

Discovery of Novel Acetylcholinesterase Inhibitors through AI-Powered Structure Prediction and High-Performance Computing-Enhanced Virtual Screening (Supporting Information)

Beatriz Chafer-Dolz,^{*,†} José M. Cecilia,[‡] Baldomero Imbernón,[¶] Estrella
Núñez-Delicado,[¶] Victor Casaña,[†] José M. Cecilia,[§] and José P.
Cerón-Carrasco^{*,§}

[†]*Bio Logic Crop Science, S.L. Amadeo de Saboya 1-4, 46010, Valencia, Spain*

[‡]*Universitat Politècnica de Valencia (UPV). Camino de Vera S/N, 46022, Valencia, Spain.*

[¶]*Universidad Católica San Antonio de Murcia (UCAM). Campus de los Jerónimos, 30107,
Murcia, Spain*

[§]*Centro Universitario de la Defensa, Academia General del Aire. Universidad Politécnica
de Cartagena. C/ Coronel López Peña s/n 30729. Santiago de la Ribera, Murcia, Spain*

E-mail: bchafer@biologiccropscience.com; jose.ceron@tud.upct.es

Table S-1: Parameters for *METADOCK 2* adapted for AChE virtual screening

Metaheuristic Parameters	Description
ParamIni	<p>INEIni Initial count of ligand conformations.</p> <p>IIEFlex Enhancement of flexibility in the initial improvement functions.</p> <p>PEIIni Proportion of top conformations refined during initialization.</p> <p>IIEIni Enhancement of initial improvement functions.</p> <p>PBEIni Fraction of top conformations retained for subsequent iterations.</p> <p>PWEIni Fraction of lower-quality conformations carried to the next iteration phase.</p>
ParamSel	<p>PBESel Selects a percentage of the best conformations for combination.</p> <p>PWESel Selects a percentage of the poorest conformations for combination.</p>
ParamCom	<p>PBBCom Percentage for combining the best-ranked conformations.</p> <p>PWWCom Percentage for combining the worst-ranked conformations.</p> <p>PBWCom Percentage for combining the best and worst conformations.</p>
ParamMut	<p>PMUCom Proportion of top combinations subjected to mutation.</p> <p>IMUCom Degree of mutation applied to combined elements.</p>
ParamImp	<p>PEIImp Proportion of top combined conformations to be enhanced.</p>
ParamInc	<p>IIEImp Degree of enhancement applied to elements after combination.</p>
ParamEnd	<p>PBEInc Proportion of the best conformations included in the reference set.</p> <p>NIREnd Limit on steps without improvement before termination.</p> <p>MNIEnd Maximum number of iterations allowed, irrespective of improvement.</p>

Table S-2: Configurations of METADOCK 2 utilized in the evaluation, labeled from M1 to M4

	<i>Configurations</i>			
	M1	M2	M3	M4
INEIni	128	128	256	1024
IIEFlex	20	30	40	50
PEIIni	100	100	50	50
IIEIni	50	150	200	200
PBEIni	100	50	50	100
PWEIni	0	50	50	0
PBESel	50	50	80	100
PWESel	50	50	20	0
PBBCom	30	20	100	15
PWWCom	0	20	10	0
PBWCom	0	20	10	0
PMUCom	20	0	10	0
IMUCom	50	0	30	0
PEIImp	50	50	100	100
IIEImp	200	100	100	50
PBEInc	50	90	50	90
NIREnd	4	4	4	5
MNIEnd	4	4	4	5

The seminal AChE within insecticide framework was originally published more than 20 years ago for the fruit fly *Drosophila melanogaster*. The work conducted by Sussman and coworkers resolved both for the three-dimensional structures of AChE and of its complexes with two potent inhibitors (*Protein Sci.* 2000, **9**, 1063). It was shown that the insect AChEs enzyme are rather similar to that of vertebrate but with a smaller active-site size. More recent X-ray structures are available now in the literature. An excellent example is the crystal deposited with PDB code 6XYS entitled “Update of native acetylcholinesterase from *Drosophila Melanogaster*” (*Molecules* 2020, **25**, 1198). Modelled AChE structure for cockroach has been compared that recent experimental AChE from *Drosophila melanogaster*. Their overall structures are very similar, with a RMSD value of 1.49 Å only. That agreement further confirms the conservation of the AChE in insects.

