# Binding and Channelling of Alternative Substrates in the Enzyme DmpFG: A Molecular Dynamics Study

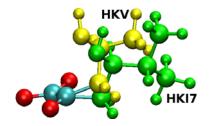
†,‡†‡Natalie E. Smith,Alice Vrielink,Paul V. Attwood,and Ben Corry.

†School of Chemistry and Biochemistry, University of Western Australia.

‡Research School of Biology, Australian National University.

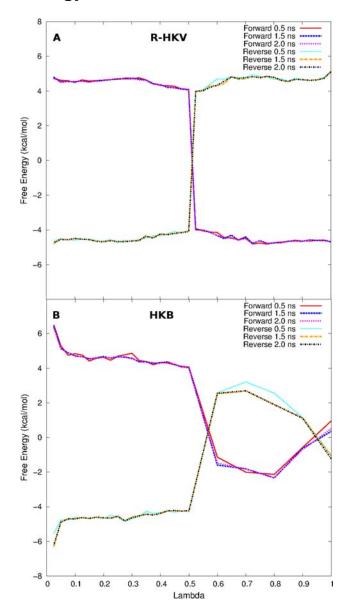
### **Supplementary Material**

Additional figure to show an example of a dual topology hydrid molecule:



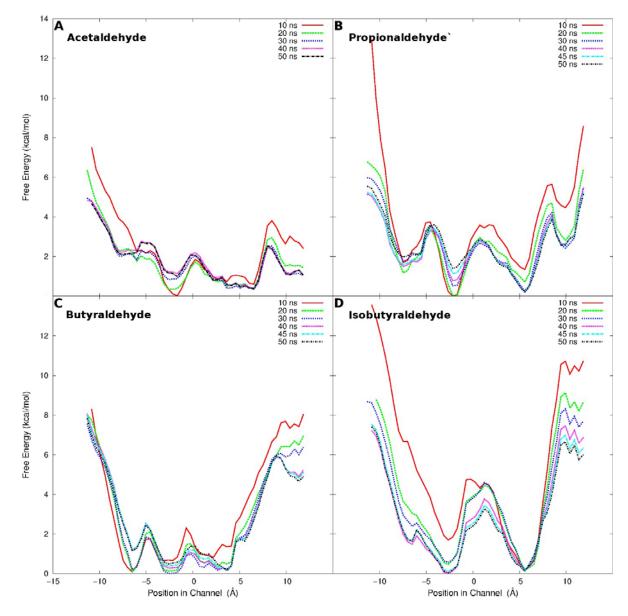
**Fig. S1:** Dual topology hybrid of HKV (yellow side chain) and HKI7 (green side chain) utilised in the FEP calculations.

Additional figure to demonstrate the convergence of free energy values obtained from Free Energy Perturbation simulations:



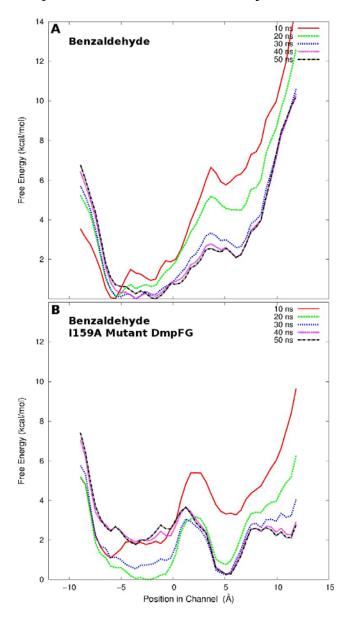
**Fig. S2:** The convergence of the free energy values obtained in each lambda window from the free energy perturbation simulations for both R-HKV relative to S-HKV (A), and HKB relative to S-HKV (B), in protein with time. Both the forward and reverse simulations are shown.

