Supplementary Information for

Role of a Remote Leucine Residue in the Catalytic Function of Polyol Dehydrogenase

Manish Kumar Tiwari,^{a,‡} Vipin C. Kalia,^c Yun Chan Kang,*,^b Jung-Kul Lee*,^a

^{*a*} Department of Chemical Engineering, 1 Hwayang–Dong, Gwangjin–Gu, Seoul 143–701, Republic of Korea. ^{*b*} Department of Materials Science and Engineering, Korea University, Anam–Dong, Seongbuk–Gu, Seoul 136–713, Republic of Korea. ^{*c*} Microbial Biotechnology and Genomics, CSIR-Institute of Genomics and Integrative Biology, Delhi University Campus, Mall Road, Delhi-110007, India.

* To whom correspondence should be addressed. E-mail: jkrhee@konkuk.ac.kr; Fax: +82-2-458-3504. E-mail: yckang@korea.ac.kr; Fax: +82-2-928-3584.

‡ Current address: Department of Chemistry, Technical University of Denmark, Kongens Lyngby 2800, Denmark.

	68	143
gi 85074831	M I V E C D H V L G H E S A G E V I A V H P S V K S I K V G I	DRVA I EP QV I CNA CEP CL TG RY NG CERVD FL ST PP V P GL L R RY V N
gi 336275479		D R V A I E P Q V I C N A C E P C L T G R Y N G C E R V D F L S T P P V P G L L R R Y V N
gi 347003842	MIVGGDHILGHESAGEIIAVHPSVKNLKVG	DRVAVEPOVICNT CEPCLTGRYNGCETVDFLSTPPVPGLLRRYVN
gi 116182132	MIVGCDHILGHESAGEIVAVHPGVTSLKVG	D R V A V E P Q V I C N A C E P C L T G R Y N G C E A V D F L S T P P V P G L L R R Y V N
gi 302409940	M I <mark>V</mark> E G D H I L <mark>G H E S</mark> A <mark>G</mark> D V I A V H P S V E H L K V <mark>G I</mark>	<mark>D R V A I E</mark> P N V I <mark>C</mark> N A <mark>C</mark> E P <mark>C</mark> L <mark>T G</mark> R Y N G <mark>C</mark> E K V E <mark>F</mark> L S T P <mark>P V P G L</mark> L R R Y V N
gi 39950617	M I V R E D H I L <mark>G H E S A G</mark> E I I A V H P S V T S L K V <mark>G</mark> I	D R V A V E P Q V I C Y E C E P C L T G R Y N G C E K V D F L S T P <mark>P V P G L</mark> L R R Y V N
gi 310790030	MIVDGDHILGHESAGEILAVHPSVTHLKVG	D R V A V E P N V I C N A C E P C L T G R Y N G C E Q V Q F L S T P P V P G L L R R Y V N
gi 171695040		DRVATEPQVICNECEPCLTGRYNGCEKVDFLSTPPVAGLLRRYVN
gi 15811375	MIVEGDHILGHESAGEVIAVHPTVSSLOIG	DRVALEPNIICNA CEPCLIGRYNGCEKVEELSTPPVPGLIRRYVN
gi 154291438	MIVEDTHILGHESAGVVLAVHPSVDSLKVG	D R V A V E P N I I C G E C E R C L T G R Y N G C E K V L F L S T P P V P G L L R R Y V N
gi 156054222	M I V E D T H I L G H E S A G V V L A V H P S V D S L K V G I	D R V A V E P N I I C G E C E R C L T G R Y N G C E K V L F L S T P P V P G L L R R Y V N
gi 322712084	M I V E G D H I L <mark>G H E S A G</mark> D V I A V H P S V T S L K V <mark>G I</mark>	D R V A V E P N I I <mark>C</mark> N A - - <mark>C</mark> E P <mark>C L T G</mark> H Y N G <mark>C</mark> E N V A <mark>F</mark> L S T P <mark>P</mark> V P G <mark>L</mark> L R R Y V N
gi 322695931	M I V E G D H I L <mark>G H E S A G</mark> D V I A V H P S V T N L K V <mark>G I</mark>	D R V A V E P N I I C N A - - <mark>C</mark> E P <mark>C L T G</mark> H Y N G <mark>C</mark> E K V A <mark>F</mark> L S T P P V P G <mark>L</mark> L R R Y V N
gi 67517338	MIVTGDHILGHESAGDVIAVAPDVTSLKVGI	D R V A I E P N V I C N A C E P C L T G R Y N G C E K V A F L S T P P V D G L L R R Y V N
gi 242774074	MIVNGDHILGHESAGVIVAVGPDVNNLKVG	DRIAVEPNIICNK CEPCLTGRYNGCENVEFLSTPPIDGLLRRYVN
gi 109622119		DRVATEPNVICHE CEPCLIGRINGCERVQFLSTPPVIGLIRRILR
gil 212531837		DRVAVEPNVICNK - CEPCLIGRYNGCESVEELSTPPVDGLLRRYVN
gi 312217694	MIVEDTHVLGHESAGTVLAVHPSVTSLKAG	DRVAIEPNVICHE CEPCLTGRYNGCERVQFLSTPPVTGLLRRYLK
