Supporting Information

Discovering Novel Alternaria solani Succinate Dehydrogenase Inhibitors by In Silico Modeling and Virtual Screening Strategies to Combat Early Blight

Sehrish Iftikhar¹, Ahmad Ali Shahid^{1,2}, Sobia Ahsan Halim^{3,*}, Jaap Wolters⁴, Vivianne G A A Vleeshouwers⁴, Ajmal Khan^{5,6,*}, Ahmed Al-Harrasi⁶, Shahbaz Ahmad¹

¹Institute of Agricultural Sciences, University of the Punjab, Lahore, Pakistan. ²Center of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan. ³Department of Biochemistry, Kinnaird College for Women, 93-Jail Road, Lahore-Pakistan. ⁴Wageningen UR Plant Breeding, Wageningen University and Research Center, Wageningen, The Netherlands. ⁵Department of Chemistry, COMSATS Institute of Information Technology, Abbottabad-22060, Pakistan. ⁶UoN Chair of Oman Medicinal Plants and Marine Products, University of Nizwa, Sultanate of Oman

***Correspondence:** Sobia Ahsan Halim, E-mail: <u>sobia.ahsan@kinnaird.edu.pk</u> <u>sobiahal@gmail.com</u>, Tel. Off: +(92) 42 99203781-84 (Ext 226), Ajmal Khan, email: <u>ajmalchemist@yahoo.com</u>

Supplementary Data

Sequences of A. solani SDH subunits.

A. solani SDH subunit A

PVIDHEYDAIVVGAGGAGLRAAFGLAEAGFNTACISKLFPTRSHTVAAQGGINAALGSM HEDDWRWHMYDTVKGSDWLGDQDAIHYMTREAPQSVIELENYGCPFSRTDDGKMYQ RAFGGQSQKYGKGGQAYRCCAAADRTGHALLHTLYGQSLRHNTKYFIEFFATDLIMED GVCKGVVAYNQEDGTIHRFIAKNTVLATGGYGRAYFSCTSAHTCTGDGMAMVARAGL PNQDLEFVQFHPTGIYGAGCLITEGSRGEGGYLLNSEGERFMERYAPTAKDLASRDVVS RSMTLEIREGRGVGPEKDHIYLQLSHLPAEVLHERLPGISETAAIFAGVDVTKQPIPVLPT VHYNMGGIPTKYTGEVITQDAQGNDQVVPGLFACGEAASVSVHGANRLGANSLLDLIV FGRAVSHTIRDNFSPGQKADPVSADAGADSISVLDQIRTSDGPKSTAEIRLQMQKVMQT DVAVFRTQESLDEGVKKIHDVDADFANVGIKDRSMIWNSDLVETLELRNLLTCAVQTA ESAANRKESRGAHAREDFPDRDDEQWMKHTLSWQKKPHGETQLGYRKVVGTTLDEAE CKAVPPFKRTY

A. solani SDH subunits B (GenBank: AGS56260.1)

MASIRAFTRLATSDLVPVRPAVFSRGFASVNDVHAREPISKIAEKIAPDASRSPVPESKTS TVPEPEPSKDAKTKTFHIYRWNPDEPTSKPKMQSYTLDLNKTGPMMLDALIRIKNEVDP TLTFRRSCREGICGSCAMNIDGVNTLACLCRIPTDTTKESRIYPLPHMYVVKDLVPDMTL FYKQYRSVKPYLQRSTAAPDGREFRQSKEDRKKLDGLYECILCACCSTSCPSYWWNQE EYLGPAVLLQSYRWIADSRDEKKAERQDALNNSMSLYRCRTILNCSRTCPKGLNPALAI AEIKKSMAFT

A. solani SDH subunits C (AGS56262.1)

MASQRVFQLGLRRAAAPGLRVQPVGRMVQRRLAATEHASQSEAAEILAKQRINRPVSP HLAIYRPQITWYASSLNRITGITLSGSLYLFGIAYLIAPYTGWHMETQSMVATVAAWPAA AKAGLKAFYAFPFFFHSFNGLRHLSWDVGIGFKNQQVIRLGW

