	4k5sA (1.90),	4k8lA (1.90),	4kcaA (1.90),	4l0lA (1.90),	4l0fA (1.90),
4l4vC (1.90),	4l4vE (1.90),	4lnlA (1.90),	4mleA (1.90),	4mokA (1.90),	4mruA (1.90),
4ms4A (1.90),	4ms4B (1.90),	4myaA (1.90),	4mymA (1.90),	4n49A (1.90),	4njhA (1.90),
4nk3A (1.90),	4nktA (1.90),	4nmyA (1.90),	4nn5C (1.90),	4ntcA (1.90),	4of8A (1.90),
2gf6A (1.91),	3gmiA (1.91),	3jtxA (1.91),	4ca6B (1.91),	4ftdA (1.91),	4je5A (1.91),
4k55A (1.91),	4nldA (1.91),	2rdeA (1.92),	3ec3A (1.92),	3k59A (1.92),	3zbtA (1.92),
4cmmA (1.92),	4lndA (1.92),	3plvA (1.93),	3vwwA (1.93),	3weuA (1.93),	4fdtA (1.93),
4fhrA (1.93),	4ifsA (1.93),	4koaA (1.93),	4kv9A (1.93),	1jztA (1.94),	1vgjA (1.94),
1z5yE (1.94),	2cxiA (1.94),	2o4vA (1.94),	2q30A (1.94),	3kwlA (1.94),	4cfsA (1.94),
4gunA (1.94),	4ioyX (1.94),	4jgeA (1.94),	4k98A (1.94),	1b12A (1.95),	1ekxA (1.95),
1g5hA (1.95),	1g8lA (1.95),	1i7qA (1.95),	1i7qB (1.95),	1itvA (1.95),	1ogdA (1.95),
1oyjA (1.95),	1qhdA (1.95),	1rlhA (1.95),	1t6sA (1.95),	1tulA (1.95),	1uw4A (1.95),
1uwsB (1.95),	2o34A (1.95),	2o95A (1.95),	2q8pA (1.95),	2qsqA (1.95),	2uuuA (1.95),
2uw1A (1.95),	2vliB (1.95),	2xdgA (1.95),	3bzmA (1.95),	3ejkA (1.95),	3hl1A (1.95),
3hrlA (1.95),	3huuA (1.95),	3m4iA (1.95),	3otnA (1.95),	3p8aA (1.95),	3rd7A (1.95),
3v6oA (1.95),	3wmrA (1.95),	3zgjA (1.95),	3zj5B (1.95),	3zntA (1.95),	4awnA (1.95),
4b2zA (1.95),	4e5xG (1.95),	4ffjA (1.95),	4gdiA (1.95),	4gj0A (1.95),	4h4zA (1.95),
4i4cB (1.95),	4ip2B (1.95),	4kctA (1.95),	4kunA (1.95),	4l0qA (1.95),	4l9rA (1.95),
4ljyA (1.95),	4meaA (1.95),	4mf9A (1.95),	4mzwA (1.95),	4n4gA (1.95),	4nheB (1.95),
4o8yA (1.95),	4o8yB (1.95),	1k7wA (1.96),	2yicA (1.96),	2z5bB (1.96),	4i3pA (1.96),
4i5qA (1.96),	1t0hB (1.97),	2d7vA (1.97),	3cvqB (1.97),	3hhsA (1.97),	3w36A (1.97),
3wbeA (1.97),	4b0aA (1.97),	4ccdA (1.97),	4icsA (1.97),	4id3B (1.97),	4ig5B (1.97),
4k6xA (1.97),	4m5dA (1.97),	2f9iB (1.98),	3ktoA (1.98),	3vniA (1.98),	4aziA (1.98),
4c9oA (1.98),	4imhA (1.98),	4ko8B (1.98),	4ktyA (1.98),	4la7B (1.98),	4ml1A (1.98),
4nnrA (1.98),	3bwsA (1.99),	3cneA (1.99),	3gnjA (1.99),	3ipfA (1.99),	3l15A (1.99),
3w9aA (1.99),	3zpxA (1.99),	4c2pA (1.99),	4egcB (1.99),	4fwgA (1.99),	4kf4A (1.99),
4knvA (1.99),	4mcuA (1.99),	4mjfA (1.99),	4nckA (1.99),	1aazA (2.00),	1aocA (2.00),
1bf2A (2.00),	1byfA (2.00),	1cvrA (2.00),	1czyA (2.00),	1d3yA (2.00),	1dp4A (2.00),
1eziB (2.00),	1f3vA (2.00),	1gsaA (2.00),	1h6kA (2.00),	1j1tA (2.00),	1k8kC (2.00),

Supplementary Material

S. Das et al.

1kcmA (2.00),	1kmoA (2.00),	1lkiA (2.00),	1n2zA (2.00),	1nbaA (2.00),	1ofdB (2.00),
1pswA (2.00),	1q8cA (2.00),	1reqB (2.00),	1sraA (2.00),	1t9fA (2.00),	1twuA (2.00),
1usrA (2.00),	1v4aA (2.00),	1vkyA (2.00),	1wdjA (2.00),	1whsA (2.00),	1wiwA (2.00),
1wzuA (2.00),	1xo0A (2.00),	1xv2A (2.00),	1ygaA (2.00),	1yzvA (2.00),	1z2wB (2.00),
1zc3B (2.00),	1zmoA (2.00),	2a67A (2.00),	2atzA (2.00),	2az4A (2.00),	2bb6A (2.00),
2bm5A (2.00),	2d0oC (2.00),	2djfA (2.00),	2e56A (2.00),	2e6mA (2.00),	2eceA (2.00),
2ervA (2.00),	2gakA (2.00),	2hiqA (2.00),	2i4lA (2.00),	2i9cA (2.00),	2ichA (2.00),
2ilrA (2.00),	2j9uB (2.00),	2oo3A (2.00),	2p08A (2.00),	2pa4B (2.00),	2q7sA (2.00),
2q7xA (2.00),	2qdjA (2.00),	2qkdA (2.00),	2qp2A (2.00),	2qv6A (2.00),	2r9lA (2.00),
2uy1A (2.00),	2v9kA (2.00),	2vq3B (2.00),	2vtwA (2.00),	2wgkA (2.00),	2x2uA (2.00),
2xsgA (2.00),	2xt2A (2.00),	2xw7B (2.00),	2ykfA (2.00),	2zayA (2.00),	3agcA (2.00),
3akjA (2.00),	3aonB (2.00),	3aqeA (2.00),	3axgA (2.00),	3bk5A (2.00),	3bm1A (2.00),
3c2qA (2.00),	3cexA (2.00),	3cnbA (2.00),	3dcmX (2.00),	3dxeA (2.00),	3e6lA (2.00),
3etoA (2.00),	3f4lA (2.00),	3f62A (2.00),	3g5bA (2.00),	3ggnA (2.00),	3gwqA (2.00),
3hc7A (2.00),	3hwpA (2.00),	3hz7A (2.00),	3i16A (2.00),	3i76A (2.00),	3igfA (2.00),
3ij6A (2.00),	3juiA (2.00),	3k6oA (2.00),	3ksjA (2.00),	3l60A (2.00),	3lgdA (2.00),
3l13B (2.00),	3llkA (2.00),	3lp5A (2.00),	3luyA (2.00),	3mkhA (2.00),	3n75A (2.00),
3na6A (2.00),	3nvxA (2.00),	3ouvA (2.00),	3qbtB (2.00),	3qtaA (2.00),	3rlbA (2.00),
3t6sA (2.00),	3tmkA (2.00),	3ttgA (2.00),	3tw0A (2.00),	3ud1A (2.00),	3umhA (2.00),
3v42A (2.00),	3vsnA (2.00),	3vvhB (2.00),	3vzbA (2.00),	3w0pA (2.00),	3w6uA (2.00),
3w9uA (2.00),	3wmdA (2.00),	3zinA (2.00),	4a5uA (2.00),	4acfA (2.00),	4bf7A (2.00),
4bprA (2.00),	4bqkA (2.00),	4br6A (2.00),	4bw1A (2.00),	4bw1C (2.00),	4c0zA (2.00),
4c1mD (2.00),	4c3xA (2.00),	4cooA (2.00),	4egtA (2.00),	4ei0A (2.00),	4evaA (2.00),
4fw9A (2.00),	4fyuA (2.00),	4g7sA (2.00),	4g7sB (2.00),	4gfuA (2.00),	4giwA (2.00),
4gqpH (2.00),	4grmC (2.00),	4h7nA (2.00),	4h8hA (2.00),	4hk1A (2.00),	4hslA (2.00),
4hu8A (2.00),	4hvcA (2.00),	4hw0A (2.00),	4i8bA (2.00),	4ic9A (2.00),	4iftA (2.00),
4ihqA (2.00),	4ijrA (2.00),	4ik1A (2.00),	4ilfA (2.00),	4je6A (2.00),	4jwmA (2.00),
4k08A (2.00),	4kluA (2.00),	4k35A (2.00),	4k70A (2.00),	4k9yA (2.00),	4kp7A (2.00),
4l4qA (2.00),	4l68A (2.00),	4l82A (2.00),	4l9aA (2.00),	4losA (2.00),	4lu3A (2.00),
4m0eB (2.00),	4m7hA (2.00),	4mdwA (2.00),	4mebA (2.00),	4mfgA (2.00),	4mjaA (2.00),
4mjja (2.00),	4mo2A (2.00),	4mqyA (2.00),	4n1xA (2.00),	4n4iA (2.00),	4n7tB (2.00),
4ncjA (2.00),	4nd8A (2.00),	4nlrA (2.00),	4nlsA (2.00),	4nnbA (2.00),	4np6A (2.00),
4ou9A (2.00),	4ovdA (2.00),	4p0mA (2.00),	1vjgA (2.01),	2aplA (2.01),	2h5nA (2.01),
2xqxA (2.01),	2zz8A (2.01),	3di5A (2.01),	3l6iA (2.01),	3uafA (2.01),	3vqrB (2.01),
3w1wA (2.01),	4gn2A (2.01),	4m9gA (2.01),	2ii0A (2.02),	4m7oA (2.02),	3t7kA (2.03),
4fwdA (2.03),	4g87A (2.03),	4h3tA (2.03),	4jo1H (2.03),	4jo1L (2.03),	4lzbB (2.03),