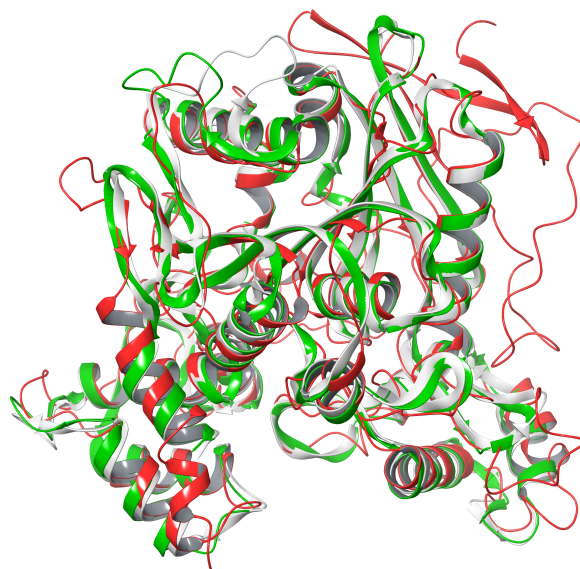


Figure S-1: Overlays for the generated AChE target models displayed as cartoons. Color scheme: Homology model in green, AlphaFold structure in white and experimental AChE of in red.

Protein Structure Alignment Results: AlphaFold vs. Homology

```

.....+.....+.....+.....
AlpaFold  SSA 13139   1 CCCCCCEEEEECCCCCCCCCEEEEECECCCCCGCCCCC
Homology   SSA 2      1 CCCCCCEEEEECCCCCCCCCEEEEECECCCCCGCCCCC
AlpaFold           13139   1 PLSEDCLYINVVAPKPRPRNAAMVWIFGGGFYSGSATLD
Homology           2      1 PLSEDCLYINVVAPKPRPRNAAMVWIFGGGFYSGSATLD

.....+.....+.....+.....
AlpaFold  SSA 13139   41 CCCHHHHHHHHCCEEEEECCCCCCCCCCCCCCCCCHH
Homology   SSA 2      41 CCCHHHHHHHHCCEEEEECCCCCCCCCCCCCCCCCHH
AlpaFold           13139   41 VYDHKTLVSEENVIIIVSMQYRVASLGFLFFDTGDVPGNAG
Homology           2      41 VYDHKTLVSEENVIIIVSMQYRVASLGFLFFDTGDVPGNAG

.....+.....+.....+.....
AlpaFold  SSA 13139   81 HHHHHHHHHHHHCCEEEEECCCCCCCCCEEEEECHHHHHHHHH
Homology   SSA 2      81 HHHHHHHHHHHHCCEEEEECCCCCCCCCEEEEECHHHHHHHHH
AlpaFold           13139   81 LFDQLMALQWIHDNIQAFGGNPNNVTLFGESAGAVSVSLH
Homology           2      81 LFDQLMALQWIHDNIQAFGGNPNNVTLFGESAGAVSVSLH

.....+.....+.....+.....
AlpaFold  SSA 13139  121 HHHHCCCCCCEEEEECCCCCCCCCECHHHHHHHHHHH
Homology   SSA 2      121 HHHHCCCCCCEEEEECCCCCCCCCECHHHHHHHHHHH
AlpaFold           13139  121 LLSPLSRNLFSQAIMESGSPTAPWAIISREESILRGLRLA
Homology           2      121 LLSPLSRNLFSQAIMESGSPTAPWAIISREESILRGLRLA

.....+.....+.....+.....
AlpaFold  SSA 13139  161 HHHCCCCCCHHHHHHHHHHHCHHHHHHHCCC-CCCCC

```

Homology	SSA	2	161	HHHCCCCCCCCCHHHHHHHHHHHHHCHHHHHHHHCCCC-CCCC
AlpaFold		13139	161	EAVGCPSSRSDIRAVIDCLRKKNATDLVNNEWG-TLGICE
Homology		2	161	EAVGCPSSRSDIRAVIDCLRKKNATDLVNNEWGTL-GICE
.....+......+......+......				
AlpaFold	SSA	13139	200	CCCCCCCCCCCCCCHHHHHHHCCCCCEEEEEEECCCCHH
Homology	SSA	2	200	CCCCCCCCCCCCCCHHHHHHHCCCCCEEEEEEECCCCHH
AlpaFold		13139	200	FPFVPIIDGAILDGPQRSLAEKNFKKTNILMGSNTEEGY
Homology		2	200	FPFVPIIDGAILDGPQRSLAEKNFKKTNILMGSNTEEGY
.....+......+......+......				
AlpaFold	SSA	13139	240	HHHHHHHCCCCCCCCCCHHHHHHHHHHHHHCCCCCHHHH
Homology	SSA	2	240	HHHHHHHCCCCCCCCCCHHHHHHHHHHHHHCCCCCHHHH
AlpaFold		13139	240	YFIIYYLTELFRKEENVYVNREEFLRSVRELNPYVNNVAR
Homology		2	240	YFIIYYLTELFRKEENVYVNREEFLRSVRELNPYVNNVAR
.....+......+......+......				
AlpaFold	SSA	13139	280	HHHHHHHCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHH
Homology	SSA	2	280	HHHHHHHCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHH
AlpaFold		13139	280	QAIVFEYTDWLNPDPIRNRDALDKMVG DYQFTCNVNEFA
Homology		2	280	QAIVFEYTDWLNPDPIRNRDALDKMVG DYQFTCNVNEFA
.....+......+......+......				
AlpaFold	SSA	13139	320	HHHHHHCCEEEEEEECCCCCCCCCCCCCCCCCCHHHHHHH
Homology	SSA	2	320	HHHHHHCCEEEEEEECCCCCCCCCCCCCCCCCCHHHHHHH
AlpaFold		13139	320	HRYAETGNNVYMYFKHRSAGNPWPSWTGVMHGDEIN YVF
Homology		2	320	HRYAETGNNVYMYFKHRSAGNPWPSWTGVMHGDEIN YVF