gi 121700771	MIVEGDHILGHESAGQVIAVASD VTTLKPG	D R V A I E P N I I C N E C E P C L T G R Y N G C E K V A F L S T P P V D G L L R R Y V N
gi 70996476	M I V E G D H I L <mark>G H E S A G</mark> Q V I A V A P D V T S L K P <mark>G I</mark>	D R V A I E P N I P C H A C E P C L T G R Y N G C L N V A F L S T P P V D G L L R R Y V N
gi 119494479	M I V E G D H I L <mark>G H E S A G</mark> Q V I A V A P D V T S L K P <mark>G I</mark>	D R V A I E P N I P <mark>C</mark> H A - - <mark>C</mark> E P <mark>C L T G</mark> R Y N G <mark>C L N V A F</mark> L S T P <mark>P</mark> V D G <mark>L</mark> L R R Y V N
gi 145230401	M I V T G D H I L <mark>G H E S A G</mark> Q V V A V A P D V T S L K P <mark>G I</mark>	D R V A V E P N I I C N A - - <mark>C</mark> E P <mark>C L T G</mark> R Y N G <mark>C</mark> E N V Q F L S T P P V D G <mark>L</mark> L R R Y V N
gi 255957067	MIVTGDHVLGHESAGQVLAVAPDVTHLKVGI	D R V A V E P N V I C N A C E P C L T G R Y N G C V N V A F L S T P P V D G L L R R Y V N
gi 169767170	MIVTGDHILGHESAGEVIAVASDVTHLKPG	D R V A V E P N I P C H A C E P C L T G R Y N G C E K V L F L S T P P V D G L L R R Y V N
gi1330935729		DRVATEPNVVCHACEPCLIGRINGCARVAFLSTPPVDGELKRIVH
gil 115397525		DRVALEPNIICNA CEPCLTGRYNGCERVAFLSTPPVDGLLRRYVN
gil 189197563	MIVEDTHVLGHESAGTVMAVHPSVTNLKPG	DRVAIEPNIICGE CEPCLTGRYNGCERVLFLSTPPVTGLLRRYLK
gi 339472744	MIVEGEHILGHESAGTVVAVHPSVTTHQIG	D R V A I E P N I I C N E C E P C L T G K Y N G C E S V Q F R S T P P I P G L L R R Y V N
gi 320593501	M V V T G D H V L <mark>G H E S A G</mark> E V V A V H A D G A K D L T G T T L K V <mark>G I</mark>	<mark>D R V A I E</mark> P N V I <mark>C</mark> G A <mark>C</mark> T P <mark>C L T G R Y N G C E R V Q F L S T P P</mark> V D G <mark>L</mark> L R R Y V N
gi 261191582	MV <mark>V</mark> T D N H I L <mark>G H E S A G</mark> T I L A V S P E V T S L K P <mark>G I</mark>	D R V A I E P N I I <mark>C</mark> N E - - <mark>C</mark> E P <mark>C L T G</mark> R Y N G <mark>C</mark> E H V R <mark>F</mark> L S T P <mark>P</mark> V D G <mark>L</mark> L R R Y V N
gi 350632917	W	D R V A I E P H I V C K A - - <mark>C E P C L T G R Y N G C K N L Q F R S S P P S H G L</mark> L R T Y V N
gi 317028626	W	D R V A I E P H I V C K A C E P C L T G R Y N G C K N L Q F R S S P P S H G L L R T Y V N
gi 134058085	WVVNDAHILGHESAGLVVKVHPSVTTLAVG	DRVAIEPHIVCKACEPCLTGRYNGCKNLQFRSSPPSHGLLRTYVN
gi1328860979		DRVATEPTTPCAK CVPCLIGRTNGCEDVLFRSTPPVPGLKRTTP
gil 31087950		DRVALETGIPCSKPTCEMCTGOYNACPEIMEWETSPYHOLMTRYHA
gi 331242651	MVVKHECGAGHESAGEVIALGEG VTDLQVG	D R V A I E A G I P C S K P T C D M C R T G Q Y N A C P E I I F C S T P P Y H G L M T R Y H A
gi 331243999	MVVKHECGAGHESAGEIIGVGEGVADVKVG	D R V A I E A G V P C S K P T C E M C R T G R Y N A C P D V V F F S T P P Y H G L L T R F H A
gi 331242635	MVVRDECGA <mark>GHESAG</mark> EVVELGEGVTDLQI <mark>G</mark>	D R V A I E A G V P <mark>C</mark> S K <mark>P T C</mark> E K <mark>C</mark> R <mark>T G</mark> C Y N A <mark>C</mark> P Q M I <mark>F</mark> F S T P <mark>P</mark> F H <mark>G L</mark> L T R Y H A
gi 296421171	***************************************	N V R - E H V L <mark>F</mark> L S T P <mark>P</mark> V S <mark>G L</mark> L R R Y V T
gi 331244001	Q	<mark>D R V A I E</mark> A G <mark>V P C</mark> S K <mark>P T C</mark> E M <mark>C</mark> R <mark>T G</mark> R Y N A <mark>C</mark> P D V V <mark>F</mark> F S T P <mark>P</mark> Y H <mark>G L L</mark> T R Y H A