Supplementary Material

S. Das et al.

4mcwA (2.03),	2bbdA (2.04),	2pkdA (2.04),	2q07A (2.04),	3d2oA (2.04),	3mpkA (2.04),
3qjgA (2.04),	4abxA (2.04),	4ccyA (2.04),	4gxzA (2.04),	4irvA (2.04),	4j5tA (2.04),
4jwhA (2.04),	4mb7A (2.04),	1qi9B (2.05),	1v72A (2.05),	1wlfa (2.05),	1zghA (2.05),
2hxrA (2.05),	2iafA (2.05),	2itbA (2.05),	2ql3A (2.05),	2rgyA (2.05),	2vs7A (2.05),
3bjqA (2.05),	3fdsC (2.05),	3hbzA (2.05),	3liuA (2.05),	3m9vA (2.05),	3oapA (2.05),
3vldA (2.05),	4b8rA (2.05),	4bq4A (2.05),	4djbA (2.05),	4esjA (2.05),	4f0qA (2.05),
4g97A (2.05),	4joiC (2.05),	4jsdA (2.05),	4k91A (2.05),	3er9B (2.06),	3lczA (2.06),
3p51A (2.06),	4lo0A (2.06),	4lwaA (2.06),	1oltA (2.07),	3tojB (2.07),	3w8sA (2.07),
3ziuA (2.07),	3clkA (2.08),	3dwaA (2.08),	3efzA (2.08),	3oqqA (2.08),	3vhlA (2.08),
4f9cB (2.08),	4m4pA (2.08),	3mdnA (2.09),	3o5tA (2.09),	3qxyA (2.09),	3r1xA (2.09),
4azcA (2.09),	4azcC (2.09),	4idoA (2.09),	4igbB (2.09),	4l5oA (2.09),	4l8vA (2.09),
4le7A (2.09),	4o23A (2.09),	1dq3A (2.10),	1ep3B (2.10),	1ez0A (2.10),	1izoA (2.10),
1lf6A (2.10),	1ofuX (2.10),	1omzA (2.10),	1qd6C (2.10),	1t4wA (2.10),	1uhvA (2.10),
1wwbX (2.10),	1wyuA (2.10),	1x3lA (2.10),	1xg8A (2.10),	1xmxA (2.10),	2a5hA (2.10),
2d0bA (2.10),	2d2oA (2.10),	2fiyA (2.10),	2gumA (2.10),	2i5eA (2.10),	2iuyA (2.10),
2nykA (2.10),	2obdA (2.10),	2p53B (2.10),	2ps5A (2.10),	2psbA (2.10),	2q2rA (2.10),
2vsyA (2.10),	2x65A (2.10),	2ynmC (2.10),	2ynmD (2.10),	2ypdB (2.10),	2zbwA (2.10),
3bczA (2.10),	3bn3B (2.10),	3bt4A (2.10),	3ch5B (2.10),	3d37A (2.10),	3dcpA (2.10),
3deeA (2.10),	3dwcA (2.10),	3e7jA (2.10),	3eleA (2.10),	3epoA (2.10),	3ervA (2.10),
3gnfB (2.10),	3guuA (2.10),	3ig5A (2.10),	3iraA (2.10),	3kyfA (2.10),	3kzsA (2.10),
3lqkA (2.10),	3mwdA (2.10),	3onqA (2.10),	3oyoA (2.10),	3p9zA (2.10),	3q3eA (2.10),
3qowA (2.10),	3rsnA (2.10),	3simA (2.10),	3t6oA (2.10),	3tdgA (2.10),	3tdqA (2.10),
3tw8A (2.10),	3v2uA (2.10),	3vatA (2.10),	3vseA (2.10),	3wqlA (2.10),	3zhrA (2.10),
4a9cA (2.10),	4b9wA (2.10),	4c1sA (2.10),	4c7lA (2.10),	4fasA (2.10),	4fdaA (2.10),
4gmoA (2.10),	4gwfA (2.10),	4i9aA (2.10),	4imdA (2.10),	4jlfA (2.10),	4jmfB (2.10),
4kg0A (2.10),	4l36A (2.10),	4le6A (2.10),	4lnaA (2.10),	4lnlA (2.10),	4lubA (2.10),
4m17A (2.10),	4m29A (2.10),	4m3mA (2.10),	4m70B (2.10),	4m7uA (2.10),	4m9cA (2.10),
4mj2A (2.10),	4nluA (2.10),	4nlwA (2.10),	4nrdA (2.10),	4nv1A (2.10),	1ttzA (2.11),
2c36A (2.11),	2q22A (2.11),	3a0oA (2.11),	3f9tB (2.11),	4kbmA (2.11),	4mjbA (2.11),
3eu8A (2.12),	4bg2A (2.13),	4ip5A (2.13),	4jj4A (2.13),	3ip3A (2.14),	3oh2A (2.14),
4c91A (2.14),	4kc5A (2.14),	4kncA (2.14),	1yreA (2.15),	1zybA (2.15),	2a1vA (2.15),
2p0oA (2.15),	2qn6B (2.15),	2r3sA (2.15),	3ih6A (2.15),	3itwA (2.15),	3nppA (2.15),
3pf0A (2.15),	3zo5A (2.15),	4b28A (2.15),	4bzwA (2.15),	4ckmA (2.15),	4h3sA (2.15),
4kx7A (2.15),	4lduA (2.15),	4lj0A (2.15),	4npbA (2.15),	4nyzA (2.15),	4oa8A (2.15),
1h9aA (2.16),	4glkA (2.16),	4l5nA (2.16),	3n40P (2.17),	4amqA (2.17),	4bupA (2.17),
4h7oA (2.17),	4noiA (2.17),	4o9xA (2.17),	2wadA (2.18),	3nngA (2.18),	3ufiA (2.18),