A. solani SDH subunits D (GenBank: AGS56264.1)

MASVMRPGLLRQACPPAHASQRMLSTATATTSRPLAQQLRPAFQRSAVPKATRIAAFH ATQRTQILPPLPQKIIGTTNDPVPVPDPDYAHGSYHWSFERIVSAGLIPLTIAPFAAGSLNP LTDSILCALLVVHSHIGFESCIIDYFPAKRVPKTRKAAMWALRAGTLTLGLALYSFETND VGITE

Supplementary Figures and Tables



Supplementary Figure S1. Schematic presentation of Computational Work Flow

SDHA	PVIDHEYDAIVVGAGGAGLRAAFGLAEAGFNTACISKLFPTRSHTVA
1YQ3A	-STRVSDSTSTQYPVVDHEFDAVVVGAGGAGLRAAFGLSEAGENTACVTKLFPTRSHTVA
120YA	SSAKVSDATSTQYPVVDHEFDAVVVGAGGAGLRAAFGLSEAGENTACVTKLFPTRSHTVA
	:*:**:*****************************
SDHA	AQGGINAALGSMHEDDWRWHMYDTVKGSDWLGDQDAIHYMTREAPQSVIELENYGCPFSR
1YQ3A	AQGGINAALGNMEDDNWRWHFYDTVKGSDWLGDQDAIHYMTEQAPAAVIELENYGMPFSR
1ZOYA	AQGGINAALGNMEEDNWRWHFYDTVKGSDWLGDQDAIHYMTEQAPASVVELENYGMPFSR

SDHA	TDDGKMYQRAFGGQSQKYGKGGQAYRCCAAADRTGHALLHTLYGQSLRHNTKYFIEFFAT
1YQ3A	TEEGKIYQRAFGGQSLQFGKGGQAHRCCCVADRTGHSLLHTLYGRSLRYDTSYFVEYFAL
1ZOYA	TEDGKIYQRAFGGQSLKFGKGGQAHRCCCVADRTGHSLLHTLYGRSLRYDTSYFVEYFAL
	* : ** : ******* : : ***** : *** : **** : **** : **** : **** : *** : ** : * : * : * : * : * : * : * : * : * : *
SDHA	DLIMEDGVCKGVVAYNQEDGTIHRFIAKNTVLATGGYGRAYFSCTSAHTCTGDGMAMVAR
1YQ3A	DLLMENGECRGVIALCIEDGTIHRFRAKNTVIATGGYGRTYFSCTSAHTSTGDGTAMVTR
1ZOYA	DLLMENGECRGVIALCIEDGSIHRIRARNTVVATGGYGRTYFSCTSAHTSTGDGTAMVTR
	::* *:**:* ***: *:***:*****:********
SDHA	AGLPNQDLEFVQFHPTGIYGAGCLITEGSRGEGGYLLNSEGERFMERYAPTAKDLASRDV
1YQ3A	AGLPCQDLEFVQFHPTGIYGAGCLITEGCRGEGGILINSQGERFMERYAPVAKDLASRDV
1ZOYA	AGLPCQDLEFVQFHPTGIYGAGCLITEGCRGEGGILINSQGERFMERYAPVAKDLASRDV
	**** **********************************
SDHA	VSRSMTLEIREGRGVGPEKDHIYLQLSHLPAEVLHERLPGISETAAIFAGVDVTKQPIPV
1YQ3A	VSRSMTIEIREGRGCGPEKDHVYLQLHHLPPQQLATRLPGISETAMIFAGVDVTKEPIPV
1ZOYA	VSRSMTLEIREGRGCGPEKDHVYLQLHHLPPEQLAVRLPGISETAMIFAGVDVTKEPIPV
	****** ****** ****** **** *** *** * ****
SDHA	LPTVHYNMGGIPTKYTGEVITQDAQGNDQVVPGLFACGEAASVSVHGANRLGANSLLDLI
1YQ3A	LPTVHYNMGGIPTNYKGQVITH-VNGEDKVVPGLYACGEAASASVHGANRLGANSLLDLV
1ZOYA	LPTVHYNMGGIPTNYKGQVLRH-VNGQDQVVPGLYACGEAACASVHGANRLGANSLLDLV