.....+......+......+

AlpaFold SSA 13139 360 CCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHCCCEEECCC
 Homology SSA 2 360 CCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHCCCCC--C
 AlpaFold 13139 360 GEPLNPAKNYQPHEIELSKRMMRYWANFAKTGNPSMEDG
 Homology 2 360 GEPLNPAKNYQPHEIELSKRMMRYWANFAKTGNPSMS--E

.....+......+......+

AlpaFold SSA 13139 400 ECCC--CCCCCECCCCCEEEEECECEEECCCHHHHHH
 Homology SSA 2 398 CCCCCCCCCCCCCCCCCCEEEEECCCCCCCCCCHHHHHH
 AlpaFold 13139 400 TWTA--TYWPVHTAYGREYLTLDVNSTETGRGPRLKQCAF
 Homology 2 398 DGTWTATYWPVHTAYGREYLTLDVNSTETGRGPRLKQCAF

.....+......+......+

AlpaFold SSA 13139 438 HHH
 Homology SSA 2 438 HHH
 AlpaFold 13139 438 WKN
 Homology 2 438 WKN

RMSD: 0.89 Angstrom

Protein Structure Alignment Results: AlphaFold vs. PDB=6XYS

.....+.....+.....+.....

AlpaFold	SSA	13139	-----
6XYS	SSA	45372	3 CCEECCCCCEEECEEEEECEEEEEEEEEEECCCECCCC
AlpaFold		13139	-----
6XYS		45372	3 DRLVVQTSSGPVGRSVTVQGREVHVYTGIPYAKPPVEDL

.....+.....+.....+.....

AlpaFold	SSA	13139	-----
6XYS	SSA	45372	43 CCCCCCCCCCCCCCECCCCCECCCCCCCCCCHHHHH
AlpaFold		13139	-----
6XYS		45372	43 RFRKPVPAEPWHGVLDATGLSATCVQERYEYFPGFSGEEI

.....+.....+.....+.....

AlpaFold	SSA	13139	1 -----CCCCCCEEEEECCCCCCEEEEEEECCCCCCC
6XYS	SSA	45372	83 HHHCCCCCCCCCEEEEEEECC---CCEEEEEEECCCCCCC
AlpaFold		13139	1 -----PLSEDCLYINVVAPKPRPRNAAVMVWIFGGGFYSG
6XYS		45372	83 WNPNTNVSEDCLYINWAPA---NGLPILIWYGGGFMTG

.....+.....+.....+.....

AlpaFold	SSA	13139	36 CCCCCCCHHHHHHHHCCEEEEEEECCCCCCCCC-C---
6XYS	SSA	45372	156 CCCCCCCHHHHHHHHCCEEEEEEECCCCCCCCCCCCCCC
AlpaFold		13139	36 SATLDVYDHKTLVSEENVIIIVSMQYRVASLGFLLFF-D---
6XYS		45372	156 SATLDIYNADIMAAVGNVIVASFQYRVGAFGLHLAPEMP

.....+.....+.....+.....