Supplementary Material

S. Das et al.

2yorA (2.19),	3ju7B (2.19),	3okyB (2.19),	3ronA (2.19),	4bx2A (2.19),	4e6wB (2.19),
4l7iA (2.19),	4lgcA (2.19),	4nawF (2.19),	4ncbB (2.19),	4ohsA (2.19),	1b3tA (2.20),
1bouB (2.20),	1c9kA (2.20),	1dmuA (2.20),	1kqnB (2.20),	1n7vA (2.20),	1nqkA (2.20),
1o57A (2.20),	1pxuA (2.20),	1qfjA (2.20),	1tv8A (2.20),	1vpqA (2.20),	1vqoA (2.20),
1vqoB (2.20),	1vqoM (2.20),	1xdyA (2.20),	1yk3A (2.20),	2aamA (2.20),	2cc3A (2.20),
2f5jA (2.20),	2fv2A (2.20),	2gpjA (2.20),	2o3aA (2.20),	2oryA (2.20),	2q0xA (2.20),
2xcmE (2.20),	2ychA (2.20),	2yi9A (2.20),	2z0uA (2.20),	3ai7A (2.20),	3aotA (2.20),
3b7fA (2.20),	3clwA (2.20),	3cwfA (2.20),	3cz8A (2.20),	3ea0A (2.20),	3gc3A (2.20),
3hrdC (2.20),	3hrzA (2.20),	3ib6B (2.20),	3jxpA (2.20),	3kluA (2.20),	3mfbA (2.20),
3n0aA (2.20),	3nctA (2.20),	3o59X (2.20),	3oovA (2.20),	3q7eA (2.20),	3rj2X (2.20),
3tegA (2.20),	3tg9A (2.20),	3vg8A (2.20),	3w3wA (2.20),	3w53A (2.20),	3wg6A (2.20),
3wnuA (2.20),	3zmKA (2.20),	3zpeA (2.20),	4bjhA (2.20),	4bjhB (2.20),	4blgA (2.20),
4dapA (2.20),	4dckA (2.20),	4eilA (2.20),	4fa8A (2.20),	4fybA (2.20),	4hbkA (2.20),
4i2yA (2.20),	4ii2A (2.20),	4j3bA (2.20),	4j3rA (2.20),	4jjfA (2.20),	4jmnA (2.20),
4kfvA (2.20),	4ks1A (2.20),	4lnsA (2.20),	4lq8A (2.20),	4lwoA (2.20),	4lziA (2.20),
4m3hA (2.20),	4m8oA (2.20),	4mayA (2.20),	4mayB (2.20),	4mzuA (2.20),	4n6rB (2.20),
4n6zA (2.20),	4n7bA (2.20),	4ncsA (2.20),	4negA (2.20),	4nloA (2.20),	4nlpA (2.20),
4oteA (2.20),	2gclA (2.21),	2yx0A (2.21),	3brdA (2.21),	3bz6A (2.21),	3zniA (2.21),
4gg2A (2.21),	4gg2B (2.21),	4i9fA (2.21),	4jhoA (2.21),	4jp8A (2.21),	4kjda (2.21),
2w2gA (2.22),	4nvrA (2.22),	2ilyB (2.23),	4oo3A (2.23),	3r4iA (2.24),	3zc8A (2.24),
4gk1A (2.24),	1ei9A (2.25),	lignA (2.25),	2f6hX (2.25),	2fiqA (2.25),	3ttyB (2.25),
3u24A (2.25),	3vpnB (2.25),	3w6gA (2.25),	3zfaA (2.25),	4ghkA (2.25),	4icqA (2.25),
4kg8A (2.25),	4mlzA (2.25),	3dljB (2.26),	3uqcA (2.26),	4lfuA (2.26),	2pw6A (2.27),
3errA (2.27),	4gi0A (2.27),	4kgqA (2.27),	2hd0A (2.28),	3k85A (2.28),	3otdA (2.28),
3we1A (2.28),	4je3A (2.28),	4ju5A (2.28),	1yf3A (2.29),	3cx4A (2.29),	3p8cB (2.29),
3whdA (2.29),	4hopB (2.29),	4kxtA (2.29),	4ldrA (2.29),	1a41A (2.30),	3crkA (2.3),
1cr1A (2.30),	1f35A (2.30),	1j6rA (2.30),	1js8A (2.30),	1l4dB (2.30),	1npeA (2.30),
1p9lB (2.30),	1swvA (2.30),	1xezA (2.30),	2b5iC (2.30),	2d4gB (2.30),	2e9xB (2.30),
2e9xC (2.30),	2ekdA (2.30),	2fp8A (2.30),	2gd9A (2.30),	2ilsA (2.30),	2in5A (2.30),
2j7nA (2.30),	2nutA (2.30),	2o3iA (2.30),	2pifA (2.30),	2v6gA (2.30),	2y9wA (2.30),
3anzB (2.30),	3dl9A (2.30),	3ei3A (2.30),	3emnX (2.30),	3fhkA (2.30),	3flpA (2.30),
3hhfB (2.30),	3ifrA (2.30),	3licA (2.30),	3m2tB (2.30),	3mjgA (2.30),	3n2oA (2.30),
3nvlA (2.30),	3tqoA (2.30),	3ug9A (2.30),	3wckA (2.30),	3zhcA (2.30),	4aezA (2.30),
4clqA (2.30),	4c2rA (2.30),	4d8mA (2.30),	4dh1B (2.30),	4eacA (2.30),	4f2dA (2.30),
4fc5A (2.30),	4g7nA (2.30),	4gouA (2.30),	4hbeA (2.30),	4hceA (2.30),	4hh8A (2.30),
4hwtA (2.30),	4iefA (2.30),	4j41A (2.30),	4j7rA (2.30),	4jegB (2.30),	4jqsA (2.30),

Supplementary Material

S. Das et al.

4jraA (2.30),	4kbbA (2.30),	4kryA (2.30),	4l54A (2.30),	4l79A (2.30),	4ladA (2.30),
4ldyA (2.30),	4lfhD (2.30),	4lfhG (2.30),	4lgvA (2.30),	4lo6B (2.30),	4mg2A (2.30),
4nchA (2.30),	4nlqA (2.30),	4nlvA (2.30),	4nlyA (2.30),	4o24A (2.30),	4ocaA (2.30),
3bjuB (2.31),	3ct9A (2.31),	3e7eA (2.31),	3zetA (2.31),	4je4A (2.31),	4je4B (2.31),
1q87A (2.32),	3d xpA (2.32),	4kw2A (2.32),	4lrzE (2.32),	4lvpA (2.32),	1o65A (2.33),
2o0tB (2.33),	3cihA (2.33),	3k jxA (2.33),	4bw0B (2.33),	4g2aA (2.33),	4ggvA (2.33),
2vakA (2.34),	4lqx A (2.34),	1pn4B (2.35),	1rtwA (2.35),	1xksA (2.35),	2fppB (2.35),
3bywA (2.35),	3n6xA (2.35),	3s1sA (2.35),	3wd7A (2.35),	3z1eA (2.35),	4a01A (2.35),
4bjuA (2.35),	4bv4R (2.35),	4bxjA (2.35),	4ic1A (2.35),	4j18A (2.35),	4oc9A (2.35),
1xr4A (2.37),	3e01B (2.37),	3k9tA (2.37),	4dguA (2.37),	4j56A (2.37),	4lmoA (2.37),
4n01A (2.37),	4djdD (2.38),	4n0gA (2.38),	3wbzA (2.39),	4narA (2.39),	1b35B (2.40),
1p74A (2.40),	1sjiA (2.40),	1t77A (2.40),	2ahmE (2.40),	2bh1X (2.40),	2gekA (2.40),
2i0mA (2.40),	2mprA (2.40),	2o5nA (2.40),	2oajA (2.40),	2ph7A (2.40),	2pt7G (2.40),
2trcP (2.40),	2x24A (2.40),	2x27X (2.40),	2z86A (2.40),	3aqbA (2.40),	3axyB (2.40),
3c0tA (2.40),	3ceaB (2.40),	3cs3A (2.40),	3cx3A (2.40),	3h7mA (2.40),	3hg9A (2.40),
3ic8A (2.40),	3io5A (2.40),	3lkbA (2.40),	3luaA (2.40),	3m62A (2.40),	3mydA (2.40),
4aipB (2.40),	4bltA (2.40),	4ccgA (2.40),	4fwvA (2.40),	4g8aA (2.40),	4h0oA (2.40),
4h73A (2.40),	4hnqA (2.40),	4ig4A (2.40),	4iz6A (2.40),	4k71B (2.40),	4k94C (2.40),
4lcwA (2.40),	4lp4A (2.40),	4m52B (2.40),	4m66A (2.40),	4mb8A (2.40),	4mhcA (2.40),
4mmsB (2.40),	4mz1A (2.40),	4nx3B (2.40),	4nx3D (2.40),	4olzA (2.40),	4oojA (2.40),
4oseA (2.40),	4c20A (2.41),	4fyeA (2.41),	1m4uA (2.42),	4im4A (2.42),	2cx6A (2.43),
3gwjA (2.43),	4j6tA (2.43),	1ya5T (2.44),	4fv1A (2.44),	4g59A (2.44),	4n2oA (2.44),
3o5yA (2.45),	4ba7A (2.45),	4h71B (2.45),	4lneB (2.45),	2arhA (2.46),	3wd8A (2.46),
4bifA (2.46),	4cc9A (2.47),	4h2rA (2.47),	4i2xE (2.48),	4mafA (2.48),	2fpnA (2.49),
3skvA (2.49),	3vywA (2.49),	3wlcA (2.49),	4binA (2.49),	4g4sO (2.49),	4iglA (2.49),
4iglB (2.49),	4kpoA (2.49),	1bt3A (2.50),	1bw0A (2.50),	1h21A (2.50),	1iw8B (2.50),
1jb0D (2.50),	1mspA (2.50),	1poiC (2.50),	1rcwA (2.50),	1sj7A (2.50),	1uunA (2.50),
1xo1A (2.50),	2alxA (2.50),	2b81A (2.50),	2bm8A (2.50),	2ffsA (2.50),	2i2xA (2.50),
2jh pA (2.50),	2nlyA (2.50),	2oggA (2.50),	2opiA (2.50),	2ph5A (2.50),	2qlzB (2.50),
2vuhB (2.50),	2ylhA (2.50),	3aj1A (2.50),	3ayfA (2.50),	3eatX (2.50),	3f3bA (2.50),
3floA (2.50),	3gn3A (2.50),	3klyA (2.50),	3kfwX (2.50),	3mt1A (2.50),	3s6n2 (2.50),
3safA (2.50),	3vmmA (2.50),	3vxmD (2.50),	3vxmE (2.50),	3w2wA (2.50),	3wajA (2.50),
3zdrA (2.50),	3ziaA (2.50),	3zigA (2.50),	4c50A (2.50),	4icnA (2.50),	4issA (2.50),
4j0mA (2.50),	4j57A (2.50),	4j6uA (2.50),	4l23B (2.50),	4l8nA (2.50),	4ldqA (2.50),
4llaA (2.50),	4lorA (2.50),	4m3cA (2.50),	4m44A (2.50),	4m8rA (2.50),	4mp4A (2.50),
4msvA (2.50),	4n3pA (2.50),	4n8vG (2.50),	4ne7A (2.50),	4nf aA (2.50),	4nltA (2.50),