Fig. S1 Multiple sequence alignment of polyol dehydrogenase enzymes. Multiple sequence alignment was carried out using Discovery Studio 3.5. Protein sequence of 45 members of the polyol dehydrogenase family was aligned (only amino acid residues from 68 to 143 are shown; sequence number is according to NcLAD protein sequence). The target amino acid residue (L136; PDB ID 3M6I) is indicated with orange background. The conserved residues are marked with yellow background. Gene IDs are given on the left side.



Fig. S2 Effect of point mutations on the secondary structure. Expression (left panel) and purification (right panel) of NcLAD wild-type and selected mutants. Lane M contains the protein markers and lanes WT, L136P, L136A, L136C, L136H, L136E, L136K and L136W correspond to the purified mutant enzymes of NcLAD with a molecular weight of ~41 kDa.



Fig. S3 The effect of point mutations on the globular structure and secondary structure elements of NcLAD. (A) NcLAD wild-type PDB coordinates (red color) as a reference structure, superimposed with four mutants L136A (grey color), L136C (yellow color), L136H (blue color), and L136P (magenta color). (B) Secondary structure element cartoon predicted using Kabsch and Sander's method represents the altered secondary structure elements in the mutants (L136H and L136P) of NcLAD. The secondary structure elements of NcLAD are color coded, with helices in red, strands in blue, and coils in beige.



Fig. S4 An orthogonal top-view of the NcLAD wild-type structure (ribbon diagram), showing the distribution of hydrophobicity at the substrate-binding pocket (5Å from bound L-arabinitol). The solvent hydrophobic surface was calculated with a probe radius 1.4 Å. Surface colors range from blue (negative hydrophobicity) to brown (positive hydrophobicity). L-arabinitol is shown in the stick model with yellow color carbon.



Fig. S5 Correlation between relative activities and free energies of binding for five (wild-type, L136C, L136A, L136H, and L136P) molecular patterns (enzyme-substrate complex) discussed in this study.

NcLAD/HjLAD	Relative activity (%) of NcLAD/HjLAD
Wild-type	100/100
L136A/L149A	44/43
L136C/L149C	72/80
L136I/L149I	39/47
L136M/L149M	38/46
L136V/L149V	29/21
L136W/L149W	1.8/NI
L136K/L149K	1.2/NI
L136E/L149E	0.32/NI
L136H/L149H	0.70/0.30
L136P/L149P	0.00/0.00

Table S1 Relative activities of NcLAD and HjLAD wild-type enzymes and their variants

NI- Not investigated

Structure	RMSD (Å) based on C-alpha pairs
Wild-type	0.00
L136A	0.50
L136C	0.63
L136H	1.50
L136P	1.60

 Table S2 RMSD values for globular geometry of the four superimposed structures of NcLAD