AlpaFold	SSA	13139	72 --C-CCCCCCHHHHHHHHHHHHHHHCCCCCCCCCEEEEE
----------	-----	-------	--

6XYS	SSA	45372	196	CCCCCCCCCHHHHHHHHHHHHHHHHHCCCCCCCCCEEEEE
AlpaFold		13139	72	--T-GDVPGNAGLFDQLMALQWIHDNIQAFGGNPNNVTLF
6XYS		45372	196	SEFAEEAPGNVGLWDQALAIRWLKDNAHAFGGNPEWMTLF
.....+......+......+......				
AlpaFold	SSA	13139	109	ECCHHHHHHHHHHHHHCCCCCCCCCEEEEECCCCCCCCCEC
6XYS	SSA	45372	236	ECCHHHHHHHHHHHHHCCCCCCCCCEEEEECCCCCCCCCEC
AlpaFold		13139	109	GESAGAVSVSLHLLSPLSRNLFSQAIMESGSPTAPWAIIS
6XYS		45372	236	GESAGSSSVNAQLMSPVTRGLVKRGMMQSGTMNAPWSHMT
.....+......+......+......				
AlpaFold	SSA	13139	149	HHHHHHHHHHHHHHCCCCCCH--H-HHHHHHHHHCH
6XYS	SSA	45372	276	HHHHHHHHHHHHCCCCCCH--CCCCCCHHHHHHHCH
AlpaFold		13139	149	REESILRGLRLAEAVGCPSSRSI---R-AVIDCLRKNA
6XYS		45372	276	SEKAVEIGKALINDCNCNAS--MLKTNPAHVMSCMRSVDA
.....+......+......+......				
AlpaFold	SSA	13139	185	HHHHHHCC-CCCCCCCCCCCCCCCCCCCCCHHHHHHHCC
6XYS	SSA	45372	314	HHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHH-H
AlpaFold		13139	185	TDLVNNEW-GTLGICEFPFVPIIDGAILDGPPQRSIAEKN
6XYS		45372	314	KTISVQQWNSYSGILSFPSAPTIDGAFLPADPMTLMKT-A
.....+......+......+......				
AlpaFold	SSA	13139	224	-CCC-EEEEEECCCHHHHHHHHHCCCCCCCCCCHHH
6XYS	SSA	45372	353	CC-CCEEEEEEECCCHHHHHHHHHCCCCCCCCCCHHH
AlpaFold		13139	224	-FKK-TNILMGSNTEEGYYFIIYYLTELFRKEENVYVNRE
6XYS		45372	353	DL-KDYDILMGNVRDEGTYFLLYDFIDYFDKDDATALPRD

.....+.....+.....+.....

AlpaFold	SSA	13139	262	HHHHHHHHHCCCCCHHHHHHHHHHHHCCC-CCCCHHHHHH
6XYS	SSA	45372	392	HHHHHHHHHCCCCCHHHHHHHHHHHHCCCCC-C-HHHHHH
AlpaFold		13139	262	EFLRSVRELNPYVNNVARQAIVFEYTDWL-NPDDPIRNRD
6XYS		45372	392	KYLEIMNIFGKATQAEREAIIIFQYTSWEGN-P-GYQNQQ

.....+.....+.....+.....

AlpaFold	SSA	13139	301	HHHHHHHHHCHHHHHHHHHHHHHHCCEEEEEEECCCCC
6XYS	SSA	45372	430	HHHHHHHHHCHHHHHHHHHHHHHHCCEEEEEEECCCCC
AlpaFold		13139	301	ALDKMVG DYQFTCNVNEFAHRYAETGNNVYMYF KHR SAG
6XYS		45372	430	QIGRAVGDHFFTCPTNEYA QAL AERGASVHYYYFTHRST

.....+.....+.....+.....

AlpaFold	SSA	13139	341	CCCCCCCCCCHHHHHHHCCCCCCCCCCHHHHHHHHHH
6XYS	SSA	45372	470	CCCCCCCCCCHHHHHHHCCCCCCCCCCHHHHHHHHHH
AlpaFold		13139	341	NPWPSWTGVMHGDEIN YVFG EPLNPAKNYQPHEIELSKRM
6XYS		45372	470	SLWGEWMGVLHGDEIEYFFGQPLNNSLQYRPVERELGKRM

.....+.....+.....+.....

AlpaFold	SSA	13139	381	HHHHHHHHHCCCCEEECCCECCCCCCECCCCEEEEE
6XYS	SSA	45372	510	HHHHHHHHHCCCCC-C---C--CCCCCCECCECEEEEC
AlpaFold		13139	381	MRYWANFAKTGNPMSSEDGTWTATYWPVHTAYGREYLTLD
6XYS		45372	510	LSAVIEFAKTGNPA-Q---D--GEEWPNFSKEDPVYYIFS

.....+.....+.....+.....

AlpaFold	SSA	13139	421	CC---ECEECC-CCHHHHHHHH-----
----------	-----	-------	-----	-----------------------------

6XYS SSA 45372 544 CCCCCCE-EEHHHH-HHHHHHHHHHHHHHHHC
AlpaFold 13139 421 VN---STETGRG-PRLKQCAFWKN-----
6XYS 45372 544 TDDKIEK-LARGPL-AARCSFWNDYLPKVRWS

Alignment RMSD: 1.49 Angstrom