Supplementary Material

S. Das et al.

4ob9A (2.50),	4q1zA (2.50),	1aiwA (NA),	1ayjA (NA),	1bdsA (NA),	1c01A (NA),
1cx1A (NA),	1d8bA (NA),	1dcjA (NA),	1de3A (NA),	1e8rA (NA),	1f2rI (NA),
1ib8A (NA),	1ksqA (NA),	1m2eA (NA),	1t4yA (NA),	1u6pA (NA),	1ug2A (NA),
1v32A (NA),	1v5rA (NA),	1w9rA (NA),	1wijA (NA),	1wilA (NA),	1wirA (NA),
1witA (NA),	1wqkA (NA),	1x3aA (NA),	1x5jA (NA),	1xpwA (NA),	1y7jA (NA),
1yhpA (NA),	1yuaA (NA),	1zg2A (NA),	2a2pA (NA),	2a7yA (NA),	2b3wA (NA),
2bgoA (NA),	2comA (NA),	2cryA (NA),	2dawA (NA),	2db2A (NA),	2di7A (NA),
2dipA (NA),	2dirA (NA),	2dk3A (NA),	2dktA (NA),	2dkyA (NA),	2e6jA (NA),
2e72A (NA),	2e7mA (NA),	2ec4A (NA),	2ekiA (NA),	2eqgA (NA),	2eqoA (NA),
2hgcA (NA),	2jm3A (NA),	2jo7A (NA),	2joqA (NA),	2jqoA (NA),	2jsnA (NA),
2ju5A (NA),	2k2vA (NA),	2k54A (NA),	2k87A (NA),	2k89A (NA),	2k8sA (NA),
2k9mA (NA),	2kafA (NA),	2khvA (NA),	2kskA (NA),	2kvrA (NA),	2kwuA (NA),
2kxyA (NA),	2l0zA (NA),	2l25A (NA),	2l3uA (NA),	2l6oA (NA),	2la7A (NA),
2lfpA (NA),	2lg7A (NA),	2ljxA (NA),	2loxA (NA),	2lpeA (NA),	2lqkA (NA),
2ltkA (NA),	2lv5A (NA),	2m3xA (NA),	2m4vA (NA),	2m5sA (NA),	2m9uA (NA),
2ma6A (NA),	2malA (NA),	2mc3A (NA),	2md7B (NA),	2mewA (NA),	2yujA (NA),
3zj1A (NA),	3zuaA (NA),	4bs2A (NA),	4by9D (NA),	4by9E (NA),	

Fig S1: GO directed acyclic graph for PDB: 1m56 (Cytochrome c oxidase) showing function representation at different GO levels. Each node represents a GO MF term. The level of each GO term is indicated beside the arrow. The terms span from general descriptions (level 1) to very specific ones (deeper levels represented by larger numbers). The same GO term can have multiple levels depending on how the path is traced to the root GO term. The term depicted in bold is annotated for the PDB; PDBs are annotated at the deepest possible level. (GO:0003674 – molecular function; GO:0003824 – catalytic activity; GO:0016491 – oxidoreductase activity; GO:0015002 – heme-copper terminal oxidase activity; **GO:0004129** – cytochrome-c oxidase activity; GO:0016675 – oxidoreductase activity, acting on heme group of donors; GO:0016676 – oxidoreductase activity, acting on heme group of donors, oxygen as receptor; GO:0005215 – transporter activity; GO:0022891 – substrate-specific transmembrane transporter activity; GO:0015075 – ion transmembrane transporter activity; GO:0008324 – cation transmembrane transporter activity; GO:0022890 – inorganic cation transmembrane transporter activity; GO:0015077 – monovalent inorganic cation transmembrane transporter activity; GO:0015078 – hydrogen ion transporter activity).

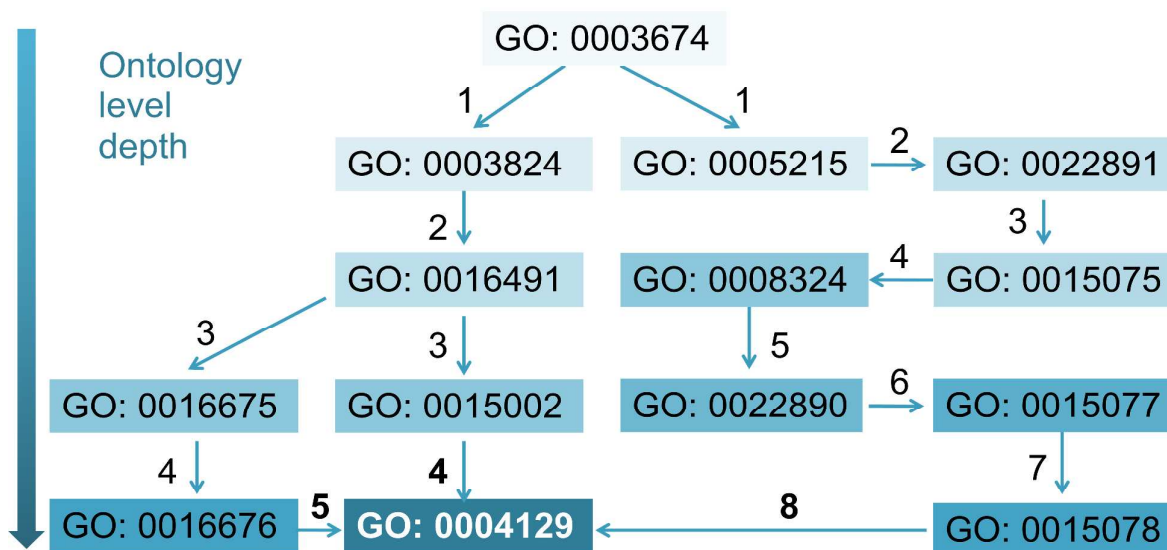


Table S2: Examples of proteins from PDB (resolution $\leq 2.5\text{\AA}$) which have functionally important fragments^L (propensity ≥ 20 , p-value ≤ 0.05) from the F0 dataset. ED values are indicated where less than 50% CC values were >0.9 and less than 25% CC values were >0.95 .