SDHA	VFGRAVSHTIRDNFSPGQKADPVSADAGADSISVLDQIRTSDGPKSTAEIRLQMQKVMQT
1YQ3A	VFGRACALTIAETCKPGEPVPSIKPNAGEESVANLDKLRFADGTIRTSEARLNMQKTMQS
1ZOYA	VFGRACALSIAESCRPGDKVPSIKPNAGEESVMNLDKLRFANGTIRTSELRLSMQKSMQS
	***** : :* :. **: . :. :** :*: **::* ::*
SDHA	DVAVFRTQESLDEGVKKIHDVDADFANVGIKDRSMIWNSDLVETLELRNLLTCAVQTAES
1YQ3A	HAAVFRTGSILQEGCEKLSQIYRDLAHLKTFDRGIVWNTDLVETLELQNLMLCALQTIYG
1ZOYA	HAAVFRVGSVLQEGCEKILRLYGDLQHLKTFDRGMVWNTDLVETLELQNLMLCALQTIYG
	··**** · *:** :*: : *: :: **·:**:********
SDHA	AANRKESRGAHAREDFPDRDDEQWMKHTLSWQKKPHGETQLGY
1YQ3A	AEARKESRGAHAREDYKLRIDEFDYSKPLQGQQKRPFEEHWRKHTLSYVDVKSGKVTLKY
1ZOYA	AEARKESRGAHAREDFKERVDEYDYSKPIQGQQKKPFQEHWRKHTLSYVDVKTGKVSLEY
	* *************************************
SDHA	RKVVGTTLDEAECKAVPPFKRTY
1YQ3A	RPVIDRTLNEEDCSSVPPAIRSY
1ZOYA	RPVIDKTLNEADCATVPPAIRSY
	* *:. **:* :* :*** *:*

Supplementary Figure S2 (A). Multiple sequence alignment of SDHA.

SDHB 1YQ3B 1ZOYB	MASIRAFTRLATSDLVPVRPAVFSRGFASVNDVHAREPISKIAEKIAPDASRSPVPESKT
SDHB 1YQ3B 1ZOYB	STVPEPEPSKDAKTKTFHIYRWNPDEPTSKPKMQSYTLDLNKTGPMMLDALIRIKNEVDP AQTAAAATSRIKKFSIYRWDPDKPGDKPRMQTYEVDLNKCGPMVLDALIKIKNELDS AQTAAATAPRIKKFAIYRWDPDKTGDKPHMQTYEIDLNNCGPMVLDALIKIKNEIDS : : : *.* ****:**: .**:**: ***:**:**:*****
SDHB	TLTFRRSCREGICGSCAMNIDGVNTLACLCRIPTDTTKESRIYPLPHMYVVKDLVPDMTL
1Y03B	TLTFRRSCREGICGSCAMNIAGGNTLACTKKIDPDLSKTTKIYPLPHMYVVKDLVPDLSN
1ZOYB	TLTFRRSCREGICGSCAMNINGGNTLACTRRIDTNLDKVSKIYPLPHMYVIKDLVPDLSN ************************************
SDHB	FYKOYRSVKPYLORSTAAPDGR-EFROSKEDRKKLDGLYECILCACCSTSCPSYWWNOEE
1YQ3B	FYAQYKSIEPYLKKKDESKQGKEQYLQSIEDRQKLDGLYECILCACCSTSCPSYWWNGDK
1ZOYB	FYAQYKSIEPYLKKKDESQEGKQQYLQSIEEREKLDGLYECILCACCSTSCPSYWWNGDK ** **:*::***:. : :*: :: ** *:*:*********
SDHB	YLGPAVLLQSYRWIADSRDEKKAERQDALNNSMSLYRCRTILNCSRTCPKGLNPALAIAE
1YQ3B	YLGPAVLMQAYRWMIDSRDDYTEERLAQLQDPFSLYRCHTIMNCTRTCPKGLNPGKAIAE
1ZOYB	YLGPAVLMQAYRWMIDSRDDFTEERLAKLQDPFSLYRCHTIMNCTGTCPKGLNPGKAIAE ******:*:*:***: ****: ** *:: :*****:**:*
SDHB	IKKSMAFT
1Y03B	ΙΚΚΜΜΑΤΥΚΕΚΑΑΑΑ
1ZŐYB	IKKMMATYKEKKASV
	*** **
	Supplementary Figure S2 (B). Multiple sequence alignment of SDHB.
SDHC	MASQRVFQLGLRRAAAPGLRVQPVGRMVQRRLAATEHASQSEAAEILAKQRINRPVSPHL
1YQ3C	SRPLSPHI
1ZOYC	NRPLSPHI