### Additional figure to demonstrate the convergence of the free energy profiles obtained in this study:

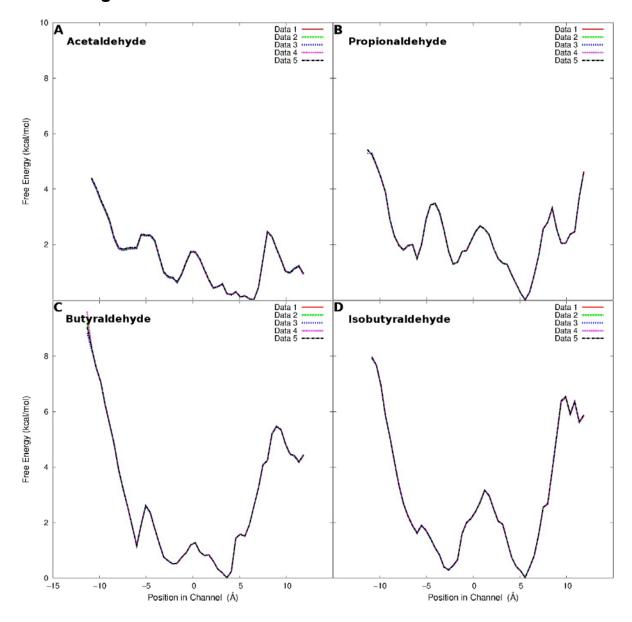


**Fig. S3:** Convergence of each free energy profile with time for acetaldehyde (A), propionaldehyde (B), butyraldehyde (C) and isobutyraldehyde (D).

Additional figure to demonstrate the convergence of the free energy profiles for benzaldehyde obtained in this study:

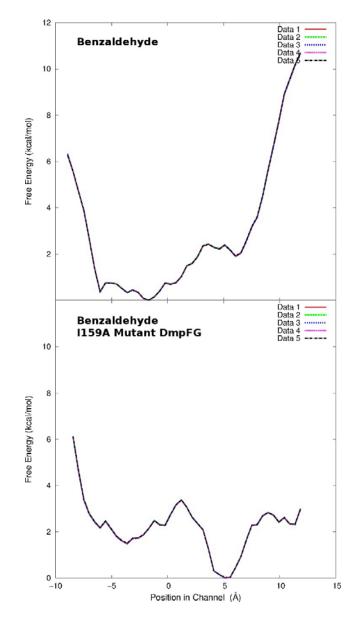


**Fig. S4:** Convergence of the free energy profile of benzaldehyde in WT (A) and I159A mutant (B) DmpFG with time.



Additional figure to determine the statistical error in each data set:

**Fig. S5:** The free energy profiles obtained when each complete data set (50 ns) was randomly shuffled and split into 5 separate data sets. A free energy profile was obtained for each data set allowing a standard error to be determined for each aldehyde.



### Additional figure to determine the statistical error in each data set:

**Fig. S6:** The free energy profiles obtained when each complete benzaldehyde data set (50 ns) was randomly shuffled and split into 5 separate data sets. A free energy profile was obtained for each data set allowing a standard error to be determined.