Type of match	F0 PDB (ref.) GO MF, propensity, p-value	F0 Fragment and its function ^L	PDB from test dataset (ref.)	Test Fragment and its function ^L	Overall sequence identity	% Match in CC		
						>0.90	>0.95	ED
<i>cis</i> -match, dynamics match GO match	2bo9B¶(1) Human latexin (carboxypeptidase inhibitor) F0008201, F0008191 53.2, 0.02	122: IPDNFG IPDXXG with <i>cis</i> -Pro is a conserved motif in the latexin family of proteins. The fragment is part of the protein interaction surface with carboxypeptidase.	1wnhA(2) Mouse latexin	122: IPDNFG The fragment is a conserved motif involved in protein- protein interaction	83.6	100	98	-
	4klxA(3) <i>Mycobacterium tuberculosis</i> dihydrofolate reductase (DHFR) F0004146 51.3, 0.001	54: VRPLPG The fragment is part of the folate binding site. The <i>cis</i> conformation of P56 is conserved in other DHFRs.	2w3wA <i>Mycobacterium avium</i> dihydrofolate reductase	58: VRPLPG The fragment is part of the folate binding site which contains a bound lipophilic anti-folate.	62.3	83	35	-
<i>cis</i> -match, dynamics match GO match	3r41A¶(4) Fluoroacetate dehalogenase from photosynthetic bacterium <i>Rhodopseudomonas palustris</i> F0016787 214.1, 0.00006	36: LLHGFP The fragment is part of the active site. Backbone amide of F40 is part of oxanion hole that stabilizes transient negative charge in intermediate reaction step. F40A mutant disrupts dehalogenation activity.	1ehyA(5) Epoxide hydrolase from <i>Agrobacterium radiobacter</i>	34: LLHGWP HGXP is a conserved motif found in epoxide hydrolase and other α/β hydrolase-fold enzymes, and is part of the active site. W38 is suggested to be part of an oxanion hole. The motif has been additionally suggested to stabilize the putative hydrolytic water molecule.	23.6	65	32	-

<i>cis</i> -match, dynamics match GO match	4ew1A¶(6) human glycinamide ribonucleotide transferase F0016742 59.7, 0.0000006	108: HPSLLP The fragment is part of the active site. It transfers formyl group from 10- Formyltetrahydrofolate to β-glycinamide ribonucleotide. His108 is involved in formyl transfer.	1s3iA(7) N-terminal hydrolase domain of 10- formyltetrahydrofo late dehydrogenase from rat	106: HPSLLP His106 is a key active site residue, residues 104-108 are part of the active site structure. It converts 10- Formyltetrahydrofolate to tetrahydrofolate. A βme molecule close to the fragment mimics formate (product from first step of the reaction).	16.6	69	40	-
	1kmvA(8) Human dihydrofolate reductase F0004146 32.9, 0.000002	112: VWIVGG* V115 interacts with the antifolate inhibitor while V115, G116 and G117 interact with the coenzyme NADP via hydrogen bond and vander Waals contacts. The <i>cis</i> conformation of G117 is conserved in other DHFR structures.	3jtwA Putative dihydrofolate reductase from <i>Pedococcus pentos aceus</i>	112: VWIVGG* The fragment is part of the ligand-binding pocket which contains electron density for some unknown ligand. It has high structural similarity to other dihydrofolate reductases (e.g. 2xw7, DALI(9) z-score: 20.6, rmsd: 2.4Å, sequence identity: 28%) and riboflavin biosynthesis proteins(e.g. 3ky8, DALI z-score: 21.7, rmsd: 2.1Å, sequence identity: 28%), both of which contain the DHFR domain. This protein has been assigned a GO of 5-amino-6- (5- phosphoribosylamino)uracil reductase activity in Uniprot(10), which may be a mis-annotation	19.1	91	74	-

<i>cis</i> -match, dynamics match GO mis-match	2jdcA(11) Glyphosate <i>N</i> - acetyltransferase from <i>Bacillus</i> <i>licheniformis</i> F0016746 72.8, 0.013	131: DTPPVG The fragment is part of the substrate-binding site and precedes the catalytic residue H138. The side chain of V135 is involved in van der Waals interactions with the competitive inhibitor 3-phosphoglycerate in the ligand-bound structure (PDB: 2jdd).	2omzB(12) Human epithelial- cadherin	44: DTPPVG The fragment is part of the binding interface between epithelial-cadherin and bacterial invasion protein internalin from <i>Listeria</i> <i>monocytogenes</i> .	15.1	79	54	-
	4c6eA(13) Dihydro-orotase domain of human CAD F0016812 32.4, 0.0000005	1475: REPGGT The fragment is part of the active site. R1475 side chain interacts with dihydroorotic acid and the product N- carbamoyl-L-aspartate. The closed conformation of the flexible loop (involved in catalysis) is favored by interactions between F1563 and R1475.	2pbrA Thymidylate kinase from <i>Aquifex Aeolicus</i>	35: REPGGT Comparison with PDB: 4edh (holo-structure of homolog) shows that this fragment forms the base of the TMP- binding pocket, where TMP is the phosphate group acceptor.	12.2	55	10	15.8

<i>cis</i> -mismatch, dynamics match GO match	1bx7A(14) Hirustasin (serine protease inhibitor) from the leech <i>Hirudo medicinalis</i> F0030414, F0004867 47.9, 0.02	42: NGCEYP This fragment is part of the secondary binding loop. The binding of hirustasin to tissue kallikrein (serine protease) is accompanied by a change in the relative orientation of the sub-domains in hirustasin changes and the <i>cis</i> to <i>trans</i> isomerisation of Pro47. C44 forms disulfide bridge with C24, which is one among the bridges that impart structural rigidity.	3bg4D(15) Guamerin (serine protease inhibitor) from <i>Hirudo</i> <i>nipponia</i> leech	48: NGCEYP This fragment is part of the binding interface to chymotrypsin. C50 forms disulfide bridge with C30, which is one among the bridges that impart structural rigidity.	47.5	93	68	-
	2octA¶(16) Human cystatin B F0004869, F0030414 41.9, 0.02	69: VFQSLP The <i>trans</i> -to- <i>cis</i> isomerization of P74 (conserved throughout the cystatin superfamily) results in the formation of a tetramer from two domain-swapped dimers (these higher oligomers lead to amyloid fibrils). P74 is found in <i>trans</i> - conformation in other reported structures in monomeric or dimeric form.	3kseD Human cystatin A	69: VFKSPL P74 is observed in the <i>trans</i> form in monomers (this PDB) and domain-swapped dimers (PDB: 1n9j)(17). In this PDB, the fragment from cystatin A (a thiol proteinase inhibitor) is at the binding interface with cathepsin L (thiol proteinase).	53.1	83	64	-

<i>cis</i> -mismatch, dynamics match GO mis-match	2uu8A¶(18) Ni, Ca concanavalin A from <i>Canavalia</i> <i>ensiformis</i> (Jack bean) F0005537 108.5, 0.009	207: ADGIAF* ²⁰⁷ AD ²⁰⁸ is part of the monosaccharide binding site (19). The ligation of metal ion to metal-free concanavalin A causes <i>trans</i> -to- <i>cis</i> isomerization of D208, which leads to the proper binding orientation of D208 and the carbohydrate specificity determining loop (20).	2kucA putative disulphide- isomerase from <i>Bacteroides</i> <i>thetaiotaomicron</i>	3: ADGIAF The selected fragment is located on the protein surface but is away from the putative active site (³⁶ CGPC ³⁹) motif in disulfide isomerases (12.3Å between the Cα atoms of A7 and C39 (21).	10.5	87	28	-
	2ciwA(22) Chloroperoxidase from fungus <i>Caldariomyces</i> <i>fumago</i> F0004601 31.2, 0.03	3: GSGIGY(P) The enzyme is extensively glycosylated. The above fragment is close to N12, which has N-glycosylation. The carbonyl of <i>cis</i> -P9 is involved in hydrogen bond with <i>N</i> - acetylglucosamine.	luxxX(23) Xylanase U from <i>Ruminiclostridium</i> <i>thermocellum</i>	29: GSGIGY(I) The fragment is part of the binding site. Y34 is a conserved residue and interacts with the bound xylopentaose.	13.0	22	11	31.3