SDHC	AIYRPOITWYASSLNRITGITLSGSLYLFGIAYLIAPYTGWHMETQSMVATVAAWPAAAK
1YQ3C	SIYKWSLPMAMSITHRGTGVALSLGVSLFSLAALLLPEQFPHYVAVVKSLSLSPALI
1ZOYC	TIYRWSLPMAMSICHRGTGIALSAGVSLFGLSALLLPGNFESHLELVKSLCLGPTLI :**: .: * :* **::** .: **.:: *: * : : : :
SDHC	AGLKAFYAFPFFFHSFNGLRHLSWDVGIGFKNQQVIRLGW
1YQ3C	YSAKFALVFPLSYHTWNGIRHLVWDMGKGFKLSQVEQSGVVVLILTLLSSAGIAAIS
1ZOYC	YTAKFGIVFPLMYHTWNGIRHLIWDLGKGLTIPOLTOSGVVVLILTVLSSVGLAAM- * .**: :*::**:*** **:* *:. *: : *
	Supplementary Figure S2 (C). Multiple sequence alignment of SDHC.

* *.* *.:*.* *: . : ** *. ::.: .* * *:: * *: *
* *.* *.:*.* *: . : ** *. ::.: .* * *:: * *: *
LTDSILCALLVVHSHIGFESCIIDYFPAKRVPKTRKAAMWALRAGTLTLGLALYSFETND
* ** ** ** * *
GSSKAASLHWTSERAVSALLLGLLPAAYLYPGP
ASSKAASLHWTGERVVSVLLLGLLPAAYLNPCS
QRTQILPPLPQKIIGTTNDPVPVPDPDYAHGSYHWSFERIVSAGLIPLTIAPFAAGSLNP
MASVMRPGLLRQACPPAHASQRMLSTATATTSRPLAQQLRPAFQRSAVPKATRTAAFHAT



Ramachandran Plot statistics

	No. of residues	%-tage
Most favored regions [A, B, L]	471	91.3%
Additional allowed regions [a,b,l,p]	43	8.3%
Generously allowed regions [~a,~b,~]	l,~p] 1	0.2%
Disallowed regions [XX]	1	0.2%*
Non-glycine and non-proline residues	516	100.0%
End-residues (excl. Gly and Pro)	1	
Glycine residues 61		
Proline residues 22	2	
Total number of residues 60	0	

Supplementary Figure S3 (A). Ramachandran plot of Subunit A



Ramachandran Plot statistics

		No. of re	sidues	%	6-tage	
Most favored regions	[A,B,L]		191	(91.0%	
Additional allowed regi	ons [a,b,l,p]		17	1	8.1%	
Generously allowed reg	gions [~a,~b,~]	l,~p]	2		1.0%	
Disallowed regions	[XX]		0	(0.0%	
Non-glycine and non-pr	roline residues	5	210		100.0%	
End-residues (excl. Gly	and Pro)	2				
Glycine residues		8				
Proline residues		15				
Total number of residue	es	235				
a .			-			0.0