## Additional figure to show the sequence alignments obtained for the aldolase DmpG and its orthologues:

		*	A. A.				
DmpG Bphl TTHB246 HsaF	MTFNPSKK MKLE.GKK MSWDLSTAKF MTDMWD	LYISDVTLR VTVHDMTLR PVVVDTTLR	DGSHAIRHQY DGMHPKRHQM DGSHAHRHQY DGSHHKRHQF	TLDDVRATAR TLEQMKSTAC TVEEARATAQ TKDEVGATVA	ALDKAKVDSI GLDAAGIPLI ALDEAGVYAI ALDAAGVPV	EEE	48 47 50 46
DmpG Bphl TTHB246 HsaF	☆☆ VAHGDGLQG VTHGDGLGG VSHGDGLGG VTHGDGLGG	O SSFNYGFGRH SSVNYGFPAH SSLQYGFSRT SSFNYGFSKT	T D L E Y I E A V A S D E E Y L G A V I D E M E L I R A V R P E Q E L I K L A A	GEISH <mark>AQIA</mark> T PLMKQAKVSA ETVRRAKVAA ATAKEARIAF	LLLPGIGSVH LLLPGIGTVE LLLPGIGTRK LMLPGVGTKD	D H E D	98 97 100 96
DmpG Bphl TTHB246 HsaF	L K N A Y Q A G A L K M A K D L G V L K E A V E A G I I K E A R D N G G	R V VR VATHCT NT I R VATHCT QMVR I ATQCT S I CR I ATHCT	O EADVSKQHIE EADVSEQHIT EADISEQHFG EADVSIQHFG	Y <mark>arnlg</mark> md <b>ty</b> Qsrklgldty Makemglea <mark>v</mark> Larelgle <mark>ty</mark>	O GFLMMSHMIP GFLMMSHMRP GFLMMAHTIA	A P P	148 147 150 146
DmpG Bphl TTHB246 HsaF	EKLAEQGKL EKLVSQALL EFLAEQARL DKLAAQARI	MESYGATCIY Mqgyganciy Megygadvvy Madaccqcvy	MADSGGAMSM VTDSAGYMLP IVDSAGAMLP VVDSAGALVL	NDIRDEMRAF DDVKAFLSAV EDAYARVKAL DGVADEVSAL	KAVLKPETQV RAALKPETEL KEALSR.AKV VAELGEDAQV	GGGG	198 197 199 196
DmpG BphI TTHB246 HsaF	☆ ☆ MHAHHNLSL FHGHHNLAM FHAHNNLGL FHGHENLGL	GVANSIVAVE GVANSIAAIE AIGNTLAALA GVANSVAAVR	E G C D R V D A S L A G A T R I D A A A A G A D W V D A T L A G A K Q I D G S C	A GM <mark>G A G A G N A</mark> A G L G A G A G N T R G Y G A G A G N A R R F <mark>G A G A G N</mark> A	PLEVFIAVAE PMEVFIAVCA PLEVLAAVLD PVEALIGVFD	R R K K	248 247 249 246
DmpG Bphl TTHB246 HsaF	LGWNHGTDL MGIETGVDV AGLNPGLDV IGVKTGIDF	YTLMDAADDI FKIQDVAEDL FKLLDAAEYV FDIADAAEDV	VRPLQDRPVR VVPIMDHVIR MGPILHFQPY VRPAMPAECL	V DR ET L GL GY I DR DSL T L GY P DR DSVA I GY L DR NAL I MGY	AGVYSSFLRH AGVYSSFLLF AGVYSTFLLH SGVYSSFLKH		298 297 299 296
DmpG Bphl TTHB246 HsaF	E I <mark>AAAK Y</mark> NL KRASAK YGV KR I GKELGV VRQAER YGV	KTLDILVELG PARDILVELG DPLAILLELG PASALLHRAG	H R R M V G G Q E D R R G M V G G Q E D R R Q A V A G Q E D Q R K L I G G Q E D	MIVDVALDLL MIEDTAMTMA WILRVALELK QLIDIALEIK	AAHKENR <mark>A</mark> Rergltltaa Ekeagalad. Reldsgaavt	н	345 346 347 346

**Fig. S7:** Sequence alignments of DmpG with three of its orthologues: BphI, TTHB246 and HsaF. Levels of conservation between the orthologues are indicated by shades of blue. For clarity, a blank space has been inserted after every 10 residues of DmpG. Yellow stars indicate residues thought to have a role in either the reaction mechanism or substrate positioning, green ovals indicate channel lining residues and red diamonds indicate proposed channel gating residues.