<i>cis</i> -mismatch, dynamics mis-match GO match	4mk3A Glutathione S- transferase from <i>Cupriavidus</i> <i>metallidurans</i> F0004364 26.8, 4.5e-9	47: GKVPCL The fragment is part of the glutathione (enzyme's reactant) binding site. Residues 47-50 are involved in non-bonded contacts and hydrogen bonds with the bound glutathione sulfinate. The <i>cis</i> -Pro bend is essential for correct backbone conformation at the glutathione binding site(24).	1oe8A(25) Glutathione S- transferase from <i>Schistosoma</i> <i>haematobium</i>	130: GKVPVL This fragment is away from the active site but close to the dimer interface. The fragment ⁵¹ GRLPAV ⁵⁶ , equivalent to the fragment from 4mk3, does not get selected because the sequences differ by more than one residue.	18.8	0	0	55.9
	4gxzA(26) Suppression of copper sensitivity protein (DSBA-like thioredoxin domain) from <i>Salmonella enteric</i> serovar Typhimurium F0015035 24.6, 1.6e-9	154: QGTPAT Comparison with other proteins containing the same domain shows that the fragment is close to the active site ⁴⁸ CXXC motif and is part of the substrate-binding cleft. <i>Cis</i> -Pro157 is important for the function and stability of both oxidised and reduced forms of DsbA(27).	1gh2A(28) Human thioredoxin-like protein	65: QGTAAT This fragment precedes the fragment ⁷⁴ SATPTF ⁷⁹ , which is equivalent to the fragment from 4gxzA but is not detected because the sequences differ by more than one residue.	16.1	2	1	53.7

PDB fragments marked with * contain Xaa-Xnp *cis*-bond.

⊥Residues in *cis* conformation are in bold, and in parentheses when they occur just after the selected fragment.

Fragments marked with ¶ are either from PDBs which have no PROSITE pattern, or the fragment is outside the PROSITE pattern.

Table S3: Suitability of coarse-grained forcefield in reproducing actual protein dynamics:

comparison of full protein RMSF_{norm} curves and ACV profiles of *cis*-peptide fragment in structures with *cis*-peptides from the non-redundant dataset.

PDB ID	Correlation Coefficient (Full protein RMSF _{norm})		% of Correlation Coefficient values >N between <i>cis</i> -peptide fragment ACV profiles from CGMM & NMR frames	
	CGMM vs NMR	Martini 2.2 vs NMR	N=0.90	N=0.95
1aiw	0.75	0.74	89	50
1a7m	0.94	0.61	45	28
1apj	0.91	0.58	91	76
1aw6	0.97	0.77	85	18
1c01	0.48	0.53	76	43
1cx1	0.76	0.46	53	26
1dcj	0.90	0.45	100	91
1g9l	0.89	0.62	47	15
1h3z	0.75	0.48	81	29
1i1s	0.91	0.62	100	73
1ie5	0.82	0.28	91	53
1j7m	0.92	0.59	59	15
1mvz	0.93	0.50	65	40
1p8b	0.90	0.72	83	49
1rmj	0.90	---	65	22
1sjg	0.77	0.36	76	60
1w9r	0.96	0.69	61	30
1wh2	0.94	0.60	71	32
1wic	0.93	0.48	59	20
1x3a	0.72	0.60	100	26
1yua	0.88	0.66	91	54
2aiv	0.76	0.39	100	40
2dky	0.95	0.7	64	29
2ffw	0.89	0.69	64	36
2kpu	0.96	0.59	74	49
2kxj	0.84	0.54	98	43
2la7	0.98	0.56	58	31
2lah	0.82	0.57	70	45

‘---’ means 1μs long trajectory is not available

CC: Correlation Coefficient; CC value ≥ 0.90 are in bold.

$$\text{RMSF}_{\text{norm}} = (\text{RMSF} - \langle \text{RMSF} \rangle) / \sigma(\text{RMSF}) \quad (29)$$

Cis-peptide fragment is of length six residues centered on the *cis*-bond. Fragments were said to have matching dynamics if at least 50% CC values were >0.9 and at least 25% CC values were >0.95 .

Table S4: Numbers of hits in different classes when searching for F0 fragments in the test dataset using match in *cis*-peptide and dynamics

GO MF	<i>Cis</i> -peptide match		No <i>cis</i> -peptide (<i>trans</i> match)	
	Dynamics		Dynamics	
	Match	Mismatch	Match	Mismatch
Match	94	0	31	6
Mismatch	5	3	32	33

Table S5: Annotation of the PDBs with unknown function containing fragments from the F0 dataset with matching *cis*-peptide[⊥] and dynamics.

F0 PDB (ref.) GO MF, propensity, p-value	F0 Fragment and its function [⊥]	PDB from UF dataset [ref.]	UF Fragment and its possible function [⊥]	% sequence identity of UF to F0 PDB	% Match in CC		
					>0.90	>0.95	ED
4iedA [¶] Class D beta-lactamase from <i>Fusobacterium nucleatum</i> F0008658 61.4, 0.000002	146: LEGPLK Beta-lactamase provides resistance to beta-lactam antibiotics. The structurally equivalent fragment in a homolog from <i>Klebsiella pneumoniae</i> (PDB: 4s2n) is part of the binding site of the inhibitor avibactam (30). Close to this fragment (<12Å) is the conserved motif SXXK. Ser is required for acylation, while the carboxylated Lys is involved in Ser activation during acylation and activation of the catalytic water during hydrolytic deacylation.	3ix7A Uncharacterized protein from <i>Thermus thermophilus</i> [‡]	164: LEGPLW A short segment KALS is present within 12Å of the matching <i>cis</i> -fragment. Comparing the dynamics of this second fragment and the SXXK motif in 4ied yields 55% CC values above 0.90 but all below 0.95. These two segments are adjacent to a shallow groove (See Figure marked [‡] below). Hence this protein may have a similar function (F0008658: penicillin binding).	12	77	40	-
117aA [¶] Cephalosporin C deacetylase from <i>Bacillus subtilis</i> F0047739 203.4, 0.0049	216: VALEQP The active site is inside the tunnel of a hexameric assembly (PDB: 1ods) which can bind to different, small acetylated molecules (31). The selected fragment is at the binding interface of the adjacent monomers. In	1wv8A Hypothetical protein from <i>Thermus thermophilus</i>	54: VALELP Although the dynamics of the detected fragments match, we did not find any additional active site residues/ motifs from 117aA in 1wv8A. Hence we annotate this as the putative binding site of the protein.	13	69	34	-

	the holo structure (PDB: 1odt), the <i>cis</i> -Pro is in contact with the bound acetate ion.						
4l23B(32) Phosphatidylinositol 3-kinase regulatory subunit alpha from human F0035014 32.5, 0.03	392: FSDPLT The protein has been co-crystallized with its binding partner phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform. The selected fragment is at the interface of the two protein chains.	3k6oA Uncharacterized protein from <i>Bacteroides vulgatus</i>	223:FSYPLT In the absence of additional active site residues/ motifs from 4l23B in 3k6oA, we annotate this as the putative binding site of the protein.	16.8	67	31	-
2q88A(33) EhuB, ABC transporter from <i>Sinorhizobium meliloti</i> F0033294 227.6, 0.004	19: ANEPPF The protein helps in the uptake of compatible solutes ectoine and hydroxyectoine. The fragment is at the interface of 2 domains and part of the substrate-binding site. The conserved <i>cis</i> -peptide positions F24 in a proper conformation to interact with the substrate (ectoine/ hydroxyectoine) via cation- π interactions.	1vhoA(34) Putative peptidase/ endoglucanase from <i>Thermotoga maritima</i>	229: ASEPPF No additional active site residues/ motifs from 2q88A were found in 1vhoA. The fragment loop in 1vhoA moves towards the other domain during MD simulations. Based on the match of <i>cis</i> -peptide and dynamics, the fragment appears to be part of a putative binding site at the interface.	23	63	33	-
3vwvA (35) Protein disulfide isomerase (PDI) family protein P5 from Human F0015035	97: VQGFPT The fragment is involved in functional interaction with its receptor Peroxiredoxin-4 (Prx4), as seen from the structure with a bound C-	3h79A Thioredoxin-like hypothetical protein from <i>Trypanosoma cruzi</i>	90: VSGFPT The 2 proteins share a high structural similarity (DALI z-score: 17.6, rmsd: 1.2Å). However, the CXXC motif is absent in this protein. Hence	21	75	66	-