Supplementary Figure S3 (B). Ramachandran plot of Subunit B



Ramachandran Plot statistics

	No. of residues	%-tage
Most favored regions [A,B,L]	78	97.5%
Additional allowed regions [a,b,l,p]	2	2.5%
Generously allowed regions [~a,~b,~l	,~p] 0	0.0%
Disallowed regions [XX]	0	0.0%
Non-glycine and non-proline residues	80	100.0%
End-residues (excl. Gly and Pro)	2	
Glycine residues	9	
Proline residues	3	
Total number of residues	94	

Supplementary Figure S3 (C). Ramachandran plot of Subunit C



	No. of residues	≫o-tage
Most favoured regions [A,B,L]	81	94.2%
Additional allowed regions [a,b,l,p) 4	4.7%
Generously allowed regions [~a,~t	o,~l,~p] 0	0.0%
Disallowed regions [XX]	1	1.2%*
Non-glycine and non-proline resid	ues 86	100.0%
End-residues (excl. Gly and Pro)	2	
Glycine residues	7	
Proline residues	6	
Total number of residues	101	

Supplementary Figure S3 (D). Ramachandran plot of Subunit D



Supplementary Figure S3 (E). Ramachandran plot of complete model of SDH

Program: ERRAT2 File: /var/www/SAVES/Jobs/2076096//errat.pdb Chain#:1 Overall quality factor**: 82.010



*On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value.

**Expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit. Good high resolution structures generally produce values around 95% or higher. For lower resolutions (2.5 to 3A) the average overall quality factor is around 91%.

Supplementary Figure S4. ERRAT scores for the A. solani SDH model



Supplementary Figure S5. (A) Superimposed view of 18 Fungicides (B) Pharmacophore model established by 18 Fungicides (SDH inhibitors). The pharmacophoric features H-bond acceptor, H-bond donor, and Hydrophobic features are depicted in light blue, pink, and Green spheres, respectively (C) Superimposed view of 18 fungicides on the pharmacophore model.







Supplementary Figure S6. The binding modes of 18 known SDHIs. The interacting residues are depicted in green stick and compounds in pink color. H-bonds are displayed in black dash lines.

1.1 Supplementary Tables

Subunit	Identity	PDB code	Organism	Protein name
SDH A	69%	1YQ3	Gallus gallus	Succinate dehydrogenase
	69%	1ZOY	Sus scrofa	Succinate:ubiquinoneoxidoreductase
	69%	4YTP	Sus scrofa	Succinate dehydrogenase
	62%	3VR8	Ascarissuum	Mitochondrial quinol-fumarate reductase
	62%	4YSX	Ascaris suum	Succinate dehydrogenase
	55%	1NEK	Escherichia coli	Succinate dehydrogenase
SDH B	69%	1YQ3	Gallus gallus	Succinate dehydrogenase
	67%	3ABV	Sus scrofa	Succinate dehydrogenase
	67%	1ZOY	Sus scrofa	Succinate:ubiquinoneoxidoreductase
	66%	3VR8	Ascarissuum	Quinol-fumarate reductase
	53%	1NEK	Escherichia coli	Succinate dehydrogenase
SDH C	37%	1ZOY	Susscrofa	Succinate:ubiquinoneoxidoreductase
	37%	2H88	Gallus gallus	Succinate dehydrogenase
	37%	1YQ3	Gallus gallus	Succinate dehydrogenase
	37%	2FLQ	Geobacillus stear otherm ophilus	Nitric Oxide Synthase
	36%	2FBW	Gallus gallus	Succinate dehydrogenase
	34%	1MJT	Staphylococcus aureus	Nitric-oxide synthase
	33%	2Y5B	Homo sapiens	Ubiquitin carboxyl-terminal hydrolase
	33%	3I3T	Homo sapiens	Ubiquitin carboxyl-terminal hydrolase
	33%	3MTN	Homo sapiens	Ubiquitin carboxyl-terminal hydrolase
	31%	3VR8	Ascarissuum	Quinol-fumarate reductase
SDH D	42%	2EOB	Rattusnorvegicus	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase
	36%	1ZOY	Susscrofa	Succinate:ubiquinoneoxidoreductase
	35%	1YQ3	Gallus gallus	Succinate dehydrogenase
	33%	1NP7	Synechocystissp	DNA photolyase
	31%	2JFK	Homo sapiens	Fatty acid synthase
	31%	2JFD	Homo sapiens	Fatty acid synthase
	31%	3HHD	Homo sapiens	Fatty acid synthase