Additional figure to show the sequence alignments obtained for the dehydrogenase DmpF and its orthologues:

DmpF	M N Q K L K V A I I	GSGNIGTDLM	IKVLRNAKYL	EMGAMVGIDA	A S D G L A R A Q R	50
BphJ	M T K K I K C A L I	GPGNIGTDLL	AK.LQRSPVL	EPIWMVGIDP	E S D G L K R A R E	49
TTHB247	M S E R V K V A I L	GSGNIGTDLM	YKLLKNPGHM	ELVAVVGIDP	K S E G L A R A R A	50
HsaG	M P S K A K V A I V	GSGNI <mark>S</mark> TDLL	YKLLRSEW.L	EPRWMVGIDP	E S D G L A R A A K	49
DmpF	MG V T T T Y A G V	EGLIKLPEFA	DIDFVFDATS	A S A H V Q N E A L	LRQAKPGIRL	100
BphJ	MG I K T T A D G V	DGLIPHMQAD	GVQIVFDATS	A Y V H A D N S R K	VNALGALM	97
TTHB247	L G L E A S H E G I	AYILERPEIK	IVFDATS	A K A H V R H A K L	LREAGKIA	95
HsaG	L G L E T T H E G V	DWLLAQPDKP	DLVFEATS	A Y V H R D A A P K	YAEAGIRA	95
DmpF BphJ TTHB247 HsaG	I DL TPAAIGP I DL TPAAIGP I DL TPAARGP I DL TPAAVGP	Y C V P V V N L E E F C V P T V N L K E Y V V P P V N L K E A V I P P A N L R E	HLGKLNVN HVGKGEMNVN HLDKDNVN HLDAPNVN	☆ MVTCGGQATI MVTCGGQATI LITCGGQATI MITCGGQATI	PMVAAVSRVA PMVAAVSRVQ PLVYAVHRVA PIVYAVSRIV	148 147 143 143
DmpF BphJ TTHB247 HsaG	KVHYAEIVAS PVAYGEIVAT PVLYAEMVST EVPYAEIVAS	ISSKSAGPGT VSSKSAGPGT VASRSAGPGT VASVSAGPGT	RANIDEFTET RKNIDEFTRT RQNIDEFTFT RANIDEFTKT	T S K A I E V I G G T A G A V E K V G G T A R G L E A I G G T A R G V Q T I G G	♦0 A A K G K A I I I M A K K G K A I I I L A K K G K A I I I L A A R G K A I I I L	198 197 193 193
DmpF	NPAEPPLIMR	DTVY.VLSAA	A DQAAVAA	SVAEMVQAVQ	A Y Y P G Y R L KQ	245
BphJ	NPAEPPLIMR	DTVHCLLESE	P DQAKITE	SIHAMIKEVQ	K Y Y P G Y K L V N	245
TTHB247	NPAEPPILMT	NTVRC.IPED	EGFDREAVVA	SVRAMEREVQ	A Y Y P G Y R L KA	242
HsaG	NPADPPMIMR	DTIFCAIPTD	A DREAIAA	SIHDVVKEVQ	T <mark>Y Y P G Y R L</mark> L N	241
DmpF	Q V Q F D V I P E S	APLNIPGLGR	FSGLKT <mark>SVFL</mark>	EVEGAAHYLP	A YAGNLDIMT	295
BphJ	G P V F D		GLRVSVYL	EVEGLGDYLP	K YAGNLDIMT	278
TTHB247	D P V F E R L P		TV <mark>VS</mark> MLL	EVEGAGDYLP	K YAGNLDIMT	283
HsaG	E P Q F D E P S		AL <mark>V</mark> TTFV	EVEGAGDYLP	P YAGNLDIMT	282
DmpF BphJ TTHB247 HsaG	SAA LATAERM AAA ARTAEMF ASARRVGEVF AAATKVGEEI	AQSML AEEILAGQLT AQHLLGKPVE AKETL	NA. 312 LQPVHA. 304 EVVA 307 .VVGGAR 303			

**Fig. S8:** Sequence alignments of DmpF with three of its orthologues: BphJ, TTHB247 and HsaG. Levels of conservation between the orthologues are indicated by shades of blue. For clarity, a blank space has been inserted after every 10 residues of DmpF. Yellow stars indicate residues thought to have a role in either the reaction mechanism or substrate positioning, green ovals indicate channel lining residues and red diamonds indicate proposed channel gating residues. The pink triangle indicates IIe159 which was mutated in this MD study.