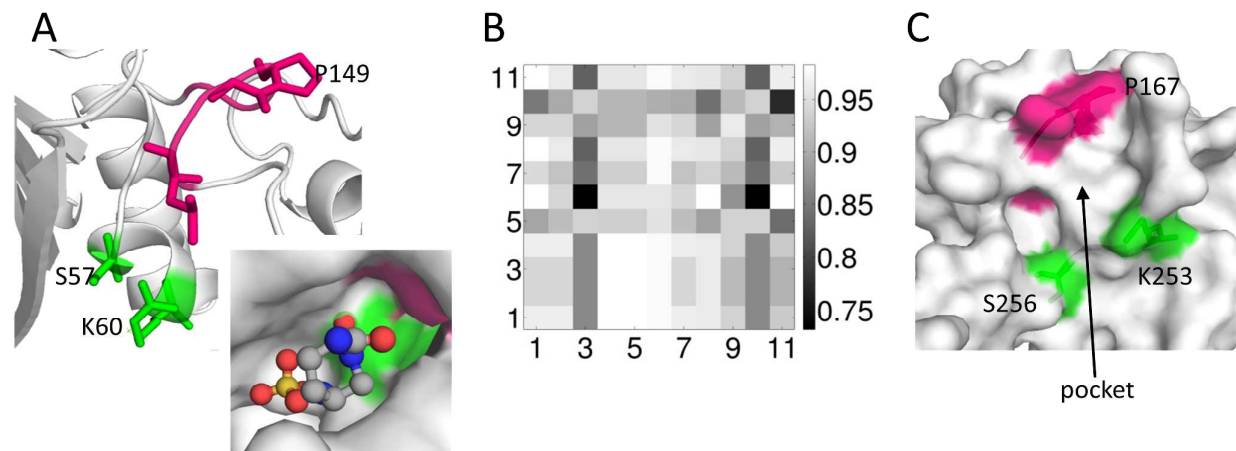
75.6, 3.5e-7	terminal peptide of Prx4 (PDB: 3w8j) and is close to the active site motif (⁵⁷ CXXC).	we label the <i>cis</i> -fragment as the putative binding site for 3h79A.					
3n10A(36) Class IV Adenylyl Cyclase from <i>Yersinia pestis</i> F0004016 58.3, 0.017	72: LWIVKG(P) The fragment is part of the substrate binding site. Residues L73, I74 and K76 are found to interact with the bound product molecule. E12 and E136 bind to Mn ²⁺ ion and stabilize it, while K14, R63, K76, K111 and R113 interact with the triphosphate-Mn ²⁺ moiety.	2acaA putative adenylyl cyclase from <i>Vibrio parahaemolyticus</i>	74: LWIVKG(P) The two proteins share a similar overall structure (DALI z-score: 27, rmsd: 1Å). The selected fragment is part of the binding site, with two phosphate groups close to K78. Moreover, the aforementioned residues interacting with triphosphate and Mn ²⁺ in 3n10A are conserved in 2acaA.	43	46	22	5

⊥-residues in *cis* conformation are in bold, and in parentheses when they occur just after the selected fragment.

⚡ GO MF present in the UNIPROT database (10), and evidence code given as IEA.

Fragments marked with ¶ are either from PDBs which have no PROSITE pattern, or the fragment is outside the PROSITE pattern.

⚡ Figure legend (for the figure below): comparison of the *cis*-fragment (pink) in the F0 protein 4iedA and UF protein 3ix7A. (A) The F0 fragment is part of the binding site and close to the catalytic motif SXXK (green). Inset depicts the motifs and the bound avibactam inhibitor (ball-and-stick model) in ahomolog (PDB: 4s2n). (B) ACVmatrix depicting matching dynamics between the selected fragments. (C) The *cis*-fragment in the UF protein is within 12Å of a stretch with sequence ²⁵³KALS²⁵⁶. These two segments line a shallow groove, which is a putative binding site.



References

1. Pallares, I.; Bonet, R.; Garcia-Castellanos, R.; Ventura, S.; Aviles, F. X.; Vendrell, J.; Gomis-Ruth, F. X., Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin. *Proc. Natl. Acad. Sci. USA* **2005**, 102, (11), 3978-83.
2. Aagaard, A.; Listwan, P.; Cowieson, N.; Huber, T.; Ravasi, T.; Wells, C. A.; Flanagan, J. U.; Kellie, S.; Hume, D. A.; Kobe, B.; Martin, J. L., An inflammatory role for the mammalian carboxypeptidase inhibitor latexin: relationship to cystatins and the tumor suppressor TIG1. *Structure* **2005**, 13, (2), 309-17.
3. Dias, M. V.; Tyrakis, P.; Domingues, R. R.; Paes Leme, A. F.; Blundell, T. L., Mycobacterium tuberculosis dihydrofolate reductase reveals two conformational states and a possible low affinity mechanism to antifolate drugs. *Structure* **2014**, 22, (1), 94-103.
4. Chan, P. W.; Yakunin, A. F.; Edwards, E. A.; Pai, E. F., Mapping the reaction coordinates of enzymatic defluorination. *J Am Chem Soc* **2011**, 133, (19), 7461-8.
5. Nardini, M.; Ridder, I. S.; Rozeboom, H. J.; Kalk, K. H.; Rink, R.; Janssen, D. B.; Dijkstra, B. W., The x-ray structure of epoxide hydrolase from *Agrobacterium radiobacter* AD1. An enzyme to detoxify harmful epoxides. *J Biol Chem* **1999**, 274, (21), 14579-86.
6. Connelly, S.; DeMartino, J. K.; Boger, D. L.; Wilson, I. A., Biological and structural evaluation of 10R- and 10S-methylthio-DDACTHF reveals a new role for sulfur in inhibition of glycineamide ribonucleotide transformylase. *Biochemistry* **2013**, 52, (30), 5133-5144.
7. Chumanevich, A. A.; Krupenko, S. A.; Davies, C., The crystal structure of the hydrolase domain of 10-formyltetrahydrofolate dehydrogenase: mechanism of hydrolysis and its interplay with the dehydrogenase domain. *J. Biol. Chem.* **2004**, 279, (14), 14355-64.
8. Klon, A. E.; Heroux, A.; Ross, L. J.; Pathak, V.; Johnson, C. A.; Piper, J. R.; Borhani, D. W., Atomic structures of human dihydrofolate reductase complexed with NADPH and two lipophilic antifolates at 1.09 Å and 1.05 Å resolution. *J. Mol. Biol.* **2002**, 320, (3), 677-693.
9. Holm, L.; Rosenstrom, P., Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **2010**, 38, (Web Server issue), W545-9.
10. Magrane, M.; Consortium, U., UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* **2011**, 2011, bar009.
11. Siehl, D. L.; Castle, L. A.; Gorton, R.; Keenan, R. J., The molecular basis of glyphosate resistance by an optimized microbial acetyltransferase. *J. Biol. Chem.* **2007**, 282, (15), 11446-11455.
12. Wollert, T.; Heinz, D. W.; Schubert, W. D., Thermodynamically reengineering the listerial invasion complex InlA/E-cadherin. *Proc Natl Acad Sci U S A* **2007**, 104, (35), 13960-13965.
13. Grande-Garcia, A.; Lallous, N.; Diaz-Tejada, C.; Ramon-Maiques, S., Structure, functional characterization, and evolution of the dihydroorotase domain of human CAD. *Structure* **2014**, 22, (2), 185-98.
14. Uson, I.; Sheldrick, G. M.; de La Fortelle, E.; Bricogne, G.; Di Marco, S.; Priestle, J. P.; Grutter, M. G.; Mittl, P. R., The 1.2 Å crystal structure of hirutastin reveals the intrinsic flexibility of a family of highly disulphide-bridged inhibitors. *Structure* **1999**, 7, (1), 55-63.
15. Kim, H.; Chu, T. T.; Kim, D. Y.; Kim, D. R.; Nguyen, C. M.; Choi, J.; Lee, J. R.; Hahn, M. J.; Kim, K. K., The crystal structure of guamerin in complex with chymotrypsin and the development of an elastase-specific inhibitor. *J. Mol. Biol.* **2008**, 376, (1), 184-92.
16. Jenko Kokalj, S.; Guncar, G.; Stern, I.; Morgan, G.; Rabzelj, S.; Kenig, M.; Staniforth, R. A.; Waltho, J. P.; Zerovnik, E.; Turk, D., Essential role of proline isomerization in stefin B tetramer formation. *J Mol Biol* **2007**, 366, (5), 1569-79.
17. Staniforth, R. A.; Giannini, S.; Higgins, L. D.; Conroy, M. J.; Hounslow, A. M.; Jerala, R.; Craven, C. J.; Waltho, J. P., Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. *Embo J* **2001**, 20, (17), 4774-81.
18. Ahmed, H. U.; Blakeley, M. P.; Cianci, M.; Cruickshank, D. W.; Hubbard, J. A.; Helliwell, J. R., The determination of protonation states in proteins. *Acta Crystallogr D Biol Crystallogr* **2007**, 63, (Pt 8), 906-22.
19. Sanders, D. A.; Moothoo, D. N.; Raftery, J.; Howard, A. J.; Helliwell, J. R.; Naismith, J. H., The 1.2 Å resolution structure of the Con A-dimannose complex. *J Mol Biol* **2001**, 310, (4), 875-84.
20. Bouckaert, J.; Dewallef, Y.; Poortmans, F.; Wyns, L.; Loris, R., The structural features of concanavalin A governing non-proline peptide isomerization. *J Biol Chem* **2000**, 275, (26), 19778-87.
21. Mac, T. T.; von Hacht, A.; Hung, K. C.; Dutton, R. J.; Boyd, D.; Bardwell, J. C.; Ulmer, T. S., Insight into disulfide bond catalysis in Chlamydia from the structure and function of DsbH, a novel oxidoreductase. *J Biol Chem* **2008**, 283, (2), 824-32.