Supplementary Table S1. Percent identities between A. solani SDHs and template sequences

Sr.	Name	Chemical structure	Docking Scores	Sr.	Name	Chemical structure	Docking Scores
1	Benodanil	O NH	-9.45	2	Benzovindiflupyr		-9.69
3	Bixafen		-9.87	4	Boscalid		-10.47
5	Carboxin	S O N	-10.24	6	Fenfuram	HN O	-9.02
7	Fluopyram	$CF_3 O CI CF_3$ N N N	-10.02	8	Flutolanil	O NH F ₃ C	-9.74
9	Fluxapyroxad		-9.96	10	Furametpyr		-10.63
11	Isofetamid	L O H S	-10.46	12	Isopyrazam	F F N N N N H	-9.57

Supplementary Table S2. Chemical structure and Docking Scores of 18 known SDHIs (fungicides).

13	Mepronil	HN O	-10.51	14	Oxycarboxin		-11.07
15	Penflufen		-9.56	16	Penthiopyrad	HN-CF3	-9.46
17	Sedaxane	F N N F H N O	-10.13	18	Thifluzamide	$\begin{array}{c} F_{3}C \\ N \\ S \\ S \\ S \\ Br \end{array} \xrightarrow{O} CF_{3} \\ CF_{3}$	-11.67

Supplementary Table S3. Chemical Structures and Docking Score of 25 compounds selected from ZINC database after docking, rescoring and interaction analysis

No.	ZINC ID	Compound	Structure	ADT Seeme	MOE
		Coue		Score	Score
1	ZINC00781454	C1		-8.74	-5.2243
2	ZINC13037761		F S S S S S S S S S S S S S S S S S S S	-8.25	-5.7331
3	ZINC78502784			-8.23	-5.2939
4	ZINC19837471			-8.22	-5.2954
5	ZINC50812116			-8.20	-4.4908

6	ZINC04498541	C2		-8.20	-5.2388
7	ZINC34378628	C12		-8.18	-5.1025
8	ZINC27260697	C3		-8.10	-4.1538
9	ZINC00795936		HO HN OCH ₃	-8.00	-5.2242
10	ZINC32180696	C11	O NH O F	-8.00	-4.6726
11	ZINC36624363			-8.00	-4.6285

12	ZINC17007371	C10	CN	-7.98	-4.4015
13	ZINC06512448	CQ	tBu	_7.93	-// 839/
15	21110000012440			-1.55	-4.0374
14	ZINC04731427			-7.87	-5.0271
15	ZINC32803988		N HN NC	-7.85	-4.3761
16	ZINC02409017	C4		-7.74	-4.5404
17	ZINC03102709	C6	Br NO HNO OH	-7.74	4.95920

1	.8	ZINC06183077	C8		-7.69	-3.4910
1	9	ZINC00689200	C5		-7.68	-3.7313
2	20	ZINC01436920			-7.68	-4.9843
2	21	ZINC18046772			-7.66	-4.4454
2	22	ZINC04903387		H ₃ CO HN O O O O O O O O H N O O O H	-7.63	-5.2351

23	ZINC05176439	C7	HO HO HN HN HN	-7.61	-4.3706
24	ZINC00707182			-7.6	-3.5623
25	ZINC12196415			-7.60	-4.3058