22. Kuhnel, K.; Blankenfeldt, W.; Turner, J.; Schlichting, I., Crystal structures of chloroperoxidase with its bound substrates and complexed with formate, acetate, and nitrate. *J Biol Chem* **2006**, 281, (33), 23990-8.
23. Pires, V. M.; Henshaw, J. L.; Prates, J. A.; Bolam, D. N.; Ferreira, L. M.; Fontes, C. M.; Henrissat, B.; Planas, A.; Gilbert, H. J.; Czjzek, M., The crystal structure of the family 6 carbohydrate binding module from *Cellvibrio mixtus* endoglucanase 5a in complex with oligosaccharides reveals two distinct binding sites with different ligand specificities. *J Biol Chem* **2004**, 279, (20), 21560-8.
24. Thom, R.; Dixon, D. P.; Edwards, R.; Cole, D. J.; Laphorn, A. J., The structure of a zeta class glutathione S-transferase from *Arabidopsis thaliana*: characterisation of a GST with novel active-site architecture and a putative role in tyrosine catabolism. *J. Mol. Biol.* **2001**, 308, (5), 949-62.
25. Johnson, K. A.; Angelucci, F.; Bellelli, A.; Herve, M.; Fontaine, J.; Tsernoglou, D.; Capron, A.; Trottein, F.; Brunori, M., Crystal structure of the 28 kDa glutathione S-transferase from *Schistosoma haematobium*. *Biochemistry* **2003**, 42, (34), 10084-94.
26. Shepherd, M.; Heras, B.; Achard, M. E.; King, G. J.; Argente, M. P.; Kurth, F.; Taylor, S. L.; Howard, M. J.; King, N. P.; Schembri, M. A.; McEwan, A. G., Structural and functional characterization of ScsC, a periplasmic thioredoxin-like protein from *Salmonella enterica* serovar Typhimurium. *Antioxid Redox Signal* **2013**, 19, (13), 1494-506.
27. Charbonnier, J. B.; Belin, P.; Moutiez, M.; Stura, E. A.; Quemeneur, E., On the role of the cis-proline residue in the active site of DsbA. *Protein Sci.* **1999**, 8, (1), 96-105.
28. Jin, J.; Chen, X.; Zhou, Y.; Bartlam, M.; Guo, Q.; Liu, Y.; Sun, Y.; Gao, Y.; Ye, S.; Li, G.; Rao, Z.; Qiang, B.; Yuan, J., Crystal structure of the catalytic domain of a human thioredoxin-like protein. *Eur J Biochem* **2002**, 269, (8), 2060-8.
29. Bhadra, P.; Pal, D., De novo inference of protein function from coarse-grained dynamics. *Proteins* **2014**, 82, (10), 2443-2454.
30. King, D. T.; King, A. M.; Lal, S. M.; Wright, G. D.; Strynadka, N. C. J., Molecular Mechanism of Avibactam-Mediated β -Lactamase Inhibition. *ACS Infect. Dis.* **2015**, 1, (4), 175-184.
31. Vincent, F.; Charnock, S. J.; Verschuere, K. H.; Turkenburg, J. P.; Scott, D. J.; Offen, W. A.; Roberts, S.; Pell, G.; Gillbert, H. J.; Davies, G. J.; Brannigan, J. A., Multifunctional xylooligosaccharide/cephalosporin C deacetylase revealed by the hexameric structure of the *Bacillus subtilis* enzyme at 1.9 Å resolution. *J. Mol. Biol.* **2003**, 330, (3), 593-606.
32. Zhao, Y.; Zhang, X.; Chen, Y.; Lu, S.; Peng, Y.; Wang, X.; Guo, C.; Zhou, A.; Zhang, J.; Luo, Y.; Shen, Q.; Ding, J.; Meng, L., Crystal Structures of PI3K α Complexed with PI103 and Its Derivatives: New Directions for Inhibitors Design. *ACS Med Chem Lett* **2014**, 5, (2), 138-42.
33. Hanekop, N.; Hoing, M.; Sohn-Bosser, L.; Jebbar, M.; Schmitt, L.; Bremer, E., Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J Mol Biol* **2007**, 374, (5), 1237-50.
34. Badger, J.; Sauder, J. M.; Adams, J. M.; Antonysamy, S.; Bain, K.; Bergseid, M. G.; Buchanan, S. G.; Buchanan, M. D.; Batiyenko, Y.; Christopher, J. A.; Emtage, S.; Eroshkina, A.; Feil, I.; Furlong, E. B.; Gajiwala, K. S.; Gao, X.; He, D.; Hendle, J.; Huber, A.; Hoda, K.; Kearins, P.; Kissinger, C.; Laubert, B.; Lewis, H. A.; Lin, J.; Loomis, K.; Lorimer, D.; Louie, G.; Maletic, M.; Marsh, C. D.; Miller, I.; Molinari, J.; Muller-Dieckmann, H. J.; Newman, J. M.; Noland, B. W.; Pagarigan, B.; Park, F.; Peat, T. S.; Post, K. W.; Radojicic, S.; Ramos, A.; Romero, R.; Rutter, M. E.; Sanderson, W. E.; Schwinn, K. D.; Tresser, J.; Winhoven, J.; Wright, T. A.; Wu, L.; Xu, J.; Harris, T. J., Structural analysis of a set of proteins resulting from a bacterial genomics project. *Proteins* **2005**, 60, (4), 787-96.
35. Sato, Y.; Kojima, R.; Okumura, M.; Hagiwara, M.; Masui, S.; Maegawa, K.; Saiki, M.; Horibe, T.; Suzuki, M.; Inaba, K., Synergistic cooperation of PDI family members in peroxiredoxin 4-driven oxidative protein folding. *Sci Rep* **2013**, 3, 2456.
36. Gallagher, D. T.; Kim, S. K.; Robinson, H.; Reddy, P. T., Active-site structure of class IV adenylyl cyclase and transphyletic mechanism. *J Mol Biol* **2011**, 405, (3), 787-803.